Arbeitspapier Nr. 130

Proceedings of the Workshop
Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages

Josef Bayer & Yvonne Viesel (eds.)
Proceedings of the Workshop
CLAUSE TYPING AND THE SYNTAX-TO-DISCOURSE RELATION IN HEAD-FINAL LANGUAGES

JOSEF BAYER & YVONNE VIESEL (eds.)

Konstanz
November 2019

http://kops.uni-konstanz.de/handle/123456789/45/browse?authority=uniknseries1&type=konstanzseries
# Table of contents

Josef Bayer & Yvonne Viesel  
*Introduction* ................................................................. 1

I. Dravidian

K. A. Jayaseelan  
*Three clause-final particles and the syntax of clausal complementation in Dravidian* ........ 7

Rahul Balusu  
*Fine tuning the Dravidian left periphery: The three 'complementizers' in Telugu* .......... 13

R. Amritavalli & Annu Kurian Mathew  
*Discourse-driven scrambling to the peripheries in child Tamil* ................................ 39

II. Indo-Aryan

Probal Dasgupta  
*Clause particles and cleft sentences in Bangla: Some preliminary generalizations* ....... 51

III. Japanese

Yoshio Endo  
*Exploring right/left peripheries: Expressive meanings in questions* .......................... 69

Norio Nasu  
*Adverb-predicate agreement in Japanese and structural reduction* ............................... 91

IV. Turkish (and German)

Tamer Akan & Katharina Hartmann  
*SOV-X: Syntactic and pragmatic constraints of the postverbal domain in Turkish* ........ 123

Jaklin Kornfilt  
*A predicate-final constraint for head-final languages* .................................................. 145

V. Uralic

Katalin É. Kiss  
*Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive* ..... 163

VI. Basque

Sergio Monforte  
*What microvariation can show us: An analysis of Basque DP* ote ................................. 175
Introduction

Josef Bayer & Yvonne Viesel
josef.bayer@uni-konstanz.de
yvonne.viesel@uni-konstanz.de

This issue of the Arbeitspapiere des Fachbereichs Linguistik of the University of Konstanz contains ten selected papers from the workshop Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages that was held at Schloss Freudental in spring 2018. The aim of this workshop was to bring together linguists who work on head-final languages with a focus on the upper clausal periphery and aspects of at-issue and non-at-issue meaning. The title was inspired by work that has been carried out with the help of grant BA 1178/9-1, given to Josef Bayer by the German Research Foundation (DFG) between April 2013 and December 2018.

Many details have been clarified in the recent past about clause typing and the syntax-to-discourse relation. Nevertheless, mainstream theorizing in this area continues to be dominated by views that emanate from generalizations about head-initial languages, for which the split-CP analysis has proposed a richly organized left clausal periphery. Strictly head-final languages show clause-initial topic constructions but nothing like a more articulated left periphery. Instead, they show a more or less richly organized right periphery as expected from Baker’s Mirror Principle. Nevertheless, many head-final languages have typologically “deviant” complement clauses with initial instead of final complementizers and complement clauses in post-verbal position. Hindi is a widely known example. The right periphery differs significantly from the left periphery by not giving rise to specifiers. If a final complementizer has a specifier, this may be the entire TP that has been raised. The concept of “complementizer” turns out to be non-uniform and occasionally not easily distinguishable from the concept of “question particle”, “discourse particle” etc. Many head-final languages show a mixed picture of clause-medial and clause-final discourse and focus particles. Right-peripheral particles have scope over the clause, but what about particles that combine with sub-sentential constituents?

Another important issue is displacement. Even the strictest head-final languages – take for instance Japanese – have the option of displacing constituents to the post-verbal domain. Theoretical proposals are so far highly heterogeneous: rightward scrambling, rightward movement to (an otherwise unattested) clause-final specifier, leftward movement followed by heavy remnant movement, copying and eliding, prosodic restructuring. Given that displacement is usually not arbitrary, what is its motivation, and what are its semantic or pragmatic effects? Does the right periphery play a role in information structure? Could it be an unusual topic position? Or some kind of “anti-focus” position? Displacement to the right periphery is constrained in various ways. Heavy XP-shift can be found next to the displacement of light elements. The categories that can be affected are far from uniform across head-final languages.

Issues of the head-final organization of syntax were discussed with a focus on clause type and the division of clause types into more fine-grained distinctions that give rise to a variety of illocutionary meanings. The interrogative type, to take a prominent example, appears next to its

---

1 The series, formerly Arbeitspapiere des Fachbereichs Sprachwissenschaft ‘Working Papers of the Department of Language Science’, changed its title because the department was renamed.
association with the standard information-seeking impact, in sub-types of "special" or "non-standard questions" (Hans-Georg Obenauer): rhetorical, surprise, disapproval, reproach, exclamative, "aggressively" non-D-linked, the hell, why-like what etc. Clause-typing occurs either on the left or on the right edge of a sentence. The fine-tuning, however, may be controlled from other sites in the clause, e.g. from clause-medial position as in German von wem könnte der Dieb den Reisepass schon geklaut haben? (from who could the thief the passport SCHON stolen have), which is a rhetorical question due to the clause-medial particle schon ('already'). Work on interrogatives was only a starting point, however, as the workshop was thematically open.

Head-final languages offer a rich source of insights into the space of parametric variation. They force linguists to rethink clausal architecture, the share between syntax and prosody, the role of information structure and as a consequence certainly also the appropriateness of theoretical models that are on the market. In order to highlight the typological character of the workshop, the contributions in these working papers are ordered according to languages and/or language families. The goals of the workshop have nevertheless been theoretical. We hope the theoretical impact on some of the pertinent questions will become visible.

The first section starts with three contributions about Dravidian languages. According to the topic of the workshop, a natural kind of phenomenon is the character and role of complementizers or complementizer-like particles in head-final languages. Two papers were related to this topic.

In his article Three clause-final particles and the syntax of clausal complementation in Dravidian, K. A. Jayaseelan discusses the role of the question particle -oo, the complementizer ennǝ, and the relativizer -a, which occur in a fixed order in Malayalam, in case they co-occur. He argues that this order can be generated only if we postulate that the complementizer, which is a quotative element derived from the verb 'say', still retains its verbal syntax and projects its own clause. The relativizer -a can then be in the C domain of the clause projected by ennǝ, and the question particle -oo can be in the C domain of the CP complement of ennǝ. A surprising consequence of this analysis is that every embedded finite clause in Dravidian – the 'ennǝ + clause' structure – is in fact bi-clausal.

Rahul Balusu's article Fine tuning the Dravidian left periphery: The three 'complementizers' in Telugu picks up on this, now with a focus on the related Dravidian language Telugu. He investigates in detail three left-peripheral morphemes that have been considered at various places in the previous literature as instances of complementizers. According to Balusu, none of these morphemes are typical complementizers. The linearly first left-peripheral morpheme -aa has all the signature properties of a polar question particle and is in many respects similar to its Hindi counterpart kyaa. The second left-peripheral morpheme, -oo, delimits the scope of questions in Telugu. This he attributes to its location in the Spec of CP, where it is base-generated, and to its semantics, which is essential for interrogative semantics, thus explaining scope delimitations. The third left-peripheral morpheme, the quotative complementizer ani, is analysed as being syntactically and semantically true to its source, a verbum dicendi, the verb say, and its complementizer nature as arising only due to its not putting forth its extended projection (in the spirit of Grimshaw 2005) and instead being merged into the matrix clausal spine at various levels.

The third contribution in this section turns to yet another Dravidian language, namely Tamil. In their article Discourse-driven scrambling to the peripheries in child Tamil, R. Amritavalli and Annu Kurian Mathew argue that the SOV-language Tamil has a pre-verbal focus and post-verbal topic position. A subject wh-word must occur in focus, and not in a topic or in a canonical S(subject) position. This leads to the distribution: *S

Introduction
children 26-29 months of age are shown to obey these word order restrictions. The authors argue that child scrambling in Tamil moves arguments to criterial positions to check topic/focus features. A possible generalization with Japanese is suggested. A non-focus account of wh- is briefly critiqued.

The second section, on Indo-Aryan, continues with the article *Clause particles and cleft sentences in Bangla: Some preliminary generalizations* by Probal Dasgupta. Intimacy-oriented discourse particles (DiPs), called Modul[ator]s in the Bangla syntax literature normally follow a finite verb or a compact wh-phrase. In his article, Dasgupta surveys interactions between a Modul and Zero Copula Construction (ZCC) in three subtypes of ZCC. He extends the discussion to other contexts now diagnosable as ZCCs – sentences in which a post-verbal constituent hosts either a Modul or some other DiP. He argues that certain sentences with these properties instantiate cleft constructions whose properties are explored here in the context of the study of DiP elements. Some preliminary generalizations are proposed.

Section 3 contains two contributions on Japanese. The phenomenon of DiPs, which was introduced in Dasgupta’s article, plays a role in the first article here, as well as in Sergio Monforte’s article in Section 6, which concludes this volume. Yoshio Endo’s article *Exploring right/left peripheries: Expressive meanings in questions* discusses non-standard questions in Japanese such as rhetorical, surprise, disapproval, exclamative, etc. (Obenauer 2006, Bayer and Obenauer 2011, Bayer 2018) within the framework of the cartography of syntactic structures. After introducing the basic ideas of the cartographic approach, Endo first examines the expressive meanings of some wh-expressions asking for reasons such as what…for, how come, etc. familiar from languages such as English, German, etc. He then turns to the main topic of examining various sentence final particles in the right periphery of the Japanese sentence to show how they contribute to creating expressive meanings in questions. Methodologically, he does this by looking at translations of Peanuts comics. Endo draws comparisons with German, where corresponding particles are placed in clause-medial position, and he speculates about the absence of similar particles in English.

The article by Norio Nasu, *Adverb-predicate agreement in Japanese and structural reduction*, turns to the related topic of sentence adverbs (S-adverbs). In cartographic work, S-adverbs have a high position in the adverb hierarchy. Nasu shows that in Japanese, S-adverbs occur with a particular inflectional form of a predicate. He argues that this phenomenon is a manifestation of the agree relation between the adverb and a functional head. An agree-based analysis correctly predicts that an S-adverb can occur in more than one position as long as it is able to c-command the functional head it agrees with. It also accounts for restrictions on the co-occurrence of more than one S-adverb in a single clause. In Japanese, an epistemic adverb cannot precede an evidential adverb. The illegitimacy of this order is reduced to an intervention effect arising from agree. Nasu’s analysis predicts that some S-adverbs in Japanese can occur at the edge of more than one functional projection as long as they enter an agree relation with the appropriate functional head. In this respect, the distribution of Japanese S-adverbs presents a departure from a principal assumption of the cartographic approach, i.e. a constituent appearing on the clausal left periphery is in a one-to-one spec-head relation with the appropriate functional head.

The two contributions that appear in Section 4 discuss mainly the head-final language Turkish but also draw comparisons with the partially head-final language German. The article by Tamer Akan and Katharina Hartmann, *SOV-X: Syntactic and pragmatic constraints of the postverbal domain in Turkish*, sets out to develop a novel syntactic account for the postverbal domain in Turkish, which establishes a tight connection between syntactic and information-structural (IS) properties of the language. The authors first analyze the properties of the Turkish postnominal domain in comparison to the SOV-language German. Turkish is much less restricted than
Akan and Hartmann argue that Turkish provides evidence for low and high information-structural projections above TP. They assume different IS-projections for each information-structural status, hence, in this order, topic, focus and discourse-anaphoric phrases. The verb is always attracted to the head of the focus phrase. Postverbal constituents are therefore derived by leftward movement of given constituents to discourse-anaphoric projections and verbal movement to the higher focus head.

While Akan and Hartmann focus on extraposition to the right clausal periphery, Jaklin Kornfilt's contribution A predicate-final constraint for head-final languages presents a general point concerning head-finality by using Turkish as a reference language. We know that extraposition is forbidden if it disrupts head-adjacency with the immediately dominating structure as in *als wir Klaus darum gebeten darum haben (as we Klaus asked for it have). Kornfilt proposes a constraint for head-final languages, a Predicate-Final Constraint (PFC), such that, even in word-order free head-final languages, predicates must be clause-final: clauses must strictly represent the head-final property of the language. This constraint is parameterized such that it is absolute for some head-final languages, but limited to embedded clauses in others (e.g. Turkish). A further weakening of the constraint in some languages, e.g. Turkish, is that the constraint holds only when the clause-final predicate is followed by material in the higher clause. Kornfilt discusses similarities between how the PFC works in German and in Turkish as well as certain differences. Her work is centrally placed in the theoretical discussions that have emerged from the so-called Final-over-Final Constraint (FOFC) and attempts to derive it.

Sections 5 and 6 contain papers on Uralic and on Basque, respectively. Katalin É. Kiss observes in her contribution Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive that in the Ob-Ugric sentence, subject agreement and object agreement also encode the topic function of the subject and the object, respectively. A [-topic] subject candidate has to be demoted, and – owing to the EPP feature of AgrS – a [+topic] internal argument must be promoted to subject. Subject demotion and internal argument promotion are realized via passivization. Subject demotion can involve subjects of transitive, unergative and unaccusative verbs alike, and internal argument promotion can target objects, oblique internal arguments, and even adjuncts. Consequently, NP-movement is not a case-driven operation; it is triggered by the interplay of the [+topic] feature of AgrS in need of checking, and the EPP. In transitive clauses with a [+topic] subject and a [+to] object, the object undergoes NP movement to Spec of AgrOP, a secondary object position. In Eastern Mansi, the object moved to Spec of AgrO and entering an agreement relation with AgrO is also assigned accusative case. Focal objects remain in the VP caseless. These facts indicate that object licensing by the verb under government, and case assignment by AgrO via specifier–head agreement are separate processes. A further discourse-motivated property of Ob-Ugric is the optional replacement of the nominative case of subjects of active verbs functioning as recurring topics with an oblique case. É. Kiss's work agrees with Amritavalli and Kurian Mathew in finding a similar motivation for NP-movement, but differs with respect to the function of focus in word order.

Finally, Sergio Monforte's article What microvariation can show us: An analysis of Basque DP 'ote' is a case study in which it is shown that syntactic microvariation offers new insights into a deeper study of phenomena discussed in other languages from distinct research approaches. Dialectal data related to Basque discourse particles, which has received little attention so far, prove that DiPs can function both as syntactic heads and as weak adverbs in Eastern Basque (cf. Coniglio 2008). Configurations formed by wh-words and DiPs are found in North-Eastern Basque. Their analysis leads to the conclusion that they constitute a single unit based on
syntactic and prosodic data. Monforte presents novel data concerning the behavior of the particle *ote* as a sentence-final particle conveying an intersubjective interpretation in some varieties of Basque, unlike its standard behavior in the TP-domain.

**Acknowledgments**

We are greatly indebted to the DFG for grant BA 1178/9-1, within which this workshop could be organized. The following colleagues and friends have generously supported us with the reviewing of abstracts: Valentina Bianchi, Daniel Büring, Vicki Carstens, Lisa Cheng, Guglielmo Cinque, Silvio Cruschina, Roberta D'Alessandro, Probal Dasgupta, Mark de Vries, Urtzi Etxeberria, Ricardo Etxepare, Caroline Féry, Hubert Haider, Fabian Heck, Caroline Heycock, Virginia Hill, Daniel Hole, Julia Horvath, Joachim Jacobs, Georg Kaiser, Beste Kamali, Jason Kandybowicz, Shin-Sook Kim, Hilda Koopman, Shigeru Miyagawa, Gereon Müller, Keiko Murasugi, Pieter Muysken, Ad Neeleman, Masayuki Oishi, Sumru Özsoy, Markus Pöchtrager, Luigi Rizzi, Michael Rochemont (†), Johan Rooryck, Mamoru Saito, Peter Sells, Andrew Simpson, Stavros Skopeteas, K. V. Subbarao, Sandhya Sundaresan, Takashi Toyoshima, Myriam Uribe-Etxebarria, Enric Vallduví, Elisabeth Verhoeven, Helmut Weiß, Kyoko Yamakoshi, James Yoon, Jochen Zeller, Jan-Wouter Zwart.

Thanks also to Annette Wilz of the university’s financial administration. With her, calculations, payments, bookings, refundings etc. were in the best hands imaginable.

Finally, a big thanks to Susanne Trissler for her help in assembling the contributions into the format of the *Arbeitspapiere des Fachbereichs Linguistik*.

**References**


This paper is about three clause-peripheral particles in the Dravidian languages, namely the relativizer, the complementizer and the question particle. The attempt to accommodate them in the C domain leads us to propose an unorthodox analysis of clausal complementation in Dravidian.

1 The relativizer -a

The Dravidian languages have a particle -a that occurs at the end of a relative clause which is standardly analysed as a 'relativizer' (Malayalam data):

(1) a. [ñaan ___ kaND-a] kuTTi
   I (Nom)  saw-REL  child
   'the) child that I saw'

It has often been assumed that the gap in the relative clause is generated by this -a moving from the gap position to the clausal periphery; the underlying assumption would be that the -a is like a relative pronoun (see Jayaseelan 1991 for this suggestion).

There is no doubt that there is a movement in this construction, because when the relativization is long-distance it shows island effects, a fact noted in early work by K. P. Mohanan and the present author; cf.

(2) a. [ñaan ___ eDuttu enn-a] ellaawarum wiśwasikkunn-a] paTam
   I (Nom) took COMP everyone believes-REL picture
   '(the) picture that everyone believes that I took'

b. * [ñaan ___ eDuttu enn-a apawaadam] ellaawarum wiśwasikkunn-a] paTam
   I (Nom) took COMP-REL allegation everyone believes-REL picture
   '(the) picture that everyone believes the allegation that I took'

While (2a) is fine, (2b) shows a complex noun phrase effect.

2 The relativizer and the complementizer

But this -a also shows up at the end of a noun complement clause where it co-occurs with the Dravidian quotative complementizer enna:

(3) [John wannu enn-a] waarrta
    John came  QUOT-REL  news
    '(the) news that John has come'

* I wish to thank Katalin Kiss for helpful comments on this paper. I also wish to thank the audience at the conference for an insightful discussion.
The presence of the relativizer here seems to support the Kaynean claim (Kayne 2010) that all complements of nouns are underlyingly relativization structures. However note that there is a problem here for the movement proposal about -a: the noun complement clause contains no gap for -a to have moved from.

There is also a puzzle here: where in the clausal periphery are the quotative and the relativizer accommodated? Why is the relativizer higher than the quotative (as is indicated by the relativizer coming after the quotative, given the head-final order)? One could perhaps say – assuming Rizzi’s (1997) analysis of the clausal periphery – that the quotative is in the position of the English complementizer ‘that’ and therefore heads the Finiteness Phrase; and that the relativizer is in some higher projection, possibly ForceP.

3 The question particle

However this analysis is made untenable when we consider a noun complement clause which is interrogative:

(4) [John wannu-oo enn-a] coodyam ‘(the) question whether John has come’
     John came-Q QUOT-REL question

There are three elements to be accommodated in the C domain (Rizzi’s “left periphery”) here: the question particle -oo, the quotative ennǝ, and the relativizer -a. The natural place for the question particle is ForceP, since it signifies the interrogative force of the clause. So the question arises: Are the quotative and the relativizer above ForceP? Is ForceP very low in the Dravidian C domain? (This was indeed the conclusion arrived at in Jayaseelan (2008).)

Rizzi (2001) postulates a position Int(errogative) which is immediately below ForceP, to accommodate Italian se (‘if’) that appears in embedded questions. If one were to say that the Malayalam question particle -oo is in Int, one could then say that ennǝ is in ForceP. But still, where do we place the relativizer -a? Are we forced to abandon the claim that ForceP is the highest projection in the C domain?

But we now show that there is a completely different analysis possible which avoids the need to tinker with the universal functional sequence in the C domain; we call it the “clausal quotative analysis.”

4 Some further facts about the complementizer

A quotative complementizer (as the name implies) is derived from the ‘say’-verb; the Dravidian ennǝ is the perfective form of the verbal root enr- ‘say’, which is obsolete in Malayalam but is still a functioning verb in Tamil. The current wisdom is that ennǝ has been completely reanalysed as a complementizer; it is generated as the head of CP, and takes a clausal complement. But ennǝ can – and often does – take a simple nominal expression as its complement; e.g.

(5) meSii “grrr” ennǝ śabdiccu ‘The machine made the sound “grrr”.
     machine QUOT sounded

1 In Tamil, when the matrix verb is enr- (‘say’), the complementizer can in fact be omitted, cf.

(i) awan [(pro) war-eeq] enr-aan ‘He said (he) would come.’
     he (Nom.) come-1stP.Sg. say-3rdP.Sg.
In (5), the complement of *enna* is just a representation of a sound; there is no C domain here to generate *enna* in. Even the noun complement construction can have a simple nominal as the complement of *enna*, cf.

(6)  “kaakka” enn-a waakka  ‘(the) word “crow”’

crow QUOT-REL word

What such data show is that *enna* is still a ‘say’-verb, which can take as its complement anything that can be ‘said’, i.e. uttered; e.g. a sound (‘Say “Boo!”’), or a word (‘Say “crow”’), or a clause (‘Say “Mary is pregnant”’). Though bleached in meaning – in (5), e.g., the machine doesn’t ‘say’ anything – *enna* retains its verbal syntax.² ³

5 Clausal complementation in Dravidian

What we have said has serious implications for the syntax of clausal complementation in Dravidian. When ‘say’ takes an object complement – irrespective of whether it is a sound, word, or clause – it goes without saying that it is outside that complement. Now consider a sentence where *enna* takes a finite clause as its complement:

(7)  John [ Mary wannu ennǝ ] paRaññu  ‘John said that Mary has come.’
    John Mary came QUOT said

We can now see that the correct analysis of (7) is that *enna* is outside its CP complement; it is not in the C domain of the embedded clause at all. The ‘say’-verb projects its own c-clause, which is nonfinite but can have its own C domain. The structure we postulate for (7) is (8) (abstracting away from word order):⁴

---

² Do we wish to entertain a “squishy” account of *enna*, saying that it has been reanalyzed as a complementizer when it takes a clausal complement, but that it is still a ‘say’-verb when it takes a nominal expression as its complement? Such a “two *enna’s” analysis would be unsatisfactory for several reasons. First of all, note that *enna* occurs indifferently with assertive and interrogative matrix verbs, showing an insensitiveness to the matrix predicate which is unexpected in a complement but is quite in keeping with an adjunct:

(i)  John [ Mary wannu ennǝ ] paRaññu  ‘John said that Mary has come.’
    John Mary came QUOT said

(ii) John [ Mary wannu-oo ennǝ ] coodiccu  ‘John asked whether Mary has come.’
    John Mary came-Q QUOT asked

Again, where do we generate *enna* in the C domain? Suppose we generate it as the head of Finiteness Phrase. Then, in a sentence like (ii) above (or like (4)), the question particle -oo – and by implication ForceP – will have to be below the Finiteness Phrase; and a “low ForceP” will make Dravidian a typological oddity.

³ Readers unfamiliar with Dravidian languages might ask: Is *enna* confined to the complements of ‘verbs of saying’? It is not. The matrix verb can be any verb that takes a clausal complement, cf.

(i)  Mary [ John kaLLan aaNǝ ennǝ ] wiśwasiccu/ samsayiccucu/aarooopiccu  ‘Mary believed/suspected/alleged that John is a thief.’
    Mary John thief is QUOT believed/ suspected/ alleged

But there is one restriction that needs to be noted on what *enna* can take as its complement: a nonfinite clause is disallowed, cf. (ii). (iii) is fine without *enna*.

    John Bill-go-INF QUOT said

⁴ We represent (8) in terms of antisymmetry. Deriving the surface order is straightforward: the complements of heads show up to their left. (See Jayaseelan 2010 for a proposal about why this word order obtains in head-final languages.)
Three clause-final particles and the syntax of clausal complementation in Dravidian

(8) John paRaññu [CP [IP PRO enne [CP [IP Mary wannu ]]]]

This literally translates as ‘John said, having said Mary came.’ (We may compare this with Dakhini Urdu locutions like: woh nahii aayegaa bolke bola, lit. ‘He said having said (he) will not come.’)

Now in the interrogative noun complement construction illustrated in (4), the relativizer -a is in the C domain of the clause that enne projects. On the other hand, the question particle -oo is in the C domain of enne’s object complement, which is a CP. That is, we have two distinct C domains here. Therefore, the ForceP that the question particle is generated in, can be the highest projection in its local C domain – and we don’t have to revise Rizzi’s picture of the left periphery. The structure we postulate for (4) is (9):

(9) coodyam [CP -a [IP PRO enne [CP -oo [IP John wannu ]]]]

6 A non-movement analysis of the relativizer -a

Returning to the relativizer -a, we already pointed out that the ‘movement-to-COMP’ analysis must be given up because there is no gap that the -a could have been moved from in the noun complement construction, cf. (3), (4), and (6). Therefore it must be generated in situ. Plausibly, it is in the position of ‘that’ in the following implementation of the raising analysis of relativization (cf. Kayne 1994:§ 8.2). (It cannot correspond to ‘the’, because the definite article in Dravidian is null.)

(10) the [CP _____ that [IP I read book ]]

We can now say that the island effects in the relative clause construction illustrated in (2) must be owing to the movement of the head noun (which is illustrated in (10)).

7 Clausal embedding in Dravidian

Our proposed analysis makes clausal embedding in Dravidian, i.e. the ‘complementizer + complement’ structure, a nonfinite adjunct of the matrix verb. We note a traditional claim of Dravidian linguistics that all embeddings in these languages are nonfinite (Steever 1988:5). In fact, Caldwell (1856 [1913]) made the strong claim that Dravidian languages allow only one finite verb per sentence. The fact that enne can embed a finite clause was excused on the plea that ‘anything can be quoted.’

But it is not correct to treat the complement of enne as being within quotation marks. In Jayaseelan (1991), I noted that in a sentence like (11) (Tamil data):

(11) awan [naan nallawan enru ] co-nn-aan
    he I (Nom) good person QUOT say-PAST-3ms

‘He said that I am a good person.’

naan ‘I’ can be interpreted as referring either to awan ‘he’ (direct discourse), or to the speaker of the sentence (indirect discourse). Also, enne + complement does not quite behave like an

5 We can bolster this suggestion by noting that the -a could be a shortened form of the Dravidian distal demonstrative aa (‘that’).
adjunct as regards extraction. While the Dravidian languages do not have wh-movement out of embedded clauses, they do have long-distance relativization, which argues that the head noun can be moved out of the complement of enna. And in the cleft construction, phrases can be extracted from the complement of enna to the focus position of the cleft. Cf.

(12) Mary-ye aaNa [John [Bill kaNDu enna] paRaññ-ata]
Mary-ACC be.PRES John Bill saw QUOT say.PERF-NOM
'It is Mary that John said that Bill saw.'

(Since clefting involves relativization – see Kayne 1994, Jayaseelan & Amritavalli 2005 – we may not be dealing with two separate facts here.) The fact that the adjunct island condition does not seem to apply to the complement of enna would suggest that some reanalysis has happened.

A proposal about the opacity of adjuncts is that adjuncts are late insertions: they are merged acyclically (or postcyclically); see Lebeaux (1988), Uriagereka (1999), Stepanov (2001, 2007). But it is also well-known that not all adjuncts are equally opaque (Cattell 1976): instrumental adverbials are transparent for subextraction (cf. What did you break the glass with?), so are infinitival adjuncts of purpose (cf. Who did she come here to look for?). On the other hand, time/place adverbials, and adverbials of cause are opaque (cf. * Who did you see a snake near?: * Who did you arrive after?: * Who did she cry, because you hit?). It would seem that the lateness of the insertion is determined by how close the adjunct is to the thematic requirements of the matrix verb. In our case, all that we need to say (then) is that the clause projected by enna (including the complement of enna) is not a late insertion, because functionally it is the internal argument of the matrix verb.

6 A remaining problem for our analysis – felt by many commentators – is: how does the matrix verb satisfy its selectional and subcategorizational requirements, if it gets a non-finite adjunct after it, and not a DP or CP as its direct object? Or, to put it differently, how does the matrix verb "skip over" the intervening clause projected by enna to recover its semantically mandated direct object from the complement of enna? This problem, we wish to suggest, should be treated as an entry point to a more adequate understanding of selectional restrictions. The problem is already present in our account of other data: for example, after the adoption of the DP-analysis of some erstwhile "noun phrases", how does the interpretation of phrases like "* eat [ the [ idea ]]" "eat [ the [ apple ]]" proceed? How does the verb "skip over" the intervening determiner to access the head noun, in order to determine that the first phrase is semantically ill-formed and the second phrase is not? Obviously, selection is not strictly local. But then, what are the types of elements that can be "skipped over"? In the case of enna, the fact that it is "bleached" in meaning may have something to do with its being non-opaque to selection.

References


Fine tuning the Dravidian left periphery:
The three ‘complementizers’ in Telugu

Rahul Balusu
kodiguddu@gmail.com

1 Introduction*

The goal of this paper is to investigate the distribution and properties of three ‘complementizers’ in Telugu, with occasional forays into Kannada and Malayalam—what turn out to be the question particles -oo and -aa, and the quotative complementizer ani (anta in Kannada and enn@ in Malayalam). While the quotative complementizer is a more likely candidate for being called a complementizer along traditional lines, the two question particles under discussion, while they have been labelled interrogative complementizers at various places in the literature on the Dravidian left periphery, are better analyzed as proper question particles, with specialized function and semantics.

The particle -aa, present in Kannada, Telugu, & Tamil but absent in the standard variety of Malayalam shows a comparatively restricted and simple pattern, but is still intriguing—it is

*We would like to thank the audience and reviewers of HFL 2018, University of Konstanz for comments and discussion, that especially helped in making the data and analysis clearer in the wider context.
restricted normally to root contexts, and that too polar and alternative questions. We analyse the question particle -aa, across Dravidian, as a polar question particle, along the lines of Bhatt & Dayal (2018) in their analysis of Hindi kyaa, that occurs in the ForceP projection in the split CP, and thus is mostly restricted to matrix contexts. It also has a singleton-set restriction on its complement, a presupposition, relegateing it’s occurrence to polar questions only. Alternative questions with -aa are essentially disjunctions of polar questions that are clausal disjuncts (larger than CP). Seemingly sub-clausal occurrences of -aa are also analysed as clausal attachments, with various mechanisms like elision or movement masking this high attachment and making it appear sub-clausal.

We build a composite picture of the quotative complementizer (QC) in Dravidian by examining its role in various left-peripheral phenomena like embedded questions and its particular manifestation in various constructions like noun complement clauses, manner adverbials, rationale clauses, with naming verbs, small clauses, and non-finite embedding, among others. The QC we conclude is instantiated at the very edge of the clause it subordinates, outside the usual left periphery, comes with its own set of projections, and is the light verb say which does not extend its projection unless needed to. It adjoins to the matrix spine at various heights (at the vP level it gets a θ-role, and thus argument properties) when it does extend its projection, and like a verb selects clauses of various sizes (CP, TP, small clause). It inherits all its properties from its source, the verb say, especially as to what projections it comes as a part of. We find that the quotative complementizers across Dravidian vary in clause-type features –The Telugu quotative complementizer ari is marked [-INTERROG]; the Kannada and Malayalam quotative complementizers anta and enne are marked [+INTERROG].

The morpheme -oo plays many roles in Dravidian –most notably the disjunction marker, but also part of existential quantifiers and correlatives, and it has been quite well documented in the literature by Slade (2011), Jayaseelan (2012), Szabóicsi (2015), among others. We find that -oo shows an interesting pattern of distribution in matrix and embedded wh-clauses underrogative and responsive predicates, and in its interaction with the quotative complementizer anilanta. We take these properties as revealing evidence to build up a semantics and syntax of -oo as a question particle that is crucial to interrogative meaning formation. We analyse the particle -oo in Telugu as being base-generated in the Spec of CP, and as being essential for the interpretation of interrogative clauses because it is needed for turning the focus-semantic value of such sentences which otherwise do not have an ordinary semantic value into the ordinary semantic value.

2 The question particle -aa in Telugu

The particle -aa, unlike -oo, is much more restricted in its distribution and has only one role –a question particle; in Kannada, Telugu, Tamil, and Malabar Malayalam (Manthodi & Balusu 2018). It does not exist in Standard Malayalam, where the only question particle is -oo. The question particle -aa in these languages and dialects shows up only in polar/alternative questions, not wh-questions, and typically in matrix contexts, (1)\(^1\). All examples in the paper, unless otherwise specified, are from Telugu.

\(^1\)We will gloss this -aa as the Polar Question Particle (PQP), anticipating our analysis.
(1) a. ravi aaDeeDu
    Ravi read
    ‘Ravi read.’

b. eeemi aaDeeDu?
    what played
    ‘What did he play?’

c. aaDeeD-aa (leed-aa)?
    play-PQP not-PQP
    ‘Did he play (or not)?’

2.1 -aa in matrix contexts

Polar and alternative questions in Telugu are obligatorily marked morphosyntactically with the question particle -aa. In Dravidian languages in general, an overt morphosyntactic cue is needed (in the form of either -aa or -oo), unlike in languages like Hindi (and other Indo-Aryan languages like Bangla, Gujarati, Punjabi, etc) where polar questions are indicated prosodically, and only optionally a question particle, polar kyaa, occurs (Bhatt & Dayal 2018).

The location of -aa in polar questions and in alternative questions in Telugu is usually clause final, except in clefts, or less prevalently in seemingly sub-clausal placement of -aa. The clause final location of -aa leads to a wide focus interpretation, of the entire sentence. Whereas, in clefts and so-called sub-clausal -aa constructions, there is a narrow focus interpretation, of a part of the sentence. This is how information structural effects involving focus are achieved.

But even in clefts we show that -aa is clause-final, here the clause being the copular pivot clause, only it is lacking an overt copula. Finally, seemingly sub-clausal -aa is also an instance, we argue, either of gapping or reduction or movement, masking the clause-final attachment of -aa. Therefore there are no instances of real sub-clausal -aa.

2.1.1 -aa in polar questions

Polar questions surface with the question particle -aa, (2), in matrix clauses. Without the particle -aa, the polar question is ungrammatical (without any bias), (3).

(2) ravi pustakam cadiveeD-aa?
    Ravi book read-PQP
    ‘Did Ravi read the book?’

(3) *ravi pustakam cadiveeDu?
    Ravi book read
    ‘Intended: Did Ravi read the book?’

Constituent questions do not surface with any question particle, (4):

(4) a. ravi eeemi cadiveeDu?
    Ravi what read
    ‘What did Ravi read?’

b. ravi eppuDu vacceeDu?
    Ravi when came
    ‘When did Ravi come?’

A seemingly sub-clausal positioning of -aa is possible, (5). This leads to narrow focus, and a focus interpretation of the element that -aa attaches to. There can be boolean disjunction within the polar question, (6), and it takes scope under the question operator. There is no ambiguity in the sentence, and an alternative reading is simply not available.

(5) ravi book-aa cadiveeDu?
    Ravi book-PQP read
    ‘Is it the book that Ravi read?’

(6) ravi book-oo paper-oo cadiveeD-aa?
    Ravi book-DISJ paper-DISJ read-PQP
    ‘Did Ravi read the book or the paper?’ [Y/N]
The boolean disjunction can also be that of non-finite clauses, (7). How high up can the boolean disjunction be in finite clauses? It can go up to the MoodP, (8). It can also go up to the evidential marker (which is possibly in the TP or CP domain), (9). The -aa by itself, without the disjunction, occurs outside the evidential marker, on the verb, (10).

(7) pustakam cadavaDam-oo raayaDam-oo jarigind-aa?
book read-NOML-DISJ write-NOML-DISJ happened-PQP
‘Did the book’s reading or writing happen?’ [Y/N]

(8) ravi vell-oo uma vacc-oo -unDavacc-aa?
Ravi go-DISJ Uma come-DISJ -might-PQP
‘Might, Ravi have gone or Uma come?’ [Y/N]

(9) nuvvu velleev-oo atanu vacceeD-oo-anT-aa?
you went-DISJ he came-DISJ-EVID-PQP
‘Apparently, did you go or he come?’ [Y/N]

(10) ravi pustakam cadiveeD-anT-aa?
Ravi book read-EVID-PQP
‘Did Ravi apparently read the book?’
A boolean disjunction of two full clauses, TPs, in a polar question is also possible, (11).

(11) nuvvu velleev-oo atanu vacceeD-oo-n-aa
you went-DISJ he came-DISJ-PQP
‘Did you go or he come?’ [Y/N]

2.1.2 -aa in alternative questions

In alternative questions -aa occurs on both disjuncts, (12). The contrastive connective, leed-aa ‘not-if’ is head of J(unction)P in (12), and occurs optionally in disjunctions, in X-OR Y-OR (Z-OR...) tuple structures, similar to other such connectives cross-linguistically (Szabolcsi 2018).

(12) nuvvu velleev-aa (leed-aa) atanu vacceeD-aa?
you went-PQP not-if he came-PQP
‘Did you go or (else / if not) did he come?’

The same is true of polar alternative questions, each alternative surfaces with an -aa, (13)-(14).

(13) ravi aaDeeD-aa leed-aaaaDa-leed-aa?
Ravi play-PQP not-if play-not-PQP
‘Did Ravi play or (else / if not) not play?’

(14) ravi aaDeeD-aa leed-aa?
Ravi play-PQP not-PQP
‘Did Ravi play or not?’

So far we have seen only full clauses as disjuncts in alternative questions. Is it possible to have either or both parts of the alternative question as so-called sub-clausal disjuncts? This is
possible in Hindi with kyaa (Bhatt & Dayal 2018). It turns out that seemingly sub-clausal juncts are also possible in alternative questions in Telugu, (15)-(16).

(15) ravi coffee-aa tea-aa taageeDu? (16) ravi coffee taageeD-aa tea-aa?
Ravi coffee-PQP tea-PQP drank Ravi coffee drank-PQP tea-PQP
‘Is it coffee or is it tea that Ravi drank?’ ‘Did Ravi drink coffee, or tea?’

2.1.3 -aa in cleft structures

A case where -aa disjuncts seem to appear sub-clausally on the surface involves clefting, (17), where there is a cleft marker on the verb (CLM). But these are also in fact clausal disjuncts, with -aa suffixed to the copula, which becomes overt when negated, (18), of the pivot clause in a biclausal cleft structure, with pivot and cleft clauses (Jayaseelan & Amritavalli 2005).

(17) ravi coffee-aa tea-aa taag-indi?
Ravi coffee-PQP tea-PQP drank-CLM
‘Is it coffee or tea that Ravi drank?’

(18) ravi coffee kaad-aa tea kaad-aa taag-indi?
Ravi coffee EQ.NEG-PQP tea EQ.NEG-PQP drank-CLM
‘Is it not coffee or is it not tea that Ravi drank?’

As seen in (18), -aa attaches to the copula kaadu (which is the suppletive form of the negative equative copula, EQ.NEG), in line with the clausal (left-peripheral) attachment hypothesis we have been entertaining for this morpheme. Of course, this is also possible in polar cleft questions, (19)-(20). It is also possible to have disjunction in the cleft pivot, (21).

(19) ravi coffee-aa taag-indi? (20) ravi coffee kaad-aa taag-indi?
Ravi coffee-PQP drank-CLM Ravi coffee EQ.neg-PQP drank-CLM
‘Is it coffee that Ravi drank?’ ‘Is it not coffee that Ravi drank?’

(21) ravi coffee-oo tea-oo-n-aa taag-indi?
Ravi coffee-DISJ tea-DISJ-PQP drank-CLM
‘Is it Coffee or Tea that Ravi drank? [Y/N]

The cleft constructions are the preferred way to achieve information structural effects in polar and alternative questions, as the above examples demonstrate. The seemingly sub-clausal placement strategy to deliver information structural effects of focus and topic are rather less preferred.

2.2 -aa in embedded contexts

In embedded contexts, in both polar and alternative contexts, the preferred question particle is actually -oo, (95)-(96).

(22) neenu atanu cadiveeD-oo leed-oo aDigenu
I He read-DISJ not-DISJ asked ‘I asked if he had read or not.’
(23)  nuvvu velleev-oo atanu vacceevD-oo aDigeenu
    you went-DISJ he came-DISJ asked
    '(I) asked whether you went or he came.'

Embedded wh-questions also surface with the particle -oo, (24). We will discuss in detail the distribution and analysis of this -oo in a later section.

(24)  ravi eemi cadiveeD-oo aDigeenu
    Ravi what read-DISJ asked
    '(I) asked what Ravi read.'

In embedded contexts -aa is not acceptable under plain responsive, i.e. veridical predicates, (25), marginal with negated responsive, i.e. under non-veridical predicates, (26), and acceptable though less preferred to -oo under rogative predicates, (27)-(28).

(25)  *atanu tinnaaD-aa telusu
    He ate-PQP know
    'Intended: (I) know if he ate.'

(26)  ??atanu tinnaaD-aa naaku teliy-adu
    He ate-PQP I-DAT know-not
    'I don't know if he ate.'

(27)  ?atanu tinnaaD-aa aDigeenu
    He ate-PQP asked
    'I asked if he ate.'

(28)  atanu tinnaaD-aa aDugu!
    He ate-PQP ask-IMP
    'Ask if he ate!'

Thus, though -aa is not preferred in embedded polar/alternative questions, it is not altogether banned from embedded contexts. However, it shows a very interesting cline of acceptability based on the embedding predicate and other matrix operators like negation and the imperative.

2.3 Polar question particle signature of -aa

As seen up until now, the particle -aa in Telugu is restricted to polar and alternative questions. Seemingly sub-clausal attachment of -aa is possible like polar kyaa in Hindi. -aa also displays selectiveness in embedding, or quasi-subordination (Dayal & Grimshaw 2009). A summary of all the findings is given in (29). The cline of grammaticality in the table refers to the acceptability under veridical, non-veridical, rogative predicates and imperatives, as discussed in the previous sub-section.

(29)  | Polar questions       | Matrix  | Embedded |
     |                     | -aa     | -oo (*/??/?-aa) |
     | Alternative questions| -aa (obligatory on all disjuncts) | -oo (*/??/?-aa) |
     | Wh- questions        | —       | -oo         |
2.4 Analysis of the question particle -aa

2.4.1 An earlier account of -aa

Amritavalli (2013) analyses the -aa in Dravidian as a question-operator in the matrix clause, (30), and that -aa is covert in wh-questions, (31), using examples from Kannada.

(30) makkaLu ba-nd-ar-aa KANNADA (31) yeSTu jana sattaru aa
children come-pst-3pl-Q how-many people die.pst.3pl Q
‘Did the children come?’ ‘How many people died?’

Amritavalli (2013) also proposes that the -aa in embedded wh-clauses in Kannada is a covert interrogative complementizer that co-occurs with the quotative complementizer anta, (32), and that similarly the -aa in embedded polar questions in Kannada is an overt interrogative complementizer that co-occurs with the quotative complementizer, (33)².

(32) [fidanna yaaru baredaru-aa] anta] keeLide/kaNDu.hiDide KANNADA
this-ACC who wrote Q QC asked/discovered
‘(I) asked/discovered who wrote this.’

(33) tande [makkalu-aa] anta] keeLidar u KANNADA
father children come.pst.3pl-Q QC asked
‘The father asked if the children had come.’

Since aa-anta complements may be ambiguous between a matrix and an embedded question reading, (34), and -aa need not always scope below anta (following the surface order), Amritavalli (2013) infers that -aa can occur either as an interrogative complementizer in the embedded clause or as a question operator in the matrix clause (we don’t have access to reliable Kannada data to ascertain if this also the case in (32) and (33)).

BBC how-many people die.pst.3pl Q QC said
(i) . = ‘The BBC said how many people died.’ (with a declarative intonation)
(ii) ? = ‘How many people did the BBC said died?’ (with a question intonation)

2.4.2 Our analysis of -aa in Telugu

For the analysis of -aa as question-operator to go through, Amritavalli (2013) needs to posit a covert -aa in matrix wh-questions, and to analyse it as an interrogative complementizer in embedded contexts, with again a covert -aa in embedded wh-questions. Transposing this analysis into Telugu will again need a number of covert instances of -aa to make the question-operator analysis viable. This account will also not be able to account for the selective embedding of -aa under rogative vs. responsive predicates that we find in Telugu³.

²A reviewer asks if the quotative and the interrogative complementizer always co-occur in Kannada. We don’t have reliable access to Kannada data to answer this question.

³This selective cline of acceptability when embedded under various matrix predicates and operators is also found in Malabar Malayalam, as observed by Manthodi & Balusu (2018).
As shown in the previous section, what we find is that the -aa of Telugu parallels the distribution of polar *kyaa* in Hindi, as explicated in Bhatt & Dayal (2018), in some crucial respects—first and foremost, it is necessarily limited to polar and alternative questions, never seen in *wh*-questions. Second, it shows selectivity in where it is embedded, i.e., it is perfectly fine in **rogative**-imperatives, and ungrammatical under veridical-responsive predicates. Third, it has a so-called sub-clausal positioning, which leads to narrow focus on the constituent that -aa attaches to, leading to a narrowing of the question under discussion. This portends an analysis of the syntax and semantics of -aa along the lines of Bhatt & Dayal (2018), that can explain these properties—a morpheme residing higher up in the clausal spine than the question-operator, to explain its matrix predilection; and a morpheme that comes with a presupposition of a singleton propositional set complement, to explain its polar question restriction.

### 2.4.3 -aa in polar questions

As far as the matrix vs. embedded contrast in the distribution of -aa is concerned, it shows the characteristic properties of a root phenomenon. Therefore it should be located on the clausal spine above where normal embedding occurs. It should also be above the location where the interrogative vs. declarative split is determined, since it does not occur in declarative clauses. Following Bhatt & Dayal (2018) we take this position to be minimally the ForceP above C[+Q], as shown in (35).

\[
\text{(35) } \quad \text{ForceP} \leftarrow -aa \quad \text{CP} \quad C^0 \quad \text{TP} \\
\phi_Q \quad \text{‘Ravi ate’}
\]

Next, how do we explain the selectivity in embedding of -aa, its appearance in quasi-subordinated embedded polar questions but not otherwise? This is a larger pattern than just Telugu -aa, or Malabar Malayalam -aa, or even Hindi *kyaa* (Bhatt & Dayal 2018). It is also seen in embedded inversion in English (McCloskey 2006). Following these authors we analyse quasi-subordinated embedded polar questions as involving an extra CP layer, the ForceP layer, as shown in (36). Thus those subordinations that involve a ForceP like rogative predicates and non-veridical responsive predicates allow for -aa to be embedded under them, but those predicates that only take up to the interrogative-C layer like veridical-responsive predicates do not allow -aa to be embedded under them, as we saw in the previous section on the patterning of -aa in Telugu.

\[
\text{(36) } \quad \text{a. rogatives and non-veridical responsive: } [\text{forceP } [CP \ C^0_{+Q} \ [r \ p ]]] \quad \text{b. veridical responsiveness: } [CP \ C^0_{+Q} \ [r \ p ]]
\]

Finally, how do we explain the restriction of -aa to only polar and alternative questions, and its non-occurrence in *wh*-questions? This is the exact distribution of polar question par-
articles according to Bhatt & Dayal (2018), who propose all such particles to encode a presupposition of a singleton-set denoting complement (thus preventing them from occurring with wh-questions which denote a set of propositions). We thus follow them in proposing a similar presupposition for the Telugu -aa as shown in (37).

\[(37) \quad [-aa] = \lambda Q_{(st,t)} : \exists p \in Q[\forall q \in Q \rightarrow q = p].Q\]

Going by this lexical entry, since it takes a set of propositions, it cannot combine with declaratives. But since the set of propositions it takes is the singleton set, it cannot combine with wh-questions. Thus -aa’s distribution is restricted to polar questions. Then going by this denotation, when it occurs in alternative questions, it should also compose with only a singleton-set. This is the property we will turn to next.

2.4.4 -aa in alternative questions

The data in the previous sub-section has laid out that in alternative questions in Telugu, -aa must occur on each junct (unlike in Hindi which allows a single kyaa in an alternative question)\(^4\). This makes it clear that each -aa in an alternative question is composing with a polar question (i.e. with the introduction of each alternative in the alternative question), and together all the polar questions are disjoined to form an alternative question. This is also, surface appearance of single kyaa disregarding, the analysis proposed by Bhatt & Dayal (2018) for alternative questions in Hindi. The polar questions suffixed with -aa are disjoined by an interrogative disjunction operator (optionally spelt out as leed-aa), as shown in (38), and it has the semantics shown in (39), largely following Bhatt & Dayal(2018).

\[(38) \quad [\text{OR}_P\{\text{p},\text{q}\}]\]

\[(39) \quad [\text{OR}_Q] = \lambda Q_{(st,t)}\lambda Q'_{(st,t)} Q \cup Q'\]

We also saw that there can be one (seemingly) small disjunct on the surface, but the other one is a large disjunct as can be discerned from the overt phrase. We derive this by elision, (40).

\[(40) \quad [\text{ravi coffee taageeD-aa}] (\text{leed-aa}) [\text{ravi tea taageeD -aa}]
Ravi coffee drank-PQP if-not tea -PQP
‘Did Ravi drink coffee or tea?’ [Alt]

\(^4\)We attribute this difference between -aa and kyaa to the affixal vs independent-word status of -aa and kyaa respectively.
Thus in Telugu tupling with -aa (disjuncts of size ForceP) unambiguously leads to alternative questions. The only way to get a polar question interpretation is to have small disjuncts with the boolean disjunctive particle -oo (which is another role that -oo performs, where it is attached to phrases), and this does not have an alternative question interpretation. Why doesn’t the low occurring boolean disjunction operator get to scope over the question operator that -aa signals higher up in the clause, therefore delivering an alternative question interpretation? This is the puzzle we take up in the next subsection.

2.4.5 Scope of -aa and boolean disjunction in questions

Why can’t a sentence like (41a) have an alternative question interpretation where the disjunction indicated by the -oo is scoping over the question operator indicated by the -aa, as shown in (41b)?

(41)  a. #ravi coffee-oo tea-oo taageeD-aa?
     Ravi coffee-DISJ tea-DISJ drank-PQP
     'Intended: Did Ravi drink coffee or Did Ravi drink tea?'

     b. ORBOOL > -aa > CP[+Q]

If boolean disjunction takes scope over -aa (and the question-operator below it), the problem is the type-mismatch between what -aa delivers (a singleton-set) and what boolean disjunction expects (a multiple-set), as shown in (42). Thus an alternative question interpretation is not possible for a sentence like (41a).

(42)  a. [coffee or tea] = λP(ε,0). P(coffee) v P(tea)

     b. [-aa]=λQ(ε,0)Q

2.4.6 Information structural effects of -aa in cleft questions & sub-clausal positioning

To achieve information structural effects of not-at-issue and at-issue in polar and alternative questions, one strategy employed is that of clefts, discussed in §2.1.3. The cleft pivot, which is marked with -aa, is at-issue, and the rest of the cleft clause is not-at-issue. This information structural partition falls out naturally from the syntax-semantics of clefts, as is widely discussed in the literature, which we won’t go into here. The partitioning can be tested with favored continuations in gapping (43), and polar congruence (44), as discussed in Bhatt & Dayal (2018):

(43)  a. ravi coffee-aa taag-indi?
     Ravi coffee-PQP drank-CLM
     'Is it coffee that Ravi drank?...
     b. Tea-aa? 'or Tea?'
     c. #Uma-aa? 'or Uma'

(44)  a. ravi coffee-aa taag-indi?
     Ravi coffee-PQP drank-CLM
     'Is it coffee that Ravi drank?...
     b. leedu, Tea 'No, it was Tea.'
     c. # leedu, Uma 'No, it was Uma.'

Seemingly sub-clausal or clause-medial positioning of the PQP also has information structural effects, with the sub-clausal material that -aa attaches to as being at-issue, and the rest of
the material as being not-at-issue. This in in contrast with Hindi polar *kyaa*, where the material to the right of *kyaa* is at issue and the material to the left of it is not-at-issue, as observed by Bhatt & Dayal (2018). While Bhatt & Dayal propose that the material to the left of *kyaa* is in Topic positions, we propose that the material that -*aa* attaches to is in the Spec of ForceP, of which -*aa* is the head, (45b)-(46b). This material in the Spec of ForceP is in focus, and for which alternatives are generated due to its focus semantic value, leading to a narrow-focus informational structural effect.

(45) a. *ravi*-aa  *coffee taageeDu?*
   Ravi-PQP coffee drank
   'Is it Ravi that drank coffee?'

b. [Diagram]

(46) a. *ravi* *coffee-aa taageeDu?*
   Ravi coffee-PQP drank
   'Is it coffee that Ravi drank?'

b. [Diagram]

Thus, sub-clausal -*aa* leads to a narrow-focus interpretation of the polar question, whereas clause-final position of -*aa* leads to a wide-focus interpretation of the question (as the entire TP is now in the Spec of ForceP). This is similar to the focalizing effect of clitic heads to and *ki* in Bangla, with both possibilities, narrow-focus of particular phrases and wide-focus of the entire proposition, as observed in Bayer et al. (2014). They also provide a similar analysis, with movement of the focused element to the Spec of the phrase hosting the clitic.

2.5 Telugu -*aa* is a PQP

Telugu -*aa* has all of the three signature properties for a Polar Question Particle (PQP) that Bhatt & Dayal (2018) discuss – restriction to polar/alternative questions, selectiveness in ap-
pearing inside embedded polar/alternative questions, and finally, seemingly sub-clausal or flexible syntactic positioning. It is thus well-suited for being identified as another PQP cross-linguistically, that resides in ForceP and has a presuppositional requirement of a singleton-set question as complement. In the larger South Asian linguistic picture, we surmise that the Dravidian languages in relative contact with the Indo-Aryan languages developed the PQP -aa.

3 The quotative complementizer ani

Complementizers –morphemes that play the role of identifying clauses as complements –are known to have quite varied lexical sources (Bayer 1999), and in fact there may be no complementizer in natural language which does not have a lexical source, i.e a ‘pure’ complementizer. In this section we examine in detail the polyfunctional quotative complementizer (QC) in Telugu and build a composite picture by examining its role in various left-peripheral phenomena and its particular manifestation in various constructions like noun complement clauses, manner adverbials, with naming verbs, small clauses, and non-finite embedding, among others. The QC we conclude is instantiated at the very edge of the clause it subordinates, outside the usual left periphery, comes with its own set of projections, and is the light verb SAY which does not extend its projection (in the sense of Grimshaw 2005, i.e, the verb does not form the extended projections, the functional projections, IP or CP), unless called for (to host an operator or a morpheme that occurs in its higher extended functional projection, IP or CP). It adjoins to the matrix spine at various heights (at the vP level it gets a 0-role, and thus argument properties) when it does extend its projection, from F0 (the lexical category V) to F1 (the functional category T) or F2 (the functional category C), and like a verb selects clauses of various sizes (CP, TP, small clause).

We find that the QC retains a lot of what it derives from its lexical source, the verb ‘say’, and can unfurl a vP, an extended projection at F0, without any argument structure projections under it. It can take some rudimentary structure at the IP level, extending its projection to F1. But most importantly, it can also project another C-layer, extending its projection to F2, and therefore can embed its own C-level projections, Force, Evidence, Allocution, etc. It is into the C-domain of this QC structure that the relative complementizers -a and -ee in Dravidian can be located. We also find that the QC, located outside the usual left periphery of the clause that it subordinates (Jayaseelan 2014), comes with its own set of projections, that are the same for its source, the verb say. It adjoins to the matrix spine at various places –vP, IP, CP. Projection of the vP shell, and the other F-level projections reflects the verbal origins of the QC.

We exemplify our analysis mostly using the Telugu QC ani, as it is very illustrative (being more transparent in form to function mapping), but also draw from QC properties of Malayalam, Kannada, Bangla, and Meiteilon.

3.1 The QC + relativizer in Dravidian

We find an intriguing pairing of the QC and the relativizer (REL) in Dravidian (Jayaseelan 2014), in so called Noun Complement Clauses –unlike the Eastern Indo-Aryan group (Bangla, Assamese, Oriya), which have a QC but not a REL, (47)-(49).
The QC is at the leftmost edge of the complement, past even the question particle in CP (49) (Jayaseelan 2017). So where does the relativizer fit in? Where in the left periphery are the question particle, the QC and the relativizer accommodated? These are the questions raised by Jayaseelan (this volume) that we begin with.

We propose that this is where we see the QC unfurling its extended projection (Grimshaw 2005) to F2, the CP, (50).

The relativizer -a is in the C domain of the clause projected by the QC enn®. The question particle -oo is in the C domain of the object complement of enn®, a CP clause. There are thus two separate C domains here, one, of the embedded verb, and the other, of the QC.

This extending of its functional projection by the QC is also evident in other C-domain morphemes pairing with the QC, like the –conditional -Tee, (51), and the –concessive -aa, (52).

Finally, how the noun and its complement get together is another matter which we won’t get into here, except to sketchily say that this starts off as an equative copular structure –-[SC [DP vishayam ‘matter’] [CP Ravi raaDu ‘Ravi will not come’] – out of which extraction takes place leaving a gap and the relativizer in the left periphery to yield the phrase in (48) Ravi raaDu ann-a vishayam ‘The matter that Ravi will not come’. The structure for the sentence in (47) would thus be as in (53).
3.2 The polyfunctional behavior of the QC

3.2.1 QC with manner adverbials

The QC also shows up with onomatopoeic adverbials, (54).

(54) 'Dhap!' enna viinu. MALAYALAM
    QC fell
    '(It) fell with a thud.'

Here enn displays its verbal quotative nature and takes anything that can be 'said' as its complement, though the verbal root en 'say', of which enn is the perfective form, itself is obsolete in Malayalam (Jayaseelan 2014). The manner adverbial we propose is adjoined to the matrix clause at the level of the VP, (55).

(55) [vP [vP viinu...]] [vP enna [\textquoteleft Dhap!]]] MALAYALAM
    fell QC onomatopoeic sound

Telugu data is particularly instructive here, as it can show not only perfective, (56), but also progressive marking with the manner adverbial QC, (57).

(56) 'grr' an-i aagindi TELUGU
    QC-PERF stopped
    '(it) stopped with a 'grr'.'

(57) 'grr' an-Tuu aagindi
    QC-PROG stopped
    '(it) stopped with a 'grr'.'

This tells us that the QC this time is extending its projection to F1, TP (Grimshaw 2005), and the adverbial is attached to the matrix spine again by adjunction, (58).

(58) [vP [vP aagindi . . . ]] [vP -Tuu [vP an [\textquoteleft grr!]]]] TELUGU
    stopped PROG QC onomatopoeic sound

3.2.2 QC in naming constructions

The 'naming' structure is also a noun complement construction, with just a nominal as the complement of the noun, (59)-(60).

(59) kaakka en-a waakka MALAYALAM
    crow QC-REL word
    'The word 'crow''.

(60) ravi an-ee vyakti TELUGU
    Ravi QC-REL.NON.PST person
    'A person called Ravi'.
We analyze these as starting of as small clauses out of which the noun is extracted leaving a gap for the relativizer to fill, (61).

\[(\text{61}) \quad \text{waa}kka \left[ C_P \rightarrow a \left[ V_P \text{enn} [SC \quad \text{\_\_} \text{kaakka}] \right] \text{MALAYALAM} \right. \]

\[\text{word} \quad \text{REL} \quad \text{QC} \quad \text{crow} \]

In a designation/naming ECM structure, the QC subordinate clause, a small clause (SC) headed by the QC, attaches to the matrix again without unfurling any further, i.e. a VP-level adjunction:

\[(\text{62}) \quad \text{dii(n)-ni} \quad \text{Charminar ani pilustaa-mu} \quad \text{TELUGU} \]

\[\text{this-ACC Charminar QC call-3PL} \]

We call this ‘Charminar’.

The SC embedding interestingly gets translated in some Indian Englishes as \textit{We call this as Charminar}, because in English it is as that is the small clause complementizer. Moulton (2015) also notes parallel patterns between ECM structures in English and QC clauses in Bangla.

3.2.3 The QC and the evidential

The QC, the very bleached reportative (without argument structure), takes a nominalizer -Ta and forms the reportative evidential in Telugu:

\[(\text{63}) \quad \text{tinnaa-Du an-Ta} \quad \text{TELUGU} \]

\[\text{ate-3MSG QC-NMLZ} \]

‘Apparently, he ate.’

The reportative evidential can also co-occur with the QC:

\[(\text{64}) \quad \text{Tea taagutaadDu ani an-Ta} \quad \text{TELUGU} \]

\[\text{tea} \quad \text{drink-will QC QC-NMLZ} \]

‘(He) says that he will drink tea.’

This pattern is common in Meiteilon too (Kidwai 2014):

\[(\text{65}) \quad \text{ma ca thek-keni h\_y\_na h\_y\_ba ni} \quad \text{MEITEILON} \]

\[\text{he} \quad \text{tea drink QC}_1 \quad \text{QC}_2 \quad \text{BE} \]

‘He says that he would drink tea’

We analyze these as adjunctions of the QC in the left periphery of the matrix clause.

3.2.4 The QC and topic marker

The QC, the \textit{SAY}-shell, also takes the conditional -Tee, and the reflexive middle -kun, in its extended projection at F2 and its extended projection at F1 respectively, to form topic markers:
(66)  rao anu-koo/an-Tee  eemii tinaDu       TELUGU
     Rao QC-RFLX/QC-COND what   eat-not
     ‘As for Rao, he doesn’t eat anything.’

This is CP adjunction of the QC phrase, at TOPP into the matrix spine. It can also compose
with a distal determiner aTLu:

(67)  uma pustakam cadivindi-ann-aTLrao ceppeeDu       TELUGU
     Uma book    read-QC-DIST.DET  Rao said
     ‘Uma read the book, says Rao.’

Similar patterns occur in Meiteilon (Kidwai 2014):

(68)  Sita-na lairik-tu pa-re-háybadu  Ram-na hay       MEITEILON
     Sita-AGN book-DEF read-PERF-DIST.DET Ram       say
     ‘He says that he would drink tea’

3.3 QC in ‘embedded’ clauses

The QC can, of course, take a finite clause as its complement, normal ‘embedding’:

(69)  joby suzi wannu enna paRaññu       MALAYALAM
     Joby Suzy came  QC  said
     ‘Joby said that Suzi came’.

(70)  rao uma vaccindi ani naa-kuceppeeDu       TELUGU
     Rao Uma came   QC I-DAT tell
     ‘Rao told me that Uma came’.

The QC is not in the C-domain of its complement clause, but outside it. Here the QC stays
a VP, and doesn’t project any extended projections. It then attaches to the matrix spine as VP
adjunction. Clausal embedding with the QC is thus adjunction to the matrix verb.

The nominal QUOTE hosted in the SPEC of QUOTEP under VainP, (71), is what we propose
receives the object $\theta$-role from the matrix verb.

(71)  $[v \, \text{P ani} \, [\text{QuotP} \, \text{QUOTE \{k DIRECT\}}} \ldots$

The QC stays a VP or projects a VP. It then attaches to the matrix spine as vP or VP
adjunction. It if attaches as VP adjunction, wh-phrases in the QC embedded clause take
narrow scope, and if it attaches to the vP, they take wide scope (Kidwai 2014). In Telugu
and Bangla, they take wide scope, suggesting they are vP adjoined. In Kannada they can take
narrow or wide scope (Amritavalli 2013) indicating VP and vP adjunctions respectively. In both
cases it is the nominal QUOTE hosted in the SPEC of QUOTEP under VainP that receives the
object $\theta$-role from the matrix verb.
3.3.1 QC and non-finite complement clauses

Besides a full CP, the QC can also embed non-finite clauses, infinitives, (72), and small clauses, (73), as these examples from Telugu illustrate.

(72) *rao-ni cadava(m)-ani ceppeenu* TELUGU
    *Rao-ACC read.INF-QC told*
    ‘(I) told Rao to read.’

(73) *rao katti ani telusu*
    *Rao knife QC know*
    ‘(I) know that Rao is sharp (intelligent).’

3.4 The QC and its source

The QC *ani* in Telugu can be dropped under the verb that is its source, *anu* ‘say’. It is the only verb (along with *anu-koo* ‘think’ & *ani-pinc* ‘feel’, both from *anu*) that allows *ani* to be dropped:

(74) *rao uma vaccindi (ani) ann-aaDu* TELUGU
    *Rao Uma came QC said-3MS*
    ‘Rao said that Uma came’.

(75) *rao uma vaccindi (ani) anu-kun-(n)aaDu*
    *Rao Uma came QC said-REFLX-3MS*
    ‘Rao thought that Uma came’.

(76) *rao-ki uma vaccindi (ani) ani-(p)inc-indi*
    *Rao-DAT Uma came QC said–CAUS-3S*
    ‘Rao felt that Uma came’.

This we attribute to the verb *anu* having the same features as *ani*, being its source, and thus can fully take over its role.

3.5 Final thoughts on *ani*

The QC never stopped being a verb. The QC is instantiated outside the usual left periphery of the clause that it subordinates (Jayaseelan 2014). It comes with its own set of projections, and is the light verb *SAY* which does not extend its projection unless called for to accommodate morphemes in the higher functional projections (Kidwai 2014). It adjoins to the matrix spine at various places –vP, IP, CP. When it does extend its projection, it accommodates IP and CP level affixes. The QC should therefore not be equated with the complementizer *that* of English. It is the relativizer -a in Dravidian that is parallel to the complementizer *that* in English. It too, like *that*, is derived from the demonstrative *aa*, (Jayaseelan 2014).
4 The question particle -oo

The question particle -oo has many roles in Dravidian—most notably the boolean disjunction marker (as we saw in earlier sections), but also part of existential quantifiers and correlatives, and it has been well investigated in the literature by Slade (2011), Jayaseelan (2012), Szabolcsi (2015), and others. In this paper, we focus on the question particle role of -oo in Telugu.

4.1 The distribution of question particle -oo in Telugu

4.1.1 -oo in Matrix contexts

-oo shows an interesting pattern of distribution in matrix wh-clauses in Kannada (Amritavalli 2003, 2013) and Telugu. In matrix clauses in both Kannada & Telugu, a plain question interpretation arises only when the question particle -oo is left out, and the wh-clause is unmarked with any kind of question particle, (77).

(77) enta duuram velleeDu?
    how far went
    ‘How far did (he) go?’

The particle -oo is good in matrix wh-clauses, in both Kannada & Telugu, only when interpreted either as being embedded under wonder (78); or as an exclamation (79); depending on the intonation — (?) or (!).

(78) enta duuram velleeD-oo ?
    how far went-oo
    ‘I wonder how far (he) went.’

(79) enta duuram velleeD-oo !
    how far went-oo
    ‘How far (he) went!’

We won’t investigate the part that -oo plays in the formation of exclamatives in this paper, but refer the reader to Balusu (2018) for an extensive analysis of the role of -oo in wh-exclamatives. Instead we will look in some detail at the participation of -oo in the formation of indirect questions, which is more in line with the focus of this paper.

In matrix contexts, -oo in polar & alternative questions again leads to a ‘wonder’/indirect reading:

(80) coffee taageeD-oo tea taageeD-oo?
    coffee drank-oo tea drank-oo
    ‘I wonder if (he) drank coffee or tea?’ [Only Alt]

A polar question with -oo is marked, a polar alternative question is preferred:
(81) ??coffee taageeD-oo? 
coffee drank-oo
'I wonder if (he) drank coffee'

(82) coffee taageeD-oo leed-oo? 
coffee drank-oo not-oo
'I wonder if (he) drank coffee or not?'

The 'wonder' use of -oo goes beyond the English wonder, in that can be used in contexts like in (83).

(83) nii peeru eemiT-oo 
your name what-OO
'May I know your name please'
*I wonder what your name is'

What -oo delivers is a non-canonical meaning, namely, a non-intrusive question—as Farkas (2018) calls such uses with oare interrogatives in Romanian:

(84) oare pe.cine a invitat Rodica? ROMANIAN 
oare who.Acc has invited Rodica
'Who has Rodica invited, I wonder'

Like canonical questions, the Speaker raises an issue and thereby signals he wishes to have it resolved. Unlike canonical questions, the Speaker signals that he does not wish to put the Addressee on the spot for providing the answer. It is infelicitous in contexts where the Addressee is mandated to answer the question:

Doctor to patient

(85) *ivala enni goolilu tiisukunnaav-oo 
today how-many tablets take-OO
'How many tablets did you take today, I wonder.'

But it is appropriate in contexts where Addressee competence is presupposed, but Addressee may have reasons to withhold the answer:

Chef to guest

(86) mikku ee-kuura naccind-oo 
you which-curry like-OO
'Which curry did you like, I wonder.'

It is also appropriate as 'engaging' questions, in which Addressee competence assumption is absent:

Host of party to co-host

(87) enta mandi vastaar-oo 
how-many people come-will-OO
'How many people will come, I wonder.'
It is even appropriate in some self-addressed contexts (though not in *I-can’t find-the-value* questions):

_**Exasperated friend to a troublesome friend**_

(88) *ninnu enduku piliseen-oo*

you why called-oo

‘Why did I call you, I wonder.’

It does not necessarily have to be addressed to the Hearer, it can express a wish to know the answer:

(89) **A movie goer to a friend**

a. **Speaker:** *eppuDu avutund-oo*

when finish-oo

‘When will it finish, I wonder.’

b. **Addressee:** *nannu enduku aDugutunnaavu?*

me why asking

‘Why are you asking me?’

But it is not good when the Speaker knows the answer – ‘obvious answer’ rhetorical questions:

_**A race car driver at a car rental**_

(90) *naaku toolaDam enta baagaa vacc-oo*

I.DAT driving how well come-oo

‘How well do I know driving, I wonder.’

It is also not good when addressee is competent and there is no reason for non-cooperation:

_**To the session chair who is keeping track of time**_

(91) *inkaa enta seepu und-oo*

still how-much time be-oo

‘How much time is still there, I wonder.’

### 4.1.2 -oo in embedded contexts

In embedded clauses, in Telugu, -oo marked *wh*-clauses can appear under both rogative and responsive predicates, but never with the quotative complementizer *ani*, (92)-(93). An unmarked *wh*-clauses can occur embedded in Telugu only under rogative predicates (with the quotative present) — with responsive predicates there is only a matrix scope reading, (94).

(92) *eemi cadiveen-oo (*ani) aDigueeDu*

what read-oo QC asked

‘(He) asked (me) what (I) read.’

(93) *eemi cadiveen-oo (*ani) ceppeeDu*

what read-oo QC told

‘(He) told (me) what (I) read.’

(94) *eemi cadiveen (*ani) ceppeeDu*

what read QC told

‘What did (he) say that (I) read?’
In Telugu, -oo marks the scope of the wh-clause. When unmarked, both matrix and embedded scope are available for the wh-clause in Kannada (Amritavalli 2003, 2013), and only matrix scope is available in Telugu.

As we saw in §2.2, in embedded contexts, in both polar and alternative contexts, the preferred question particle is -oo, (95)-(96).

(95) a. cadiveen-oo leed-oo aDigeeDu
   read-DISJ not-DISJ asked
   'He asked if (I) read or not.'

   b. Rao-ki cadiveen-oo leed-oo telusu
   Rao  read-DISJ not-DISJ knows
   'Rao knows if (I) read or not.'

(96) naa-ku tea kaavaal-oo coffee kaavaal-oo aDigeeDu
   I-DAT tea want-DISJ coffee want-DISJ asked
   '(He) asked whether I want tea or whether I want coffee.'

4.2 Analysis of the question particle -oo

4.2.1 -oo as question-operator

Jayaseelan (2001, 2012) makes the theoretical claim that the question-operator is the disjunction operator, universally. For Jayaseelan, the homophony of the question particle and the disjunction marker -oo is not accidental, and explains why the question-operator (always abstract in English), is realized as the disjunction marker -oo in Malayalam, and -ka in Japanese.

Interestingly, in Malayalam only polar/alternative questions are marked with -oo (both matrix and embedded). Malayalam constituent questions are not marked with -oo — whether matrix or embedded:

(97) aare wannu? MALAYALAM
   who came
   'Who came?'

(98) avan [aare wannu enne] paraññu/coodiccu MALAYALAM
   he who came QC said/asked
   'He told/asked who came.'

Jayaseelan (2001) proposes that there is a superficial deletion rule in Malayalam that deletes an underlying -oo in constituent questions:

(99) aare wannu-oo? (100) avan [aare wannu-oo enne] paraññu/coodiccu

The claim is that all Malayalam questions—both polar questions and constituent questions—are marked by a clause-final -oo. A question clause has the disjunction operator in the head position of ForceP. The clause-final ‘question marker’ -oo is the realization of the disjunction operator and it is generated as the head of ForceP.
What if we transpose this claim to Kannada & Telugu matrix constituent questions? This gives us an interesting contrast: when covert, -oo forms a direct question; when overt, -oo forms an indirect question or exclamative:

(101) *evaru vacceeru-ee?*  (102) *evaru vacceeru-oo?*  
who came who came-oo  
‘Who came?’ ‘I wonder who came?’

### 4.2.2 -oo as interrogative complementizer

Amritavalli (2003, 2013) treats -oo as an interrogative complementizer that precisely delimits the scope of an embedded question in Kannada. For her, -oo is a question complementizer like *whether*, rather than a question operator. So apparent matrix *wh*-questions, when suffixed with -oo, are understood as indirect or embedded questions, because of a ‘silent’ matrix non-veridical predicate. -oo occurs in C₁ with responsive predicates and in C₂ with rogative predicates in her analysis:

(103) \[
\text{Speech Act} \quad \text{[C₁ Force} \quad \text{(Topic*)} \quad \text{[C₂ (Focus)} \quad \text{(Topic*)} \quad \text{[C₄ Finite. [[[T [R \ldots]]]]]]]}
\]

to ka no JAPANESE  
antæ/oo oo/aa wh- annoo-du KANNADA

The problem with this analysis is that if -oo is an interrogative complementizer, it should not occur on both the juncts of a polar question, like in (104).

(104) *neenu ravi vacceerD-oo leed-oo kanukkunnaanu*  
I Ravi came-oo not-oo found out  
‘I found out whether Ravi came or not.’

A more damaging piece of data for this analysis is that there is another morpheme which is a more viable candidate for the interrogative complementizer – the complex *eem-oo*, which like the English *whether* is also composed of a *wh*-word and a disjunctive element, as shown in (105).

(105) *neenu ravi vacceerD-eem-oo kanukkunnaanu*  
I Ravi came-what-oo found out  
‘I found out whether Ravi came.’

### 4.2.3 Our analysis of question particle -oo in matrix contexts in Telugu

Following Farkas (2018), we propose that here -oo weakens the Addressee compliance effect of canonical questions, to form non-intrusive questions. -oo questions widen the range of projected futures, which now include not only ones in which the Addressee resolves the issue just raised, but also a future in which he choses not to, leaving the Addressee more choice, allowing for no Addressee response. Addressee compliance no longer requires the Addressee to resolve the issue raised.
The Addressee therefore can comply with an -oo question without resolving the issue, though futures in which he does resolve it would be preferable, since those involve information increase. -oo thus marks a question for being non-intrusive, and -oo’s role is to signal widening of the projected states to include, besides context states in which the Addressee resolves the issue, a context state in which he does not.

Interrogatives place an inquisitive proposition on the Table and project canonical states in which the Addressee volunteers information that settles the issue. Adding a non-intrusive marker allows the Addressee to comply without volunteering such information, either because he doesn’t have it or because he doesn’t wish or is not willing to provide it. The Speaker is still seen as wishing to have the issue resolved; if he didn’t, he could have remained silent.

-oo marked interrogatives are appropriate when the Addressee is assumed to know the true answer, but the context justifies asking a ‘softened’ question, and also in contexts in which the Addressee is not assumed to know the answer.

4.2.4 Our analysis of question particle -oo in embedded contexts in Telugu

Cable (2010) proposes that question particles originate adjoined to a structure containing a wh-item or project a QP layer containing it, (106)-(107).

(106) Q-adjunction:

\[
\text{XP} \quad \text{XP} \quad \text{Q} \quad \text{wh}
\]

(107) Q-projection:

\[
\text{QP} \quad \text{Q} \quad \text{XP} \quad \text{wh}
\]

The constituent with a Q-feature is attracted to the CP by the interrogative probe that is seeking to check its Q-feature. In wh-in-situ languages with clause-edge question particles, the question particle undergoes overt movement to the Spec of CP, (108).

(108)

\[
\text{CP} \quad \text{C} \quad \text{TP} \quad \text{XP} \quad \text{XP} \quad \text{Q} \quad \text{XP} \quad \text{wh}
\]

This is the analysis that is proposed for Sinhala, and for Japanese, where it is argued for using data that is claimed to show the base location of the question particle, or with data showing intervention effects of focus-marked elements on question particle interpretation. Hagstrom (1998) proposes that what is usually considered the interrogative complementizer in Japanese, -ka, is actually a question particle that is generated clause internally, and then moves overtly to C, thus showing overt movement and subject to intervention effects and the like. He also
argues that, in Sinhala, the question particle overtly stays in its base position, i.e. clause-internally, but undergoes covert movement to the C-domain, just like -ka.

But there is no data from Telugu that shows either the base generation of the question particle or intervention effects. We therefore take the question particle to be base generated in the Spec of CP in Telugu.

The semantics of the question particle, following Kotek (2014), is to take the focus semantic value of the wh-sentence and return this as the ordinary semantic value, (109), another kind of domain widening, in a way like the domain widening role of -oo with alternatives, in matrix question contexts, and also in boolean disjunctions.

\[(109)\]
\[
\begin{align*}
\text{a. } & [Q \alpha_0]^0 = [Q \alpha_0]^f \\
\text{b. } & [Q \alpha_0]^f = \{[Q \alpha_0]^f\}
\end{align*}
\]

The question particle is thus essential for the interpretation of an interrogative sentence. This explains why the question particle -oo delimits the scope of the wh-phrase in Telugu, and this also explains why embedded wh-phrases can take higher scope, (110).

\[(110)\]
\[
\text{ravi eemi tinnaaDu ani Kiran ceppeeD-oo naaku telusu}
\]
\[
\text{Ravi what ate QC Kiran said-OO I-DAT know}
\]

'I know what Kiran said that Ravi ate.'

5 Conclusion

In this paper we hope to have shown that none of the so called complementizers in Telugu (and other Dravidian languages) which appear to be like complementizers –ani, -oo, and, -aa – are really complementizers when looked at closely. -aa found in Kannada, Telugu, Tamil, and Malabar Malayalam is the Polar Question Particle. It stays in the root clause. It only appears in embedded contexts under quasi-subordination. In alternative questions, it attaches to each member, it marks the scope, and delivers an unambiguous reading. The quotative complementizer ani never stopped being a verb. It is instantiated outside the usual left periphery of the clause that it subordinates (Jayaseelan 2014). It comes with its own set of projections, and also extends its projections from F0 to F1 to F2 (Grimshaw 2005), based on where it attaches to the matrix clausal spine and what morphemes it accommodates in its own extended projection. -oo in Telugu is the question particle that is base generated in the Spec of CP and which plays a crucial role in interrogative sentence semantics, thus also delimiting the scope of the wh-phrase.

References


Discourse-driven scrambling to the peripheries in Child Tamil

R. Amritavalli & Annu Kurian Mathew
amritavalli@gmail.com
annukmathew@gmail.com

1 A prediction about word order in subject questions

1.1 A pre-verbal focus position

Jayaseelan (2001, 2010) notices that wh- question words have to occupy the immediately pre-verbal position in the Dravidian SOV language Malayalam. He argues that they move to an IP-internal focus position. His data and analysis generalize to the other literary SOV Dravidian languages Kannada, Telugu and Tamil. Here we discuss child Tamil.

Jayaseelan’s analysis crucially accounts for an obligatory movement of a subject question word, apparently downward and rightward, to a pre-verbal position ‘below’ the object, which is not a COMP position. Illustrating this for Tamil, in (1) the subject question word yaari ‘who’ must appear ‘below’ and to the right of the object, but to the left of the verb; the COMP position is sentence-final.

(1) Tamil. onn-e yaari adjcc-aa? OSV
   you-ACC. who beat.PST-3p.
   ‘Who beat you?’

How does the wh-subject appear in such a sentence-internal position? Jayaseelan adopts two assumptions. One, that wh- moves into a focus position, as earlier suggested for Hungarian (Brody 1990) and Basque (Laka and Uriagareka 1987), among other languages. Two, Antisymmetry: i.e. that (1) has an underlying VO order. The wh- word moves to a pre-verbal focus position. The other arguments vacate the VP, such that the object moves past the subject.

(1) a. [IP … [FocP yaari ‘who’ [IP yaari who’ [VP onn-e ‘you-ACC.’]]

1.2 A post-verbal topic position

A post-verbal topic position was first noticed by Tirumalesh (1996) for Kannada, and reiterated by Jayaseelan for Malayalam, cf. (2). As both authors point out, indefinite NPs, which cannot be topics, cannot appear post-verbally. (In (2), the topic follows Finite NEG. Our Tamil child data show that the topic also follows the question particle, presumably in ForceP.)

*The child data reported here are from the second author’s work towards her doctoral dissertation. We thank Keiko Murasugi and K. A. Jayaseelan for helpful discussion. Our thanks also to two anonymous reviewers, and Jaklin Komiflt for comments and questions. Any errors remain our own.
1.3 A prediction about word order in subject questions

Given a pre-verbal Focus position and a post-verbal Topic position in Tamil as in Malayalam, and given that wh- must move to Focus,¹ we now make a simple prediction: a wh- word in Tamil (as in Malayalam) moves to a pre-verbal (focus) position, but it cannot move to a post-verbal (topic) position. This entails that a subject question (with a transitive verb) (i) must manifest a non-canonical, scrambled order, and that (ii) this must be the order OSV. A subject question cannot manifest a scrambled order OVS, or a non-scrambled order SOV.

Turning to child Tamil, we check if subject questions obey the three restrictions stated above. A subject wh- word must scramble to an immediately pre-verbal position (ok OSV, *SOV). But it must not scramble to a post-verbal position (*OVS). If these three word order restrictions are met, we can infer that child scrambling moves arguments out of canonical positions to “criterial” positions to check topic/ focus features. And indeed we find that subject wh- words are never left in situ. They always occur in an immediately pre-verbal position, very early in the data: between 26-29 months. They, however, never scramble to a post-verbal position, arguing for very early, error-free acquisition of focus and topic feature-driven scrambling to the peripheries.

2 Related Empirical and Theoretical issues

Interestingly, Japanese, an SOV language, also prohibits post-verbal wh- words. A scrambled order SVO occurs in declaratives, but SVO is prohibited in object questions (Murasugi and Sugisaki 2008). These authors show that the restriction on scrambling wh- to a post-verbal position is acquired early, and offer this as evidence for the child’s knowledge of canonical as against movement-derived word orders. They however do not offer an explanation for the restriction against post-verbal wh-. We suggest that in Japanese as in Dravidian, the post-verbal position may be a topic, which cannot host the inherently focused wh-word. We return to this point in section 4.

Mathew (2014, 2015) proposes an alternative account of (1) (repeated below) for Malayalam, that does not assume Antisymmetry. On her account, the wh- subject is only apparently immediately pre-verbal; it remains in situ. But a wh- is an indefinite, and indefinite subjects trigger object topicalization into “a Topic position available in the left periphery of Malayalam, a la Rizzi (1997)” (cf. her examples “A person Priya saw, ok Priya-ACC. a person saw, ’2014:26), because the leftmost element is interpreted as a default topic: a reading “unmistakable in the case of a non-subject at the left edge.” Thus (1) has the structure (3), not (1a). ((1a) is repeated below for comparison.)

¹ Cf. Rooth (1996:271): “The position of focus in an answer correlates with the questioned position in wh questions ...” Question-answer congruence suggests that wh- is focused. The function of focus is to evoke alternatives. For questions and answers, the relevant alternative set has its basis in the semantics of questions.
(1) a.  \[\text{[\text{FOC} y\text{aari} 'who' \text{[\text{FOC} y\text{aari} 'who' [\text{VP \text{on}\text{-n-e 'you-ACC.'}]]}]\]}

Mathew's proposal also correctly derives the OSV order in (1), and blocks the OVS order for subject questions: given that the \textit{wh}- word is an indefinite, it cannot be a topic, so it cannot be post-verbal.

Now there are instances where multiple elements intervene between the subject and the verb, in the canonical order. In these instances as well, a \textit{wh} subject occupies the immediately pre-verbal position, to the right of all the intervenors. Mathew (M) therefore proposes that when there is a \textit{wh}-subject, all "items that might otherwise appear between the \textit{Wh} and the verb" including "PPs, adverbs etc.," are topic-marked, and move to topic positions at the left periphery. We note, however, that the topicalization of even a single intervenor would suffice to satisfy the prohibition against leftmost indefinite elements. The topicalization of all intervenors is not therefore needed to remove the \textit{wh}-indefinite from a default topic (i.e., leftmost) position. It is stipulated in order to account for the consistently immediately pre-verbal position of \textit{wh}.

We must also point out (in response to a reviewer, who maintains that the cartographic analysis is "the issue at stake" in the analysis of (1)) that M's analysis, no less than Jayaseelan's (J's), appeals to "cartographic encoding," i.e. to "elements pertaining to information structure finding their own place in the functional sequence" (Mathew 2014:15). The difference is that instead of focus movement to a focus position for the question word, M resorts to topic movement of all elements that intervene between the subject question word and the verb. We are aware of no other, non-cartographic, account of the data in (1). As such, we cannot agree with the reviewer's observation that "children have early movement to criterial positions only if the cartographic analysis is correct to begin with."

The real "issue at stake" is M's rejection of Antisymmetry. This issue is of tangential interest here. We are aware of purely syntactic arguments (that we do not have the space to fully articulate) that favour the Antisymmetric account (Jayaseelan 2001, 2010, and p.c.) as arguably more complete and coherent. Our concern here, however, is with the word order of subject questions in the acquisition of Tamil (section 3).\footnote{A reviewer asks if S(wh)OV is a possible but non-preferred order in adult Tamil/Dravidian. We maintain, with Jayaseelan and Mathew, and against the prevailing view of Dravidian languages as \textit{wh}-in-situ, that S(wh)OV is not a possible order in non-rhetorical questions. In fact, our initial motivation for investigating question word order in child Tamil was to check if acquisition data concur in this respect with our adult native speaker intuitions; which, indeed, they do (cf. sections 1.3 above, and 3, 3.2.1-3.2.2 below).}

We aim to add to a claim that scrambling is, in some languages, acquired very early, and in a principled way; with the caveat that there are syntactic and semantic reasons to prefer an analysis where a \textit{wh}- word moves into a focus position.

As we shall see in section 5, when we briefly return to the syntactic dispute, child data corpora are accidental enough, and child Tamil utterances elliptical enough, that they cannot on their own decide between J's and M's accounts. Nevertheless, we highlight some core issues that M does not address. In particular, section 5.1, which responds to a reviewer query about a "nominalized" structure that allows a post-verbal \textit{wh}- (e.g., an apparent OVS subject question), deals with what we analyze (following Jayaseelan) as a cleft. This structure does not occur in our child data, but is a prominent site of dispute in J's and M's accounts. We differentiate a post-verbal cleft focus from the post-verbal topic in a finite clause. A Subject that is a post-verbal cleft focus can carry focus markers (and/or a yes-no question particle), whereas a Subject that is a post-verbal topic cannot do so. We thus confirm that the post-verbal \textit{wh}-Subject in the "nominalized" structure is not in a topic position, but is a cleft focus.
3 The Acquisition Data

As a pre-condition to the discussion of "scrambled" or criterial movement-driven word orders, we first establish the predominance of, and therefore knowledge of, canonical (S)OV order in the Tamil child. We then show that a subject wh- is never in the canonical S position, although 124 subjects occur in canonical S position in two-argument SOV sentences in our data. A subject wh- (indeed, all wh-) is always immediately pre-verbal. It is never post verbal. Thus subject questions in transitive sentences are always in a non-canonical word order; and they always have the scrambled order OSV, but never the order OVS.

3.1 Canonical word order and scrambling in child Tamil

In Tamil as in other languages, evidence for canonical OV word order, namely the occurrence of OV, N-Postposition, and pre-verbal complements to be, is seen even at the two word stage. This stage occurs in Tamil at 16-22 months, as seen in longitudinal data from the Vanitha database (CHILDES, a girl) and the MPI-CIEFL database (a boy).

For scrambling, we analyse 22,811 utterances between 23-32 months, i.e. after the two-word stage, from the two subjects mentioned above and a third, male subject from the MPI-CIEFL database. We analyse in all 54 hours of longitudinal data. Excluding utterances irrelevant to word order such as intransitive SV utterances, single argument utterances, or utterances of a single inflected verb with all arguments dropped, 4485 utterances are identifiably head final: they consist of a Verb and a pre-verbal Object, Complement or Adjunct, without or with the Subject.

Of these 4485 utterances, 4231 utterances (over 94%) are in canonical order, and 254 or 5.66% are in identifiably non-canonical orders. The latter are utterances with a verb and two overt arguments in the orders OSV, SVO, OVS, or DO-IO V; or utterances with the verb be where S and a complement or adjunct occur in non-canonical order.

We take this proportion of non-canonical to canonical word order utterances as robust evidence for knowledge of canonical and non-canonical orders in the 23-32 month old Tamil child.

3.2 Wh- questions between 23-32 months

In all our data, if a verb is overt, the wh- word is immediately pre-verbal. It is never post-verbal.

There are in all 630 wh- questions in the data. But 580 of these consist either of the wh- word plus a verb, or of copula-drop utterances (wh- word plus argument/adjunct). They are thus uninformative about the position of the wh- word vis-à-vis other arguments, complements or adjuncts; or not obviously informative about the position of the wh- word vis-à-vis the copula.

For our purposes, therefore, we consider 50 wh- questions occurring in identifiably non-canonical word orders (a subset of the 254 non-canonical order utterances mentioned in section 3.1).

---

3 There are 68 first-person or self-referent subjects: ‘I throw the ball, I put on shoes, (name) put on shoes,’ and 32 first-person dative subjects: ‘I want pen, I know that.’ Eleven 2nd person subjects occur in directives: ‘You draw a cockroach.’ There are 4 other 2nd person subjects: ‘Did you bring chocolates?’ and 9 3rd person subjects: ‘Mini akka is taking photos, Nurse puts (in) a needle, Mother gave a band-aid.’

4 Created with funding and technical support from the Max Planck Institute, Nijmegen, The Netherlands, at the (then) Central Institute of English and Foreign Languages, Hyderabad, this database has 52 weeks of 1-hour weekly video and audio recordings of natural speech interaction for 4 children (2 male, 2 female), starting ages 15 to 32 months. We thank Bhuvana Narasimhan for initiating the collaboration with the first author. The second author was a member of the research team, with Gayathri Raman. Others who helped create the database were Mini, Vijaya, Sanjay Gupta, and Vindhya Singh.
3.2.1 Subject questions

Of the 50 \(wh\)-questions in non-canonical orders, 25 are subject questions that clearly evidence \(wh\) in an immediately pre-verbal position, and not in the utterance-initial canonical subject position. Six subject questions that appear from the 28\textsuperscript{th} month on are OSV ((4-5) below; (4a-b) are consecutive utterances).

(4) a. OSV \(co\ \pi is\ \ \text{taatti} \ \text{kuttaa}\) 2;03.20
    chalk piece who (baby talk) give.PST.3PL.
    ‘Who gave (the) chalk piece?’

b. \(\text{idi} \ \text{yaari} \ \text{kud\text{\text{"a}}a}\) 2;03.20
    this who give.PST.3PL.
    ‘Who gave this?’

(5) OSV \(\text{idi} \ \text{aar} \ \text{uutna}\) 2;04.23
    this who blow.PST.3PL.
    ‘Who blew this up?’

Occurring earlier (starting at 26 months) are subject questions with \(be\) in the Complement-\(S\)-\(be\) order. The normal order is \(S\)-Complement-\(be\). There are 17 such questions.

(6) a. Complement-\(S\)-\(be\) \(\text{ulla} \ \text{enna} \ \text{irikki}\) 2;01.26
    inside what be.NONPST.3SG.N
    ‘What is inside (this)?’

b. \(\text{anga} \ \text{aar} \ \text{irkaa}\) 2;04.17
    there who be.NONPST.3PL.
    ‘Who is there (at that place)?’

Finally, two subject questions show a locative adjunct before the subject, and not after it.

(7) Adjunct-\(SV\) \(\text{meela} \ \text{aar} \ \text{peesaraa}\) 2;04.10
    above who speak.NONPST.3PL.
    ‘Who speaks/ is speaking up there?’

3.2.2 Intermediate conclusion

Subject questions between 26-29 months occur in the scrambled order XSV, i.e. OSV, Complement SV or Locative SV. They do not manifest in the canonical orders SOV, S-Complement-V or S-Locative-V. Recall that over 94\% of child utterances are in the canonical order; and that 124 non-\(wh\) transitive subjects manifest in canonical (SOV) position. These data therefore observe 2 of the 3 word order restrictions on subject questions: a subject \(wh\)-word must occur pre-verbally; it cannot occur in the initial, canonical subject position (*SOV, ok OSV).

3.3 The post-verbal Topic

We now turn to the third restriction: the OVS order is prohibited in subject questions, because the post-verbal position is a topic position that cannot host a \(wh\)-word.

Post-verbal topics occur in our data at around 26 months. There are 173 utterances with post-verbal topics, including S, O, IO, Complement, Adjunct, and Dative Subject. Of the 173 post-verbal topics, 71 are Subjects. Cf. the OVS utterance (8). There are 14 such OVS utterances.

(8) OVS \(\text{ad.d\text{\text{"a}}} \ \text{sonn\text{\text{"e}}} \ \text{naani}\) 2;02.01
    that. EMPH,REFL,FOCMKR. say.PST.1SG I
    ‘I said precisely that.’
But not a single post-verbal subject is a wh-word. In particular, there are no subject questions in the OVS order.

There are indeed 6 OVS questions, but these are all object questions, with pre-verbal wh-.

(9) OVS    enna      padraan      ivan?  2;01.18
            what       do.NONPAST.3M.SG  he.PROX.

‘What’s he doing?’

With this, we see that subject questions in Tamil children 23-32 months of age obey all 3 word order restrictions *SOV, OSV, *OVS.

4 A possible generalization with Japanese

We have just mentioned that there are 6 object questions in the OVS order. Let us now add that there are no object questions in the SVO order. This is the fact that invites comparison with Japanese child data. Murasugi and Sugisaki (2008) report that 2 longitudinal corpora for Japanese attest, by age 2;5, both (i) utterances in a scrambled VO order, and (ii) direct-object wh- questions, “reasonably often.” But there are in these corpora no object wh- questions in the VO order; and this absence of SVO object questions is in accordance with adult Japanese.

The Japanese and Tamil corpora are comparable: 33000 lines of child speech (Japanese), 22,811 utterances (Tamil), as are the child ages (upto 29 months for Japanese, 23-32 months for Tamil; all data cited here are in the 26-29 month range).

Murasugi and Sugisaki do not speculate why the SVO order is illicit for object questions in Japanese. Their interest is to show that the child exhibits knowledge of this restriction, which justifies the inference that SVO must indeed be a derived or scrambled word order for the Japanese child. We have made the corresponding inference regarding knowledge of canonical order in child Tamil at the outset, by comparing the proportion of canonical and non-canonical word order utterances in our database. We went on to explain the restriction against a post-verbal question word as an exclusion of a wh-word, which is inherently focused, from a Topic position. We now suggest that our hypothesis might extend to Japanese: the post-verbal position must be a topic position.5

5 A reviewer observes that in Turkish/Turkic, too, wh-constituents and other focused elements cannot show up post-verbally, and children seem to be aware of this restriction starting at a rather early age.

5 "We will not explore scrambling or the number of possible Topic positions here; the aim of the discussion here being to show that the scrambled object is dislocated to a Topic position at the C-domain,” Mathew (2014:24, n.4.).

5 What focus, or topic movement of everything else?

The data presented so far are compatible with either the Antisymmetric, wh- as focus account of Jayaseelan, or the account of Mathew of wh- as indefinite, triggering topic movement to the left-periphery. We now briefly consider data that potentially differentiate these accounts.

J and M both acknowledge the possibility of sentence-internal scrambling of the direct object over the indirect object. M gives the example ‘Rajan [that cat] [to Priya] gave’ (2014:20), but she does not elaborate on how the SOV structure could accommodate IP-internal topicalization.6 J (2001) posits IP-internal topic positions above the IP-internal pre-verbal focus position (as in Rizzi’s left periphery), in addition to left-peripheral and right-peripheral topic positions.

6 “We will not explore scrambling or the number of possible Topic positions here; the aim of the discussion here being to show that the scrambled object is dislocated to a Topic position at the C-domain,” Mathew (2014:24, n.4.).
The IP-internal and left-peripheral topicalization of the direct object ‘water’ are illustrated below (example (10) = J’s (23)). Topicalization yields a definite reading for this argument.

(10) a. ñaan ḍɪvə marattiṅa wellam ozhiccu. S IO DO V (Canonical order)
   I poured water to a tree.

b. ñaan wellam ḍɪvə marattiṅa ozhiccu. S DOtopic IO V (IP-internal topic)
   I poured the water to a tree.

c. wellam ñaan ḍɪvə marattiṅa ozhiccu. DOtopic S IO V (Left-peripheral topic)

Our data do have 3 instances of DO-over-IO Topicalization, in (11a-c). But the Subject is not overt in them (two utterances are imperatives, and the third is a pro-drop utterance). So these data do not speak to the question of IP-internal topicalization versus left-peripheral topicalization. Hence our contention that child data cannot per se choose between the Antisymmetrical and the wh-in-situ clause structures; corpus data are accidental (lacking the crucial example), and the utterances often elliptical.

(11) a. raju maamaa, ina puucci-ya… ina puucci-ya kunjaali-kki viqi 2;05.17
   Uncle Raju, this insect-ACC…this insect-ACC Kunjaali-DAT leave-IMP
   ‘Uncle Raju, this insect, this insect, set (it) on Kunjaali (grandmother).’

b. adi ena-k konṭaα. 2;07.26
   that I-DAT bring-give-IMP
   ‘Give me that.’ (The child wants the handkerchief his mother wipes his nose with.)

c. pro caakle ḫʊnā kki tari-v-eens. 2;02.28
   pro chocolate you-DAT. give-FUT.-1.SG
   ‘(I’ll) give you the chocolate.’ (Collecting a bunch of chocolates from her father’s hand, the child promises to share them with him.)

Moving on to right-peripheral or post-verbal topics, we have seen that a non-wh-element in a wh-question can occur in this position: in (9) above, a subject is the post-verbal topic in an object wh-question. Again, compare the locative wh-questions (12) and (13): in (13), the object is a post-verbal topic.

(12) idi enga pooḍṭo? 2;04.03; 06.28min
   this where put-MODAL
   ‘Where should (I) put this?’

(13) ed-la pooḍṭo id-a? 2;04.03; 07.19min
   what-LOC put-MOD. this-ACC.
   ‘In what should (I) put this?’

We note that rightward topicalization in (9) and (13) results in the wh-word surfacing in the leftmost position, in apparent violation of the prohibition against leftmost wh- indefinites (pace Mathew). These data show that topicalization does not effect only the removal of a wh- from leftmost position, as Mathew claims. They question if wh- subjects inevitably trigger leftward topic movement. We have already questioned, in section 2, the plausibility of across-the-board leftward topicalization of all elements intervening between a wh-subject and the verb.7

---

7 Not all superficially left-displaced elements may be Topics. In the successive utterances below, the Subject and locative arguments of be are clearly “interchanged.” Is Subject (which is intonationally focused) in pre-verbal Focus? Or has Location moved left to Topic? (If the Subject ‘dog’ is a specific
Post-verbal topics in Tamil (and Dravidian more generally), as in Turkish, serve to “background” information (Kornfilt 2005:176-177, n.8), or to introduce background information. We thus find in our data a number of 1st and 2nd person post-verbal subjects, that are otherwise prime candidates for pro-drop. This does not argue (however) that they are not topics. Nuances in types of topics/ topic positions might be related to nuances of definiteness (Schwarz 2009, 2013). Post-verbal topics are perhaps “familiar” topics or, like subjects, “continuing topics” (Jenks 2018). Jayaseelan (2001, n.13) observes that even an indefinite noun phrase, where it “has been the subject of conversation, and therefore ... is part of the ‘given information’ (i.e. Topic),” “is not always interdicted from appearing in the post-verbal Topic position— or for that matter, in the pre- subject Topic position” (citing a Malayalam example ‘I too got a letter!’ where ‘a letter’ occurs post-verbally).

Multiple topics can occur post-verbally, as pre-verbally (Jayaseelan op.cit, n.14). Anaphor and quantifier binding do not differentiate pre- and post-verbal topics in Tamil; unlike in Turkish, both reconstruct.

### 5.1 A post-verbal Cleft Focus

In the “nominalized” structure (14), the post-verbal argument is Focus. As shown in (15), it can be marked with the focus particle taan (an emphatic reflexive, cf. (8) above), and/or the emphatic particle -ee.

(14) **unn-e** paart-ada Karṇan/avan.
you-ACC. saw-NOML. (name)/ he
‘It is Kannan/ It is he that saw you.’

(15) **unn-e** paart-ada avan-taan / avan-ee taan.
you-ACC. saw-NOML. he-FOC.(EMPH.REFL.)/ he-EMP. FOC.(EMPH.REFL.)
‘It is indeed he himself that saw you.’

In contrast, the post-verbal Topic in a finite clause cannot be emphasized or focused.

(16) *unn-e paart-aan avan-(ee) taan.
you-ACC. saw-3MSG. he-EMP. FOC.(EMPH.REFL.)
‘Saw you, he himself (indeed).’

indefinite, it need not trigger locative topicalization: a specific/ contrastive/ partitive indefinite is acknowledged to be possible sentence-initially (Mathew 2014:25.)

(i). a. **anga** naai in-nil-idi.
there dog be-PST.3.SG.N
LOC-S-be 2:05.02
‘There was a dog there.’

b. **uḷḷa** naai in-nil-idi.
there dog be-PST.3.SG.N
2:05.02
‘There was a dog inside.’

---

8 We thank Ambalika Guha (p.c.) for this point.
9 Jayaseelan considers two derivations for post verbal topic(s): a topic field below the pre-verbal focus position, filled by VP-vacating movements; or (since the verb that these topics follow may carry negation and question markers) left-peripheral topicalization, followed by IP topicalization to a higher position in the CP.
10 Cf. (a) ava ottar-e ottar paarkale ‘they each other-ACC. did not see’ ‘They did not see each other.’
ottar-e ottar ava paarkale ‘Each other, they did not see.’ (pre-verbal topic)
avpaarkale ottar-e ottar ‘They did not see, each other.’ (post-verbal topic)
(b) yaarum/ ottarum avaroṛę tapp-e ottikale ‘QUANT. their mistake-ACC. did not acknowledge’
‘No one/ not a person acknowledged their mistake.’
avaroṛę tapp-e yaarum/ ottarum ottikale ‘Their mistake, no one acknowledged.’ (pre-verbal topic)
yaarum/ ottarum ottikale avaroṛę tapp-e ‘No one acknowledged, their mistake.’ (post-verbal topic)
The post-verbal Focus can be questioned with a yes-no particle attaching to it (17a). The
nominalized verb cannot be questioned in this way: (17b) is uninterpretable.

(17) a. *unn-e paart-ada avan-aa/ avan taan-aa/ avan-ee-vaan?
   you.ACC. saw-NOML. he -Q he FOC-Q he-EMP.-Q
   ‘Is it (indeed) he (himself) that saw you?’

   b. *unn-e paart-ada-aa avan?
   you.ACC. saw-NOML.-Q he
   ‘Did that saw you, he?’

The post-verbal Topic in a finite clause cannot be questioned with a yes-no particle (18a). The
yes-no particle must attach to the finite verb (18b).

(18) a. *unn-e paart-aan avan-aa?
   you.ACC. saw-3MSG. he-Q
   ‘Did he see you?’

   b. unn-e paart-aan-aa avan?
   you.ACC. saw-3SG.-Q he
   ‘Did he see you? (‘Saw you, did he?’)

The post-verbal Focus in the “nominalized” structure can thus be clearly differentiated from
a post-verbal Topic in a finite clause.11

The “nominalized” structure is a cleft focus construction with a copula. In Tamil, the copula is
usually dropped, but it surfaces if there is a modal.

(19) unn-e paarta-ada avan taan ir-ka-ṇum / ir-ka-łaam.
    you.ACC. saw-NOML. he EMPH.REFL. be-INF.-MOD./ be-INF.-MOD.
    ‘It must be he that saw you / It may be he that saw you.’

Consistent with J’s claim that wh- occurs in Focus, a wh-word can occur in a post-verbal cleft
focus position, as shown in (20a) (indeed, clefting is said to be the default question strategy in
Malayalam). But the cleft focus+copula can not only follow the nominalized clause, as in (19)
and (20a); it can precede the nominalized clause, as in (20b), or even “float into” it (20c).

(20) a. unn-e paarta-ada yaari ir-ka-łaam?
    you.ACC. saw-NOML. who be.INF. MOD.
    ‘Who could it be that saw you?’

   b. yaari ir-ka-łaam unn-e paarta-ada?
      who be.INF. MOD. you.ACC. saw-NOML.

c. unn-e yaari ir-ka-łaam paarta-ada?
    you.ACC. who be.INF. MOD. saw-NOML.

J generalizes the Antisymmetric structure to the cleft construction, and accounts for the three
positions of the cleft focus+copula seen in (20). M denies that the structure in question is a cleft.

---

11 A focused constituent may of course occur in a finite clause in Tamil (cf. (8) above), and may carry Q.
   (i) avan-(taan) -aa unn-e paart-aan? unn-e avan-(taan)-aa paart-aan?
      he-FOC.(EMPH.REFL.)-Q you.ACC. see-PST-3SG. you.ACC. he-FOC.(EMPH.REFL.) Q see-PST-3SG.
      ‘Did he see you?’ (‘Was it he that saw you?’)
   But such a focused constituent cannot occur post-verbally (ii).
   (ii) *unn-e paart-aan avan-(taan)-aa?  ‘Did he see you?’
   Nor can a wh-word occur post-verbally in a finite clause, as we have been at some pains to show.
She maintains that it is monoclausal, and that the copula is in C.\textsuperscript{12} If so, it is not obvious how the sentence-internal position of the focus+copula could be generated.

Finally, we may mention Jayaseelan’s argument (p.c.) from elision in question-answer pairs, which may well be relevant to our child data. We have noted in section 3.2 a preponderance, in our database of 630 wh-questions, of wh- plus verb questions (aside from copula-drop wh-questions). Now in adult Malayalam, the answer in a question-answer pair may surface as just the verb (as M notes, in favour of her analysis); or (as Jayaseelan points out in his examiner’s comments on M’s thesis), as a focused constituent (e.g., the subject, adjunct, etc.) plus the verb. The latter would, on M’s account, necessitate the deletion of a discontinuous phrase; whereas on the Antisymmetric account, it would follow from the deletion of all but the Focus Phrase (the verb moves to the head, and the focused constituent moves into the Spec, of the Focus phrase). Now if, in questions as in answers, only the Focus Phrase survives to the surface in ellipsis, the wh- plus verb child questions in our database are naturally explained: in questions as in answers, everything that is not focused may be elided. Ellipsis in questions could then be an argument that wh- moves to a Focus phrase, and is not left in situ.

6 Conclusion

Tamil acquisition data in the range of 26-29 months show that wh-subjects are consistently displaced from the canonical S position to an immediately pre-verbal focus position, but never to a post-verbal topic position. These data concur with adult intuitions about the immediately pre-verbal placement of the question word; and suggest both very early knowledge of canonical and non-canonical word orders (as in Japanese), and very early and error-free acquisition of criterial-driven movement to the peripheries. We have indicated some reasons to prefer a semantically well-motivated account of wh- as inherently focused, syntactically implemented in an Antisymmetric clause structure as movement to a pre-verbal focus position, over an account postulating across-the-board leftward topicalization of all non-wh intervenors.

Tamil and the Dravidian languages allow both left- and right-peripheral topics (with interpretive nuances, as in Turkish). Looking beyond the child data, these languages in addition allow IP-internal topics. Similarly, Foci that we analyze as cleft focus+(null) copula, can occur in any of three positions: left- or right-peripherally with respect to a nominalised clause, or (superficially and apparently) within a nominalized clause. The IP-internal Topic position and the three cleft Foci positions have so far received an account only within the Antisymmetric clause structure; which can also account for the patterns of ellipsis in questions and answers.

\textsuperscript{12} M’s proposal is that this is a mono-clausal “categorical construction,” which has a different structure than (what she calls) the ‘verb-final’ construction. Be is analyzed as an auxiliary merged in C. What we call the “cleft focus” is analyzed as a “predication base” that moves into the Spec of C.

In the interests of completeness, we may mention that J analyzes the nominalized clause as a CP complement to be. An expletive pro occupies the subject position. The verb be moves to the Focus head above vP, and the “cleft focus” is extracted to Spec of Focus. (The nominalizer D is presumably re-analysed as T to allow such extraction).

(i) Malayalam \[\text{[pro\,}\text{FOC}\,\text{Mary-ye}\,\text{[v}\,\text{aan}\,\text{aa}\,\text{CP}\,\text{iaan\,innale\,kanq-ata\,Mary-ye\,i]]}]\]

\text{Mary-ACC. be I yesterday saw.NOML.}

‘It is Mary that I saw yesterday.’

A clause-internal cleft focus+copula results when arguments move out to Topic positions above the Focus position from the complement of be. In (ii), thus, the complement subject iaan ‘I’ is topicalized.

(ii) \text{iaan\,Mary-ye\,aan\,innale\,kanq-ata.}\n
A right-peripheral focus+copula results when the entire complement clause minus Focus moves to Topic.

(iii) iaan\,innale\,kanq-ata\,Mary-ye\,aanə.\n
\[\text{[pro}\,\text{FOC}\,\text{Mary-ye}\,\text{[v}\,\text{aan}\,\text{aa}\,\text{CP}\,\text{iaan\,innale\,kanq-ata\,Mary-ye\,i]]}]\]

\text{Mary-ACC. be I yesterday saw.NOML.}
References


Clause particles and cleft sentences in Bangla: Some preliminary generalizations

Probal Dasgupta
probal.dasgupta@gmail.com

1 Introduction

This paper provides observations and some preliminary generalizations concerning the interaction of certain Bangla clause particles (Modulators) with Zero Copula Constructions. The focus is on those constructions that exhibit hitherto unnoticed clefting properties.

Section 2 surveys certain basic facts about Modulators and explores Modul’s interaction with simple examples of ZCCs (Zero Copula Constructions). Section 3 expands the ZCC analysis to explore cleft constructions, and shows that Modul plays a crucial role in certain clefts. Section 4 takes the study of particle-focus interaction beyond the boundaries of the cleft construction and beyond the Modul category. Section 5 examines the properties of interrogative constituents not just in clefts but also in an unusual construction – found in Hindi-Urdu as well – which can be designated as a ‘verb-stressed cleft’ structure. Section 6 summarizes and concludes.

2 Modulators and Zero Copula Constructions

The Bangla syntax literature has been using the term Modul[ator] for the intimacy oriented discourse particles go, re, he. A Modul like go or re, oriented to a Neu[tral] or an Int[imate] addressee respectively, appears to the right of a root sentence finite verb as in (1) and (2) or of a

---

1 This work was first presented at the Workshop on Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages held in May 2018 at Schloss Freudental, Konstanz. I thank Josef Bayer and Yvonne Viesel for hosting and all workshop-mates for stimulation and input. I also thank: Yoshio Endo for his peer review of this text; Josef Bayer for extensive comments at every stage of its gestation; Arun Ghosh, Debasri Chakrabarti, Devarati Jana, Sanjukta Ghosh, Sibansu Mukhopadhyay, Soma Paul for specific empirical input; Tista Bagchi for confirming that her 1993 account stands; Yangchen Roy for helpful comments on an earlier draft (on the basis of which the exposition has been revised) and for holding out the hope that the present study may kickstart the Malayalam-Bangla comparison which needs to follow in order for cleft and pseudo-cleft studies in South Asia to take off; Sarasij Basu for soliciting and publishing the Bangla version of this paper (Dasgupta 2019); Utpal Lahiri for inviting me to present this work to a domain-familiar South Asian audience at EFLU, Hyderabad, in November 2018; P. Madhavan and others in the audience there, for empirical and methodological input. Standard disclaimers apply.

1 The transcription used in this paper, as in much of the South Asian syntax literature, employs capital T D R for retroflex sounds, S for a palato-alveolar fricative, E O for low vowels, Y W for mid glides, M to mark nasalization for the vowel to its left, the digraph ng for the velar nasal, c j for palato-alveolar affricates. We depart from Ray et al. (1966) only in our use of the digraph where their system uses capital N. Most South Asianists use N to transcribe a retroflex nasal; resorting to the digraph is one optimal way to avoid confusion on that front.
'compact' wh phrase – an ad hoc descriptive label for size-constrained wh phrases – as in (3) and (4) below. (The far less frequently used Modul *he* is oriented to a Neu addressee but connotes disrespect; in this paper, as in the bulk of the literature, *he* is ignored.)

(1) Eka-Eka  kEno boSe  acho go alone-alone  why sitting  are.2Neu  Prt
     'Why are you sitting alone?'

(2) Eka-Eka  kEno boSe  achiS  re alone-alone  why sitting  are.2Int  Prt
     'Why are you sitting alone?'

(3) Eka-Eka  kEno  go  boSe  acho alone-alone  why  Prt sitting  are.2Neu
     'Why are you sitting alone?'

(4) Eka-Eka  kEno  re  boSe  achiS alone-alone  why  Prt sitting  are.2Int
     'Why are you sitting alone?'

One pertinent point made in earlier work (Dasgupta 2014) is that addressee intimacy level marking need not be overtly present in the clause to license a Modul. The verb in (5) and (6), where the Neu or Int Modul invokes a Neu or Int addressee respectively, agrees with a non-addressee argument in both these sentences and copies its value for the Hon[orific] feature (from this point onwards, we will simply use the gloss ‘Prt’ or ‘Modul’, leaving it as an exercise for the reader to tell the ‘Prt.Neu’ or ‘Modul.Neu’ /go/ from the ‘Prt.Int’ or ‘Modul.Int’ /re/, as nothing in the argumentation turns on this contrast):

(5) kEno  go  TiToda  rag  korechen why  Prt.Neu  Tito.senior angry Aux.3Hon
     ‘Why is Tito [a male who is senior to the speaker] angry?’

(6) kEno  re  TiToda  rag  korechen why  Prt.Int  Tito.senior angry Aux.3Hon
     ‘Why is Tito [a senior male] angry?’

Dasgupta (2014) also notes, but does not explore, the fact that zero copula constructions (ZCCs), which in Bangla show no overt agreement, can license a Modul, as in (7). We see at (8) that the predicate of a ZCC, when displaced, carries the Modul with it:

(7) tomar  boktritar  mul  kOthaTa  iMTer  mOto  SOkto  go your  talk’s  main  point  brick  like  hard  Prt
     ‘The main point of your talk is as opaque as a brick’
As opaque as a brick, that’s what the main point of your talk is like.

In this paper, we take a preliminary look at Modul-ZCC interaction in the subtypes we propose to call property ZCCs, event ZCCs, and conjunctival ZCCs, exemplified at (9)-(12), (13)-(14), and (15)-(18) respectively; it is to be hoped that these terms are self-explanatory. The ‘property ZCC’ and ‘event ZCC’ types are familiar from other languages. It is possible that Modul’s privileges of occurrence exemplify significant generalizations; but at the present juncture the study of these generalizations does not seem a promising line of inquiry; so we simply offer a few examples to give readers a sense of the lie of land:

Property ZCCs:

(9) golmaler jonne ke dayi re trouble for who responsible Prt

‘Who is responsible for the trouble?’

(10) golmaler jonne dayi ke re trouble for responsible who Prt

‘Who is responsible for the trouble?’

(11) *ke re golmaler jonne dayi who Prt trouble for responsible

(12) *golmaler jonne ke re dayi trouble for who Prt responsible

Event ZCCs:

(13) diliper biye kar SOnge re Dilip’s wedding who with Prt

‘Who is Dilip getting married to?’

(14) kar SOnge re diliper biye who with Prt Dilip’s wedding

‘Who is Dilip getting married to?’

The truly distinctive type here called the ‘conjunctival ZCC’ displays several puzzling properties:

Conjunctival ZCCs:

(15) ke re okhane daMRiye who Prt there standing

‘Who is standing there?’
The behaviour of Modul in these sentences is hardly their most interesting trait. It is characteristics (19)a-d of Modul-free conjunctival ZCCs that most saliently distinguish them from other constructions. We present the following observations in a non-standard format as this digression falls outside the scope of the paper per se. The points are being made in the hope that other workers will find it appropriate to study the phenomena more elaborately in future:

(19) **Main characteristics of the Conjunctival ZCC construction:**

(a) unlike property ZCCs and event ZCCs (where copula omission is obligatory), conjunctival ZCCs only optionally omit the copula; one can insert /ache/ ‘is’ right after the conjunctive participle in (15)-(18) and obtain a well-formed sentence;

(b) in this construction, the conjunctive participle must be monolexemic: even though /daMRiye/ ‘standing’ often means ‘waiting’ in Bangla, nonetheless one cannot replace /daMRiye/ with the bilexemic /Opekkha kore/ ‘wait Aux = waiting’;

(c) even within the class of monolexemic verbs, only some verbs turn out to be admissible in the conjunctival ZCC: in the sentences (i) /aYnaTa Ekhono beMke ache/ ‘mirror the still bent is = the mirror is still bent’, (ii) /paYer angul Ekhono phule ache/ ‘foot’s finger still swollen is = the toe is still swollen’, (iii) /ora holi khElAY mete ache/ ‘they Holi play.Loc taken.up are = they are still immersed in the revelry of Holi’, (iv) /baccaTa khElNa niye bhule ache/ ‘child.the toy with absorbed is = the child is absorbed with her toys’, it is not possible to omit the copula /ache/; comparison with verbs acceptable in the conjunctival ZCC, like (v) /egiye/ ‘advanced’, (vi) /takiye/ ‘looking’, (vii) /SuYe/ ‘lying.down’, (viii) /boSe/ ‘sitting’, (ix) /jege/ ‘awake’, (x) /ghumiye/ ‘sleeping’, suggests that the delimitation criteria are semantic in nature, though no obvious specification of these criteria immediately comes to mind;

(d) the language has several nonfinite forms that can co-occur with a copula; among these, however, it is only the conjunctival (conjunctive-aspect-marked) adverbial participle that has

---

2 I survey-quizzed twenty-odd Bangla-speaking linguists, to make sure I was not imprisoned in idiosyncratic judgments. But the paper then veered off in a direction that sidelined the whole subdomain; thus, it would be pointless to thank them by name; I am grateful to them all.

3 The conjunctive form /bhule/ literally means ‘forgetting’, here encoding the point that the child forgets everything else when she is engrossed in play.

4 I am indebted to Soma Paul, Debasri Chakrabarti and Sanjukta Ghosh for examining the conjunctival ZCC facts independently. They confirm (a) that only monolexemics are possible for them, and (b) that they also find the characterization problem intractable in this form.
From the conduct of Modul in the types of ZCC we have just surveyed we can conclude that there are no significant facts to report on the Modul front in the simple structures to which our survey has been confined. There is more to say, however, once we expand the data base in an unexpected direction, which we shall do from section 3 onwards.

3 Modul and the cleft construction

There is a pseudo-cleft construction in Bangla; as expected, it features a relative structure and a copula. This Bangla pseudo-cleft was first identified and briefly discussed by Bagchi (1993), in whose work the role of the copula was not a high priority issue. Since the Bangla copula is standardly zero, the pseudo-cleft usually instantiates the ZCC, as in (20). However, the language also harbours the Positive Polarity Copula (PPC) construction described in Dasgupta (2003). The pseudo-cleft sometimes deploys the PPC rather than a zero copula, a phenomenon exemplified in example (21). This example happens to use the Present Progressive variant of the PPC rather than /holo/ ‘PPC.Past’, the variant that bears Simple Past morphology; it does not matter which of the variants of the PPC is used in a particular sentence; they are syntactically and semantically identical. Likewise, the relative pronoun used in (20) and (21) is the generic ja; in such sentences, some speakers strongly prefer the pronoun jeTa, closer to ‘which’ than to ‘what’; nothing in the argument turns on pronoun choice:

(20) chobi ja caY SeTa pOroTa
    Chhabi what wants that paratha
    ‘What Chhabi wants is a paratha’

(21) chobi ja caY SeTa hocche pOroTa
    Chhabi what wants that PPC.Pres.Prog paratha
    ‘What Chhabi wants is a paratha’

New issues arise when we turn from the correlative structure instantiated in (20) and (21) to the structure exemplified in (22), where the clause particle /je/ (and through it the entire clause ‘that Dipti is angry’) is coindexed with the sequent pronoun /SeTa/ ‘that’. The gloss provided at (22) is in keeping with the preference for preposition stranding in idiomatic English – a more literal gloss would have been ‘it is about this unexpected delay that Dipti is getting annoyed’. Notice that when we gloss (22) in English it is natural here to use a ‘true cleft’ and not a ‘pseudo-cleft’ gloss. Throughout this paper, unless we specify otherwise, the term cleft is used in its broad sense, which includes pseudo-clefts:

(22) dipti je birokto hocche SeTa ey OprottaSito deri niye
    Dipti Prt annoyed Aux that this unexpected delay about
    ‘It is this unexpected delay that Dipti is getting annoyed about’
One can ‘intensify’ the clefting effect (i.e. the focusing) by adding an Emphasizer /i/ to the focus phrase in (22), thus obtaining (23). Pending a dissertation-level investigation of such foci, not yet available for Bangla or other modern Indo-Aryan languages, I venture to conjecture that without /i/ the focus in (22) is (using É. Kiss’s (1998) terminology) an **identificational** focus that is **exhaustive**, and that adding /i/ makes it **contrastive** as well in (23); this is my tentative unpacking of our informal remark that adding the Emphasizer /i/ ‘intensifies’ the focal character of the phrase. (Applying the standard diagnostic tests strengthens my conjecture, but trying to show this right here would involve running the risk of multiplying digressions beyond even my elastic sense of necessity; later workers who choose to translate the generalizations offered here into a formally explicit account will do the exercise anyway.) Throughout this study, whenever we use the term ‘focus’, we ordinarily mean an identificational focus:

(23) dipti je birokto hocche SeTa ey OprottaSito deri niye  
Dipti Prt annoyed Aux that this unexpected delay about.Emp  
‘It is precisely this unexpected delay that Dipti is getting annoyed about’

When this focus expression /ey deri niye(i)/ ‘about (precisely) this delay’, with or without the Emphasizer /i/ but carrying focal intonation, appears to the right of a simple clause, despite the absence of the particle /je/ and the sequent pronoun /SeTa/, the cleft interpretation persists, as we see at (24) and (25):

(24) dipti birokto hocche ey OprottaSito deri niye  
Dipti annoyed Aux this unexpected delay about  
‘It is this unexpected delay that Dipti is getting annoyed about’

(25) dipti birokto hocche ey OprottaSito deri niyei  
Dipti annoyed Aux this unexpected delay about.Emph  
‘It is precisely this unexpected delay that Dipti is getting annoyed about’

We claim that the syntax of (24) and (25) structurally supports this interpretation. In particular, we claim that (24) and (25) have a zero copula; that /ey OprottaSito deri niye(i)/ ‘about (precisely) this unexpected delay’ counts as the predicate of this ZCC; and that the pre-focus material /dipti birokto hocche/ ‘Dipti is getting annoyed’ is a gap-laden open sentence whose gap is associated with the filler /ey OprottaSito deri niye(i)/, ensuring that this filler counts not just as a predicate but as a specifical predicate. For concreteness, we take it that /dipti birokto hocche/ harbours a covert wh element of the sort diagnosed by Dasgupta (2016) in pre-demonstrative relative clauses. This assumption helps make sense of the fact that the pre-focus material in (25) can be fortified either by adding a full sequent pronoun /SeTa/ ‘that’ as in (26) or by adding just a nominal classifier /Ta/, here amounting to something half-way between a nominal element and a clause particle, as in (27), a sentence that exemplifies a particularly casual style of speech and does not represent the norms of written prose. Later in the paper we shall have occasion to return to uses of this /Ta/ element that have breached the speech-prose barrier.
With these preliminary considerations on the table, we are now ready to place the Modul category in the context of the cleft instantiations of the ZCC. Consider sentence (28):

(28) carTe bidghuTe upoma khuMje peYechiS kon boyer SeS pataY re four bizarre tropes found have which book's last page.Loc Prt

‘On the last page of which book is it that you have found four bizarre tropes?’

We are concentrating on (28), the variant that eschews all the overt devices. But it is possible in (28), exactly as in (23), to insert the clause particle /je/ and the sequent pronoun /SeTa/. We also have the option of tweaking (28) along the lines of (26), by inserting just the pronoun /SeTa/ without a supporting particle /je/, or even to mimic (27) and insert the bare classifier /Ta/. We save space by not exhibiting all those variants. Our point is about the unavailability of (29), in contrast to the well-formedness of (6), repeated as (30) (we are using ‘huh’ in the gloss for (29) to give a very rough sense of the flavour of the ill-formedness involved):

(29) *kon boyer SeS pataY re carTe bidghuTe upoma khuMje peYechiS which book's last page.Loc Prt four bizarre tropes found have **‘On the last page of which book, huh, have you found four bizarre tropes?’

(30) kEno re TiToda rag korechen why Prt.Int Tito.senior angry Aux.3Hon

‘Why is Tito angry?’

The point we wish to stress is that a wh plus Modul sequence like /kEno re/ ‘why Prt.Int’ in (30) is possible only if the wh-expression is ‘compact’ in some sense that has proved difficult to pin down exactly (‘a single word’ is a reasonable approximation but fails for certain outlier examples). This is the pattern instantiated at (29). Against this background, the availability of the ‘diffuse’ wh plus Modul sequence /kon boyer SeS pataY re/ ‘which book’s last page.Loc Prt’ in (28) but not in (29) calls for explanation. Our conjecture that the cleft construction finds that sequence a valid niche in (28) provides the explanation required.

On this account, we expect – and we find – that the word-string shared by (31) and (32), featuring a compact wh-word plus Modul, has two different structures associated with distinct intonations. In contrast, we expect (and find) that the word-string familiar from (28), shared by (33) and (34), is acceptable only when intoned and interpreted as in (34). For only in the cleft structure (34) can the diffuse interrogative plus Modul sequence find a valid niche:
In the discussion of the behaviour of Modul in these constructions, we have simply accepted the fact that Modul appears in final position; and we have been looking at the expanded ZCC alone. In order to improve our understanding of the big picture, we need to do better than this.

4 Modul and other particles in focal contexts

Consider (35)-(37), a build-up involving 'light' and 'heavy' postverbal material (note that we refrain from using the restrictive terms 'compact' and 'diffuse' here, since 'light items quite often breach the one-word ceiling). In (35), the 'light' phrase /eTa niye/ 'about this' is acceptable in postverbal position, and the Modul /re/ can occur only in immediate adjacency to the verb; the variant that places Modul in clause-final position is perceptibly worse, as shown within brackets.

(35) dipti birokto hocche re eTa niye [??eTa niye re] Dipti annoyed Aux Modul this about [??this about Modul]

‘Dipti is getting annoyed, you know, about this’

In (36) and (37), as the postverbal material gets heavier, the string’s acceptability keeps diminishing. As in the case of (35), the variant placing Modul at the very end of the sentence continues to be sharply ill-formed (in (37), we abbreviate the word /OprottaSito/ and its gloss 'unexpected' to ensure easy inspectability of the alignment between the Bangla words and their glosses):

(36) ?dipti birokto hocche re ey deriTa niye ["ey deriTa niye re] Dipti annoyed Aux Modul this delay.Cl about [this delay.Cl about Modul]

‘Dipti is getting annoyed, you know, about this delay’
Dipti is getting annoyed, you know, about this unexpected delay.

However, when we switch over to an intonation contour that treats the postverbal material as focal rather than as leaked, the story changes dramatically, as we see at (38). On the form side of the equation, Modul must occur immediately to the right of the finite verb, not to the right of the postverbal phrase. At the interpretive level, the sentence elicits a cleft reading. We save space by not showing focal counterparts to (36) and (37); they have the same properties:

(38) Dipti annoyed Aux *this about.Emph Modul [Modul *this about.Emph]

It is about this, you know, that Dipti is getting annoyed'

To stop readers from imagining that the sequence 'focus phrase plus Modul' per se invariably triggers a cleft sentence reading, we need to introduce one important fact that deflates this erroneous impression. Recall the PPC construction exemplified at (21) in section 3: we provide a fresh example, (39). Inserting a Modul to the right of the focus, as in (40), is fine. We save space by not showing explicitly that placing Modul between the PPC /hocche/ and the focus /diliper praner bondhu/ 'Dilip’s bosom friend' leads to ungrammaticality.

(39) Pushan’s neighbour Tridib is a bosom friend of Dilip’s'

(40) Pushan’s neighbour Tridib is a bosom friend of Dilip’s'

The phenomena are reminiscent of the clefts discussed earlier in this section, but these examples of the Positive Polarity Copula Construction emphatically do not instantiate clefting. The point is that a postverbal constituent plus Modul sequence must have two properties: first, the constituent must be a focus (if it is a low key ‘leaked’ item, the sentence crashes); second, it requires support from some copula, either the null copula of a cleft ZCC or the overt copula of a PPCC.

The PPCC is not the only pertinent construction that is familiar from earlier work. Consider the quirky modal particle /ba/, exemplified in (41), that was highlighted in the Q-ba study (Dasgupta 2005). This quirky particle’s pertinence to the issue of cleft sentences becomes apparent in (42), where a sequence consisting of the focused constituent plus /ba/, placed in postverbal position, triggers a cleft interpretation. This particle /ba/ counts as quirky because of its restricted distribution. It selects either a special subjunctive (see note 5) or an interrogative laden with the emphazizer particle /i/ (as in these examples). Note that only in the company of /ba/ does an interrogative constituent in Bangla ever carry such a focus particle:
Clause particles and cleft sentences in Bangla: Some preliminary generalizations

Exploring further the place of these quirky phenomena in the overall picture that has been emerging from our discussion, we naturally proceed to ask how Modul interacts with this Q-ba construction. The moment we look at the relevant examples, (43) and (44) – where, as in the case of (29) earlier, we use ‘huh’ as a somewhat lame gloss for Modul, since the usual ‘you know’ sounds incongruous in this context in English – we run into the fact that, although (43) shows that Modul can in principle occur in the same clause as a quirky /ba/ particle, nevertheless (44) makes the point that it cannot share interrogative-contiguous space with that element – a phenomenon of some syntactic interest that is tangential to our main narrative here, so we relegate it to a footnote.5

(41) tora kake-i ba Dakte ceYechili?
you.Pl whom-Emph Prt to.invite wanted
‘Who indeed did you want to invite?’

(42) tora Dakte ceYechili kake-i ba?
you.Pl to.invite wanted whom-Emph Prt
‘Who was it indeed that you wanted to invite?’

The big news for us is that the good sentence (42) is exactly like the irrelevantly bad sentence (44) in that it has a cleft reading; and that the cleft reading is associated with the sequence consisting of a postverbal focused constituent and a clause particle. The fact that this clause particle is a quirky /ba/ in (42), rather than a Modul, shows that we cannot afford to restrict our story to Modul per se. Notice that we cannot wish the problem away by redescribing the quirky /ba/ as a Modul: example (43) makes that move impossible – two Moduls never occur in one and the same sentence. Once we start stretching the generalization beyond the Modul category, we are led to ask whether it is just /ba/, or other clause particles as well. The obvious place to look is Emphasizers like /i/ ‘indeed’ and Anchors like /to/ ‘of course’. Consider examples (45)-(47), which feature these elements, always placing the emphasis within the scope of the anchoring:

(43) tora kake-i ba Dakte ceYechili re?
you.Pl whom-Emph Prt to.invite wanted Modul
‘Who indeed did you want to invite, huh?’

(44) *tora Dakte ceYechili kake-i ba re?
you.Pl to.invite wanted whom-Emph Prt Modul
‘Who was it indeed that you wanted to invite, huh?’

5 It pays to look closely at the contrast in this respect between the Q-ba (interrogative plus /ba/) construction examplified in the text and the past subjunctive plus /ba/ construction that we find in (i) /dilip badam khelo-i ba/ ‘Dilip nuts eat.PaSbj-Emph Prt = it’s okay if Dilip does eat nuts’. Adding Modul to (i) produces the only moderately degraded (ii) ??/dilip badam khelo-i ba re/ ‘Dilip nuts eat.PaSbj-Emph Prt Modul = it’s okay, you know, if Dilip does eat nuts’. We conjecture that the finite verb, which supports the Emph plus /ba/ cluster, and which in general supports Modul, is just about able to carry the double burden in (ii). (43) is perfect because the verb directly handles Modul while the interrogative handles /ba/ with indirect support from the verb-associated functional projections. We take it that in (44), where the interrogative is, all by itself, called upon to carry the double burden by virtue of its indirect association with these projections, the arrangement falls apart because no verb-contiguous item activates this indirect support system. I end this note by thanking Arun Ghosh, who confirmed my judgments in this seldom studied domain. My basic account of /ba/ appears in Dasgupta (2005), where the lame gloss ‘huh’ does not rear its ugly head.
As for Raka, it was of course Dilip who invited her.

As for inviting Raka, it was of course Dilip who did it.

As for inviting Raka, it was Dilip who did it, of course.

However, even when the Emph-laden constituent occurs to the left of the finite verb, as in (48), the focal intonation plus the Emphasizer /i/ ensure that the constituent, here /dilip/, receives a focal interpretation. It then becomes a matter of convention whether one insists on glossing such a sentence as in (48), eschewing the cleft strategy in the English gloss, or whether one chooses a cleft gloss like ‘As for Raka, course, it was Dilip who invited her’. Only comparative evidence from South Asian languages with a clearly identifiable clefting strategy, like Malayalam⁶, will help us to diagnose (48), in Bangla syntax proper, as formally cleft or not. We leave that task to competent others, and meanwhile we recall one crucial fact about the interaction of the Emphasizer /i/ and the Anchor /to/ – namely, that the placement of the anchoring within the scope of the emphasis, as in (49), is completely unacceptable:

As for Raka, of course, Dilip invited her.

It was Dilip who, as for Raka, invited her.

Even if we leave pending the issue of whether an Emph-laden focal constituent in situ counts as a cleft focus in the narrow syntax of Bangla, the discussion has established that a postverbal constituent that is focally intoned and/or bears the Emph /i/ indubitably enjoys that status. To return to our main concern, such a sentence places its Modul at the very end, to the right of this cleft focus, as in (50):

It is your father, though, that they will blame, you know.

---

⁶ I thank Yangchen Roy, a linguist who happens to be a native speaker of both these languages, for following the gestation of this paper closely and for undertaking to do the follow-up required. It is to leave the domain uncluttered for future work by workers like her that the present paper concentrates on empirical generalizations and formulates its hypotheses tentatively and without insistent formal explicitness regarding their details.
Given the fact that interrogation bears intrinsic focus, we are not surprised to find that a postverbal interrogative phrase – we consider the case where Modul is present, though it need not be – also counts as a cleft focus, as in (51), where the English gloss is again forced to make do with the lame ‘huh’ because idiomatic usage precludes ‘you know’ in such a context:

(51) orā dos dicche kon cheleTake re
they blame Aux which boy.Cla.Acc Prt

‘Which boy is it that they are blaming, huh?’

We recognize (51) as instantiating the same pattern as (34) and realize that we are now ready to end this section, having reiterated the point at which section 3 came to a natural close. Before we actually close section 4, in order to clarify exactly the content of our suggestion that comparative evidence from Malayalam will make it possible to unpack our approach to Bangla, we present one example of a typical Malayalam pseudo-cleft. This example is from Madhavan (1987: 105), who gives it the serial number 35. We label it as (M35), ‘M’ for Madhavan, to hold it at arm’s length from the sequencing of our own examples. Note that underscored t, d, n are alveolars (this typologically unusual language has three distinct sets of coronal stops: dental t d n, alveolar t d n, retroflex T D N). The nominalizing /a/ suffix of the ‘wear’ verb carries neuter third person plural agreement with ‘new clothes’, the focus, which appears immediately before the copula:

(M35) nāan dhari-ccirkkunna-wa
I.Nom wear.Pres.Prog-Neuter3Plur putiya wastraṇṇaLa aaNa

‘What I am wearing are new clothes’

In Bangla, what corresponds to the Malayalam ‘nominalized clause’ is a relative clause with a phonologically null relative pronoun; and the copula at the very end of the structure in Bangla is also silent. Modulo these differences, which pertain to the typological profiles of the two languages (relative clauses with an overt relative pronoun are a highly marginal structure in Malayalam, but are a major characteristic of Bangla), the Bangla pseudo-cleft matches the Malayalam pseudo-cleft. Thanks to Madhavan’s work, the latter is well understood, which is why we maintain that comparing the two languages is likely to yield serious illumination.

5 Clefting, interrogation, and discourse particles

In this section, we begin by articulating our analysis of Bangla cleft structures. We then argue that the language harbours two sharply distinct clause types involving interrogatives and discourse particles – one of them a cleft construction, the other a distinct construction that exhibits an intriguing cluster of properties and merits the unusual characterization ‘anti-cleft’.

Our exposition so far has outlined the following picture, some of whose details remain to be made precise. In the garden variety Bangla cleft construction, the prefocal sequence, which we shall call the preamble, looks like a simple clause with a gap in it, but in fact has the syntax of a relative clause with a null relative constituent along the lines of the explicit analysis given in Dasgupta

\footnote{Massam (2017) offers a wh-free descriptive strategy that could, mutatis mutandis, be brought to bear on our material. A priori I have no reason to prefer it to the one I promote in the main text, but it is only fair to...}
(2016). The focal material, with or without a supporting particle to its right, counts as the cleft focus, linked to the relative clause preamble through the specificationally interpreted copula, which happens to be null in a Bangla cleft.

In the summary just given, some readers may find the opaque reference to Dasgupta (2016) offputting. To plug that hole, we reiterate that the relevant part of the argument in that paper had to do with pre-demonstrative relatives such as (52), featuring an overt relative pronoun (here /ja/ ‘which’), and with the fact that Bangla allows the omission of that pronoun – and of associated material (here the adposition /diye/ ‘with’) – yielding the alternative structure (53) that an abbreviated characterization can describe in terms of a null relative constituent. Readers will find a considerably more nuanced account in Dasgupta 2016, one that keeps in full view the fact that no naïve analysis that literally postulates a null relative can find a theoretically defensible way around the violation of the ‘recoverability of deletion’ principle; for our limited purposes here, we can afford to cut corners:

\[(52) \text{ja diye hire kaTte paren Emon churi} \]
\[\text{which with diamond cut can such knife} \]
\[\text{‘such a knife with which she can cut diamonds’} \]
\[(53) \text{hire kaTte paren Emon churi} \]
\[\text{diamond cut can such knife} \]
\[\text{‘such a knife with which she can cut diamonds’} \]

This account helps make sense of the affirmative cleft (50) and of its interrogative counterpart (51), here repeated as (54) to keep the example numbering continuous:

\[(54) \text{ora doS dicche kon cheleTake re} \]
\[\text{they blame Aux which boy.Cla.Acc Prt} \]
\[\text{‘Which boy is it that they are blaming, huh?’} \]

However, when the postverbal interrogative is not stressed, and when an optional DiP (Discourse Particle) Ta intervenes between the preamble and the postverbal material, as in (55), none of the elements of the account we have been constructing look applicable. The focal stress in (55) – and in its DiP-omitted but synonymous variant (56) – is on the verb /jacche/ ‘is.going’, not on the interrogative /kothaY/ ‘where’, which carries low-key intonation. We must note yet another unusual formal characteristic: this construction requires the post-DiP interrogative to be ‘compact’ in the sense explained in the early sections of this paper. And the interpretation has an extra quirk to it as well – as we indicate in the gloss, there is an element of aggressive non-D-linking associated with this construction, which we shall call the Ta-Q construction for convenience of reference:

\[(55) \text{ora doS dicche kon cheleTake re} \]
\[\text{they blame Aux which boy.Cla.Acc Prt} \]
\[\text{‘Which boy is it that they are blaming, huh?’} \]
Readers entirely unfamiliar with the facts of Bangla will want to know if the DiP /Ta/ of (55) can be equated with the bare classifier /Ta/ found in sentence (27) in section 3. The answer is slightly long-winded. The two /Ta/’s are certainly not accidental homonyms. The basic use of the classifier /Ta/ is in numeral-noun constructions like (57) and (58). Its secondary occurrences are found in definite constructions like (59), where the numeral along with the classifier is encliticized to the noun, and (60), where the numeral disappears and leaves the noun directly interacting with the enclitic classifier (that we show a hyphen in (59) but not in (60) reflects the orthographic conventions of Bangla). From (60) it is but a short step to (61), where the gerund /berono/ ‘leaving’, being a verbal noun, counts as a noun and can take a definitizer, but also counts as a verb and lends /Ta/ some of the properties of a not-quite-nominal particle:

(55) dilip jaccheTa kothaY
Dilip is.going.Prt where
‘Where on earth is Dilip going?’

(56) dilip jacche kothaY
Dilip is.going where
‘Where on earth is Dilip going?’

(57) EkTa boy
one.Cl book
‘a book’

(58) paMcTa boy
five.Cla book
‘five books’

(59) boy-paMcTa
book-five.Cla
‘the five books’

(60) boyTa
book.Cla
‘the book’

(61) Oto bhore berono(Ta) bhul hoYeche
so early leaving(Prt) wrong was
‘Leaving so early was a mistake’

The exploratory observational study by Jana (2018) systematically surveys occurrences of /Ta/, including some that are intermediate between its canonical use as a nominal classifier and its extended use as a discourse particle. What we find at (55) is certainly the DiP use of /Ta/ (we write /jaccheTa/ ‘is.going-Prt’ as a single word on account of orthographic conventions). Obviously a single conceptual leap does not get us from (61) to (55) – there are intermediate steps, some of which become clear in Jana’s data. We set aside the issue of just how the classifier /Ta/ and the DiP /Ta/ are connected; suffice it to note that they are.
We return now to the real question we face at this point in our exposition: what is going on at (55)? What is there to say about the construction instantiated here and at (56)?

While a defensible formal analysis may prove elusive for some time, we are able to point to a comparable phenomenon in the sister language Hindi-Urdu, where an ordinary question like (62) can be tweaked to yield variant (63) in which the verb proper moves to a focus position across a mandatorily compact interrogative constituent but leaves the auxiliary cluster stranded:

(62) ye Saadii kaháàM ho rahii hai
    this wedding where happen Prog Pres
    'Where is this wedding happening?'

(63) ye Saadii hó kahaaM rahii hai
    this wedding happen where Prog Pres
    'Where on earth is this wedding happening?'

The configuration of stress, constituent placement, compactness and so on in the Hindi-Urdu construction exemplified at (63) is exactly as in Bangla, modulo the different handling of the verb-auxiliary complex, a difference that Indo-Aryanists have often encountered while studying emphatic particle placement and other issues. Quite generally, where Bangla pied-pipes the inflectional material with the verb, Hindi-Urdu strands the auxiliary cluster.

We tentatively conjecture that the Ta-Q construction in Bangla does indeed instantiate an expanded ZCC – and that the preamble is a null relative, as in the standard clefts studied earlier – but that the unfocused and compact character of the interrogative constituent, coupled with the fact that having a postverbal interrogative is a defining characteristic of the construction, as is the movement to focus placing the verb in the middle of the clause, makes Ta-Q a syntactically and semantically untypical construction, leading us to treat this verb-stressed ‘anti-cleft’ construction as something of an outlier. Given the demographic fact that Hindi-Urdu constructions receive serious collective attention before their Bangla counterparts do, (63) may become the object of rigorous study soon, whereupon one can bring the results of such inquiry to bear on (55) and (56). Every effort had been made, when Dasgupta (2016) was being drafted, to obtain comparative insights from Hindi-Urdu, but those competent to provide such help did not find the time to answer questions, beyond remarking that the arguments in that paper largely carry over to Hindi-Urdu. If this is so, conceivably even the pseudo-cleft structure proposed for post-verbal heavy constituents in the present study might prove applicable to several sister languages.

Adding a Modul to the structures of current interest yields well-formed outcomes, of course:

(64) dilip jaccheTa kothaY re
    Dilip is.going.Prt where Modul
    'Where on earth is Dilip going, huh?'

---

For example, one cannot replace /kahaaM/ ‘where’ with /kis Sahar meM/ ‘which city Loc = in which city’. The transcription we use for Hindi-Urdu differs from Bangla only in the use of double vowels as a notation for vowel length, and in the use of /ai, au/ for diphthongs.
The point, of course, is that Modul cannot be niched between the focalized verb and the backgrounded interrogative in Bangla. In Hindi-Urdu, one sees clearly just where the Modul is placed – what one gets is (66), where Modul immediately follows the stranded Aux cluster:

(66) ye Saadii hó kahaaM rahii hai re this wedding hāppen where Prog Pres Modul

‘Where on earth is this wedding happening, huh?’

We are now in a position to summarize what we have done.

6 Summary and conclusion

At the outset, we introduced the descriptive category Modul, placed it within the broader class of discourse particles, and identified some of its distinctive properties. We then took a preliminary look at the way Modul interacts with the Zero Copula Construction (ZCC), an issue neglected in the earlier literature. When we considered examples of a postverbal constituent immediately followed by Modul, we found that those sentences have a cleft flavour to them. We tentatively proposed that they have a pseudo-cleft structure, with the null copula that is a standard feature of the language playing in this construction the role that an overt copula plays in familiar pseudo-cleft constructions elsewhere. Thus, we proposed that a pseudo-cleft is an expanded ZCC. To flesh out such a proposal, we linked our account to earlier work suggesting that Bangla allows null relative pronouns in left relative clauses of certain types. We finally examined certain points of interaction between these phenomena and interrogation, noting an unusual structure (involving V moving to focus, past a compact and invariably clause-final interrogative) that we decided to characterize as an outlier ‘anti-cleft’ construction; we also found a very similar verb-focused structure occurring in Hindi-Urdu, and exhibiting interesting parametric differences that will help underwrite a serious comparative account when the time comes.

References


Exploring right/left peripheries: Expressive meanings in questions

Yoshio Endo
endoling@gmail.com

1 Background

Chomsky (2001) argues that the CP zone is characterized by scope/discourse properties. Based on this idea, Rizzi (1997, 2001, 2004) claims that there are various functional heads in the CP zone for topic, focus, etc., as shown in (1) and the associated phrasal elements are ordered in the left periphery. Abels (2012) argues that left-peripheral structures like (1) should be derived from some appropriate theory of locality on chain.

(1) Force Top Int Top* Focus Mod* Top* Fin IP (Rizzi 2004)

In this paper, I will discuss various head elements ordered in the right periphery in Japanese with reference to the corresponding German cases, where special attention is paid to sentence final particles (SFPs) that contributes to creating expressive meanings. I will show that strictly ordered head elements in the right-periphery that create expressive meanings are also constrained by a theory of locality.

2 What are expressive meanings?

As Cecilia Poletto (personal communication) points out, the sequence what...for like (1) conveys the kind of expressive meaning which could be seen as an instance of what Obenauer (2006) terms a “surprise disapproval question” (SDQ). (See Endo (2015) for SDQ questions.)

(2) What are you coming to the United States for?

This paper was presented at Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages held at Konstanz, Germany in May, 2018. I am grateful to Josef Bayer for inviting me to this wonderful workshop and encouraged me to include the comic pictures in the paper that I used at the workshop. I am also grateful to the audience, especially Masayuki Ohishi, Norio Nasu and Probal Dasgupta, and a reviewer for invaluable questions and comments. Special thanks go to Cecilia Poletto, Marcel den Dikken, Luigi Rizzi, Ur Shlosnky and Andrew Radford for discussing the topic of this paper. This research is funded by a Grant-in-Aid for Scientific Research (C) of the Japan Society for the Promotion of Science (Project 16K02639).

1 According to Obenauer, SDQs and the what for questions differ in that what for questions do not necessarily require a surprise disapproval interpretation.
This type of expressive meaning is also seen in *how come* questions, another wh-expression asking for reasons, as illustrated in (3). (See Endo (2018) and Radford (2018) for *how come* questions.)

(3)  

How come the sky is blue?

Tsai (2008: 89), who attributes the observation to Andrew Simpson (personal communication), mentions the expressive meaning of *how come* questions and *why*: *why* involves no special expectation about whether or not state of affairs should hold, whereas *how come* expresses surprise that a particular state of affairs should hold. However, Andrew Radford (personal communication) notes that it is not the case that *how come* questions always express surprise that a particular state of affairs should hold. For instance, there is no surprise in what B says in (4), just curiosity and *how come* sounds less invasive than *why* in this context. I will discuss this difference in *how come* questions in 3.3.

(4)  

A: I’ve gotta go to the doctor this afternoon  
B: How come?  
A: Oh, the cut on my finger has got infected

With this background of expressive meanings in mind, I will turn to expressive meanings created by SFPs in Japanese in the right periphery. (See Endo 2007, 2012 for the nature of SFPs in Japanese.) Because the properties of SFPs in Japanese are less well-known, let me first introduce some basic properties of SFPs in Japanese that are relevant in the following discussion by using the comic Peanuts in the next section. (See Trotzke (2017) for expressive meanings created by discourse particles (DiPs) in German, which he calls focus of emphasis.)

### 3 Functional elements responsible for expressive meanings

*(A) Na*

The semantic import of the SFP *na* is to signal the expression of weak confirmation by the speaker about the proposition (cf. Uyeno 1971). This particle is *speaker-oriented* in that it can be used in a monologue, and thus has the feature specification [+speaker, -addressee]. To see the point, consider the sentence uttered by Charlie Brown in (5a) and the corresponding Japanese translation in (5b), where the sentence is suffixed by the SFP *na*:

(5)  

a. I wonder if it’s possible to be in love with two different girls at the same time.  

a’. *Doozi-ni futari-no ko-o aisurunante kanoo ka-na*  

at.the.same.time two-Gen girls-Acc love possible Q-SFP

---

2 A reviewer points out that certain particles in the head-final South Asian language Bangla, specifically *re* and *go*, occur either in sentence-final position or immediately to the right of an interrogative constituent in any position within the sentence. In fact, some of the SFPs in Japanese, especially *ne* and *sa*, may also appear in the clause middle position by suffixing to a phrase, as exemplified in (18).

3 All of the sentences and pictures in this paper were borrowed from Schultz (2008) and the Japanese sentences are translated by Syuntaro Tanigawa.
Here, Charlie Brown talks to Snoopy without looking at him as if he were talking to himself, where the Charlie’s sentence in Japanese in (5a’) is suffixed by the SFP na preceded by the Q particle ka. When the SFP na immediately follows the question-type particle ka to form the sequence ka-na, it carries an expressive meaning of worry, where the speaker evaluates the proposition with worry. (cf. Hirayama 2015 for the expression ka-na.)

The expressive meaning created by the sequence ka-na seems similar to what Obenauer (2004) and Bayer and Obenauer (2011) call an “I-Can’t-find-the-Value-of-x Question (CfvQ) in German, where the DIPs nur and bloß are used in the clause middle position. In CfvQ, the speaker signals that he/she has so far unsuccessfully tried to find an answer as illustrated in (6a), where the sentence does not require the addressee. This German sentence is translated into Japanese by suffixing the speaker-oriented particles na to the Q-particle ka, as shown in (6b). (cf. Endo 2012 for the similarities in DIPs between Japanese German.)

(6) a. Wo liegt nur / bloß meine Brille? (CfvQ)
   where lies NUR/ BLOSS my glasses
   ‘Where on earth did I put my glasses? (I have already looked everywhere)

b. Are, megane-o doko-ni oita ka-na?
Oh, glasses-Acc where-at put Q-male

As a reviewer correctly points out, the feature matrix [+speaker, –addressee] indicates “not requiring the addressee”, i.e. self-sufficient confidence on the part of the speaker.
As we will see later on, there are other cases in German where DiPs create expressive meaning in the clause middle position and that the corresponding Japanese sentences are suffixed by SFP. Thus, we might suggest that there is a generalization to the effect that what is happening with expressive meanings in the clause middle position by DiPs in German happens in the clause final position by SFPs in Japanese. To be more specific, Bayer and Obnauer (2011) propose a mechanism of creating non-standard questions in German by postulating the feature [QForce] for DiPs. Here, [QForce] is an unvalued uninterpretable feature and is valued by the interpretable iQForce through the operation Agree, as shown below. With the result, the interrogative force of the utterance is fine-tuned to create non-standard question of various types depending on the meaning of the type of DiPs.

\[
(7) \quad [\text{ForceP}/\text{FinP} \quad \text{Wh Force}/\text{Fin}\{iQ, \text{iQForce}[4]\} \cdots [\text{PrtP Prt}\{iQ\text{Force}[4]\}] \cdots]\]

Non-standard questions in Japanese can be created in the same way. Note that SFPs in Japanese are modal elements. Following Cinque’s (1999) hypothesis that various modal elements are hierarchically organized in the IP zone, SFPs originate and ordered according to Cinque’s hierarchy in the IP zone (Endo (2007, 2012)), where SFPs can be licensed in the same way as MPs in German to create non-standard question. One of the differences between German and Japanese is that SFPs end up in the clause final position in Japanese. Following Abraham’s (2012) idea that MPs move into the head position of ForceP to fix the illocutionary force, Endo (2007, 2012) assumes that SFPs also move into ForceP, and then the remnant IP is moved above it (See Endo (2007, 2012) for the motivation of the movement of SFPs, which satisfy the subject criterion on its way to ForceP.) Slightly departing from this idea, assume that the landing site of SFPs is the head position of Speech-actP (Hill (2014) partly because SFPs express speech-act and partly because the head of ForceP is already occupied by the Q-particle ka, as depicted below:

\[
(8) \quad \begin{array}{c}
\text{Speech-actP} \\
\text{Speech-act'} \\
\text{IP} \\
\text{Speech-act}^9 \\
\cdots \cdots \text{SFP...}
\end{array}
\]

Because Japanese is a strictly head-final language, the remnant IP moves into the Spec of Speech-actP, as shown below:

---

5 I am grateful to Luigi Rizzi (personal communication) for discussing this point.
In section 2, we will examine in more detail non-standard questions in Japanese, which are formed by head elements in the right periphery and are constrained by locality principle.

(B) Ne

The SFP ne is similar to the particle na that we saw immediately above in that it signals the expression of weak confirmation by the speaker about the proposition, but differs from the SFP na in that it is addressee-oriented, where a sentence suffixed by this particle is somewhat like what we see with English tag-questions. In addition, this particle has the feature [+empathy], which can be used to make the listener feel comfortable as if they share similar interest or information, as in in (10). Here, the clerk tells the price of a product to his/her customer with the sentence suffixed by the SFP ne to make his/her customer feel comfortable to try to show that (s)he shares similar interest with his/her customer.\(^6\)

(10) Customer: Kono oniku ikura desu ka?

this meat how.much is.polite Q

‘How much is this meat?’

Shopkeeper: Eeto, aore-wa 250yen desu ne.

well that-Top 250yen is.super-polite SFP

‘This is 10,000yen’ (Kamio 2002: 73)

The relationship between the SFP ne and the feature [+empathy] can be seen in developmental disorder as well. Watamaki (1997) reports that children with autism do not use the SFP ne at all, or even if they do, use it very infrequently. This is attributed to the fact that children with autism do not have the ability to share information with others, in contrast to normally developing children, who typically start using the SFP ne around the age between 18 and 24 months. We will discuss the relation between children with autism and other SFPs later on.

\(^{\text{6}}\) As a reviewer notes, the feature [+empathy] might be related to the notion of what Brown and Gilman (1960) call “solidarity and power”, where they argued that pronouns such as German du and French tu primarily convey solidarity, whereas pronouns like German Sie and French vous primarily convey power.
The difference between the SFPs na and ne can be clearly seen in the following conversation in (11a, b) between Charlie Brown and Peggy Jean that he fell in love with. Here, the sentences in (11a, b) are the original English sentences uttered by Peggy and Charlie Brown and the sentences in (11a', b') are the Japanese translation:

(11) a. That's cute, I like it.
b. Maybe I'll just jump into the lake right here
a'. Kawaii wa, li name SFP good name SFP
b'. Kono mizuumi-ni minage siyoo ka-na this lake-into jump. do Q-SFP

Here, Peggy Jean is looking at Charlie Brown while talking to him expecting some response from him with empathy, where the Japanese translation in (11a') is suffixed by the addressee-oriented SFP ne carrying the feature [+empathy]. In contrast, Charlie is shy and utters the sentence in (11b) without looking at Peggy like a monologue, where the Japanese translation in (11b') is suffixed by the speaker-oriented SFP na. Here, Charlie's sentence is also suffixed by the Q-particle which is immediately followed by the SFP na to create the expressive meaning of speaker's worry.

Incidentally, it is not the case that Charlie's sentences are always speaker-oriented. For instance, consider the following conversation between Charlie Brown and Lucy, where (12a) is the original sentence uttered by Charlie Brown and its Japanese translation in (12') is suffixed by the SFP ne in (12a').

(12) a. Except when the wind blows.
   a'. Kaze-ga fuka nakya ne wind-Nom blow except.if SFP

---

7 The element wa is the SFP expressing epistemic mood mainly used by a female speaker. See Endo (forthcoming) on this SFP.
Here, Charlie looks Lucy in the eye while talking to her. Unlike the previous sentences, Charlie’s sentence in (12a’) is suffixed by the addressee-oriented SFP ne expecting some response from Lucy.

When the SFP ne combines with a Q particle, it creates an expressive meaning of the speaker’s worry, where the use of the sequence ka-ne is typically restricted to a male speaker who is in a social status higher than the addressee like a teacher or pretends to be so. Thus, in the following sentence in (13a) by Linus, he tries to scold the addressee like a teacher, where the Japanese translation in (13a’) is suffixed by the ka-ne particle pair.

(13) a. Why can’t you dogs help people instead of being such a nuisance?
    a’. Kimita i-nu-wa ja-masu ru ka-warini hi-to-o ta-suke
    you dog-Top nuisance instead of people-Acc help
    ra-renai mo-n ka-ne? can’t Fin Q-SFP
(C) Yo

The SFP *yo* is somewhere in the middle between the particles *ne* and *na* in that it signals the expression of weak confirmation by the speaker about the proposition and is similar to the SFP *ne* in being addressee-oriented, but it differs from the SFP *ne* in that the particle does not carries the feature [+empathy]. This SFP indicates the speaker’s attitude that the stated proposition before the SFP is under the control of the speaker. Accordingly, it is typical to use the SFP *yo* to elicit a sense of insistence, as can be shown below, where Charlie Brown speaks to Snoopy and his utterance in Japanese is suffixed by the SFP *yo*.

(14) a. I’m home!
   Kaette.kita yo!
   home.came SFP

In this sense, the main function of the SFP *yo* is somewhat like what we see with the English vocative expression *hey*. The difference between *yo* and *ne* can be seen in the following two sentences uttered by Lucy.

(15) a. If anyone hits a ball to right field, let me know.
   a’. Dareka-ga right-ni utta-ra osiete ne
   anyone-Nom right-to hit-if tell.me SFP

---

As a reviewer correctly points out, “weak confirmation by the speaker about the proposition” entails an invitation to the addressee to supplement this admittedly weak confirmation. In other words, where the speaker is not fully confident and is admitting this incompleteness, there is an invitation to the addressee to respond by providing confirmation or disconfirmation. It is this implicit invitation that activates the feature [+addressee].
a. You'd better watch what you write in that autobiography!

a'. Jizyoden-ni kaku koton-i wa kiotuketa hoogaii wa yo.

autobiography-to write what you write watch had.better SFP SFP

In (15), Lucy is asking a favor of Charlie Brown, where she seems to have empathy to Charlie because when the speaker asks a favor of someone, he/she tries to share information with the addressee. Correlatively, in the Japanese translation in (15a’), the sentence is suffixed by the SFP ne with the feature specification [+empathy]. In contrast, in (16) Lucy gives Snoopy a warning with anger, where she does not seems to have empathy to Snoopy because when the speaker give a warning, he/she is more interested in expressing his/her emotion than thinking how the addressee feels by his/her warning. Correlatively, the corresponding Japanese sentence is suffixed by the SFP yo without the feature specification [+empathy].

Recall here Watamaki’s (1997) report that children with autism do not use the SFP ne at all, or even if they do, use it very infrequently. Watamaki also reports that children with autism do use the SFP yo like normally developing children. He attributes this to the fact that children with autism do not have to share information with others in the use of the SFP yo. (Incidentally, there are many papers dealing with the connection between autism and SFPs. See Satake and Kobayashi (1987) for the communication pattern between mother and children with autism in the
use of SFPs, Arai and Nakamura (2016) for verbal and visual training of children with autism using various SFPs, Yamamoto and Asano (2012) for evidence-based training of children with autism on the use of SFPs, Takiyoshi and Tanaka (2011) for general issues of autism and SFPs, among many others.)

Turning to the expressive meanings created by SFPs, when the SFP yo combines with a Q particle ka, it forms a rhetorical question where the speaker does not expect an answer from the addressee (Saito 2015) and seems to express irritation or worry. The situation is similar to the following German rhetorical question formed by the DiP schon in (17a), where the corresponding Japanese translation in (17a’) is suffixed by the Q-particle ka and the SFP yo. This rhetorical meaning, which does not sound like expecting response from the addressee, seems to arise from the fact that the SFP yo does not have the feature [+empathy] because when the speaker does not feel empathy to the addressee, he/she would not expect response from the addressee.  

(17) a. Wer zahlt schon gerne Steuern? (Bayer and Obenauer’s (2011) 46)  
   who pays SCHON gladly taxes  
   ‘Who gladly pays taxes’

a’. Dare-ga suki.konon.de zeikin nanke harau ka yo  
   who-Nom gladly tax epithet pay Q SFP

4 Mechanism of creating expressive meanings

Now that we have seen several expressive meanings in Japanese being created by combining the Q particle ka with a SFP, let us next examine the syntactic mechanism to create these expressive meanings out of a Q particle and a SFP in Japanese. My proposal is to use Rizzi’s (1997) idea that in the absence of an element between Force and Fin, Force and Fin are amalgamated as a single unit as shown in (18a). Because Japanese is a head-final language the linear order of these functional elements is the mirror image of (18b) below. Thus, in the absence of an element between Force and Fin, we find the amalgamated form of the Force element no and the Fin element ka in (18c), where either the Fin element no or the Force element ka is optionally pronounced:

(18) a. [ForceP]+[FinP]…  
b. [ForceP]+[FinP] (Japanese)  
c. …[FinP no]+[ForceP ka]

One of the properties of the SFP yo can be found in the fact it may trigger the deletion of the nominative Case particle ga more frequently than other SFPs like na and ne do. This particle deletion is optional. When deletion does not take place, the DP suffixed by the particle ga is interpreted as focus. This is related to the historical fact in Japanese that the nominative particle ga was originally a focus marker in Old Japanese and the meaning of focus got fossilized in the course of history. In some restricted environments, the focus interpretation on the nominative Case ga survives and thus the deletion of this particle can be seen as a strategy to eliminate the focus interpretation of the subject DP. See Endo (forthcoming) on this point. I am grateful to Marcel den Dikken (personal communication) for discussing this point.
With this background in mind, let us consider the expressive meaning in (5) we saw in the previous section, where the Japanese translation sentence uttered by Charlie Brown in (5a') is suffixed by the SFP na:

(5)  

(a) I wonder if it's possible to be in love with two different girls at the same time

(a') Doozi-ni futari-no ko-o aisurunante kanoo ka-na
at.the.same.time two-Gen girls-Acc love possible Q-SFP

I propose that some SFPs, including the SFP na, are inherently specified as [+expressive], and select the amalgamated Force-Fin pair we proposed immediately above:

(19) …[FinP no]+[ForceP ka] [SFP na/ne]
     ⇐-select-----|

In (5), Charlie’s sentence is suffixed by the SFP na, which carries the expressive meaning feature [+expressive] and locally selects the Fin-Force pair no-ka, where the Fin element no remains unpronounced.

The same explanation holds for the following sentence in (7b) uttered by Charlie Brown, which is repeated below. Here, the SFP na selects the Force-Fin pair element to create the expressive meaning of speaker’s worry, where the pronunciation of the Fin element no is not made.

---

10 At this point, one may wonder under what condition the Fin element no remains unpronounced. Makihara (1995) notes that no obligatorily appears when there is a gap between the speaker’s assumption and the real situation in wh-question asking for reason. Because a subtle meaning difference is felt with and without no, his generalization might carry over to other sentence type as well. More study is required in this area.
Exploring right/left peripheries: Expressive meanings in questions

(7)  b. Maybe I’ll just jump into the lake right here.
    b’. Kono mizuumi-ni minage siyoo ka-na.
        this lake-into jump do Q-SFP
        this lake-into jump. Do Q-SFP

Thus, my proposal is that a locality principle is operative in calculating expressive meanings created out of the Q-particle *ka* and a SFP. Here a note of caution is in order about the notion of locality. Rizzi (2017) emphasizes that selection is strictly local, where no other head may intervene between a selector and a selectee, while Agree-type relation is local but not strictly local, where only those elements that belong to the same type of the Agreed element counts as an intervening element. The strict local nature of selection that we are concerned with here is depicted in (20) below, where SFP[*E*] is a SFP carrying the feature [+expressive] that selects the Force-Fin pair. In (20a), in the presence of the intervening politeness marker *desu* between the Force+Fin pair, the SFP *ne* cannot select the Force+Fin pair because the Fin element *no* is too far away from the SFP *ne*, and then the expressive meaning of speaker’s worry disappears. In (20b), in the presence of the intervening quantifier *mo* between Force and the SFP *ne*, the SFP cannot select the Force+Fin pair, because the quantifier *mo* counts as an intervener and thus the expressive meaning of speaker’s worry disappears.

(20)  a. …[FinP no] …desu….[ForceP ka] [SFP[*E*] *ne*
      polite  ←---X-----select---------[---]  
      b. …[FinP no] +[ForceP ka] *mo* [SFP[*E*] *ne*
        also                  ←---X-----select----[---]

In other words, following the suggestion by Ur Shlonsky (personal communication), I postulate two types of the SFP *ne*, one of which carries the expressive feature [+E] locally selecting the Force+Fin pair and the other of which does not carry the expressive feature [+E] without being required to be local to the Force-Fin pair. The main function of the non-expressive SFP *ne* is simply to make confirmation to the addressee and does not require a local relation with another element. Thus, this non-expressive particle *ne* may be suffixed to an element that is even found in the non-sentence-final position, as we see in (21) below, where the SFP *ne* is suffixed to the topic phrase headed by the particle *wa*. 
(21)  a. Sometimes, every now and then...once in a while.
    b. Itumozya-nai kedo tokini-wa-ne itido-ka-ni-do
        always-not though sometimes-Top-SFP once-or twice

At this point, one may wonder if there are any SFPs that are inherently and exclusively [+expressive] and may only appear in a position strictly local to the Force-Fin pair. SFPs like *sira and *i are possible candidates. What is interesting about these particles is the fact that they gender-oriented, where the SFP *sira, is used by a female speaker while the SFP *i is used by a male speaker. As illustrated in (22a') and (22b'), the SFPs *sira and *i must have a local relation with the Q-particle *ka and thus may not have an intervener like *mo 'also'.

(22)  a. Ame-ga huru no ka *sira.
      rain-Nom fall Fin Force SFP(female)
      'I wonder if it would rain'
    b. Ame-ga huru no ka *i.
      rain-Nom fall Fin Force SFP (male)
      'I wonder if it would rain'
    a' *Ame-ga huru no ka mo *sira.
       rain-Nom fall Fin Force also SFP (female)
      'I wonder if it would rain'
    b. *Ame-ga huru no ka mo *i.
       rain-Nom fall Fin Force also SFP (male)
      'I wonder if it would rain'

The gender restrictions of the SFPs *sira and *i can be confirmed by the following sentences from *Peanuts*, where the sentence in (23a) is uttered by the female speaker Marcie and the corresponding Japanese sentence in (23a') is suffixed by the female SFP [+expressive] *sira while the sentence in (23b) is uttered by the male speaker Charlie and the corresponding Japanese sentence in (23b') is suffixed by the male SFP [+expressive] *i:

---

11 When the SFP *i combines with the affirmative element *da to create the sequence *da-i, it expresses a strong insistence.
Exploring right/left peripheries: Expressive meanings in questions

To summarize, we have seen the cartography of interrogative sentences, which are expressed in the right periphery by various types of SFPs in combination with the Q-particle ka in Japanese, as shown below. We have also seen that the expressive meanings in Japanese are constrained by locality principle.

(24) Force …Polite…Fin…Quant…SFP[+E] …SFP

5 Implication 1: Interjection

Let us next examine some implications of our approach to expressive meanings in non-standard questions in Japanese by looking at the left periphery. Recall here from the previous section that there are gender-oriented SFPs like sira, which may only be used by a female speaker, and i, which may only be used by a male speaker. This gender orientation can be confirmed by the fact that only the former, as opposed to the latter, is compatible with the female interjection ara yada ‘oh, no’, as illustrated below:

(25) a. Ara, yada ame-ga huru no ka sira.
   ‘I wonder if it would rain’
   oh, no (female) rain-Nom fall Fin Force SFP(female)

b. *Ara, yada ame-ga huru no ka i.
   ‘I wonder if it would rain’
   oh, no (female) rain-Nom fall Fin Force SFP (male)
Non-expressive SFPs, on the other, do not have gender orientation. Thus, the non-expressive SFP \(ne\), which has no local relation with the Force-Fin pair, may be used by a male speaker as in (26a) or a female speaker as in (26b).

(26) a. b. 

Recall also from the previous section that when the Force-Fin pair is selected by the expressive SFP \(ne\), it expresses worry by a male speaker. This point can be confirmed by the same co-occurrence restriction of female interjections such as \(ara\ yada\ 'oh\ no' or \(kyaa\ 'aaugh'\).

(27) *\(ara\ yada/kyaa\ ame-ga\ hut-teiru\ no\ ka-ne.\)  
\(oh\ no/aaugh\ [+female]\ rain-Nom\ fall\ Fin\ Q-SFP[+E] (=[male])\)  
‘Oh/Augh, I wonder if it is raining’

cf. \(ara\ yada/kyaa\ ame-ga\ hut-teiru\ no\ ka-sira.\)  
\(oh, no/aaugh\ rain-Nom\ fall-Asp\ Fin\ Q-SFP[+female]\)  
‘Oh, no/Augh, I wonder if it is raining’

The sentence in (27) sounds very strange due to the incompatibility between the feature [+male] carried by the expressive SFP \(ne\) and the feature [+female] carried by the female interjection

\(^{12}\) The element \(sira\) is an inherently expressive SFP that is typically uttered by a female speaker and thus is compatible with the female interjections \(maa/ara\ and \(kyaa\).\)
ara yada ‘oh, no’ and kyaa ‘augh’. How can we formally express the compatibility between gender-oriented particles and the female oriented vocative expression ara yada ‘oh, no’ and kyaa ‘augh’? I propose to capture the compatibility by the mechanism of what Endo and Haegeman (2019) call ‘adverbial concord’, where an adverbial element imposes some restriction on the associated functional element in the main clause. Although the mechanism ‘concord’ is originally used for adverbial clauses, I propose to extend this mechanism to interjections. Thus, as shown in (28), the interjections ara yada ‘oh, no’ and kyaa ‘augh’ carrying the feature [+female] impose restriction on the associated SFP to have the feature specification [-female]. Because this requirement is not met in (28), the sentence sounds strange.

(28) [ModP Ara yada/kyaa] ame-ga hutteiru no ka-ne. oh, no /augh [+female] rain-Nom fall Fin Q-SFP[+]E[=+[male]]
/X------------------adverbial concord merge----------------/

As expected, when the SFP ne is preceded by an intervener as in (29) below, the sentence is ameliorated, because a non-local SFP has no gender-orientation.

(29) Ara yada /kyaa ame-ga hutteiru no ka-mo-ne. oh, no /Augh [+female] rain-Nom fall Fin Force-also-SFP

‘Oh/Augh, I wonder if it might be raining’

Thus, in the presence of the intervening element mo between the Force element ka and the expressive SFP ne, the SFP[+expressive] ne cannot select the Force-Fin pair consisting of no and ka, where Force is too far away from the selecting SFP ne. Thus, as depicted in (30) below, the SFP ne cannot be an expressive SFP[+E] carrying the feature [+male] but is a simple SFP with no gender restriction. Because the non-expressive SFP ne imposes no gender restriction on the associated SFP, the sentence suffixed by this non-expressive SFP ne is compatible with the female interjections ara yada ‘oh, no’ or kyaa ‘augh’.

(30) …[Fin (no)] [PoliteP desu] [ForceP ka]-*[SFP ne[+E] ne [+male]]/ ↯ [SFP[-E] ne ←-------not local------------------/

The same amelioration effect is attested in case where the politeness particle desu intervenes between Force and Fin:

(31) Ara yada /kyaa ame-ga hutteiru no desu ka-ne
    oh, no /augh [+female] rain-Nom fall Fin polite Q-SFP

‘Oh/Augh, I wonder if it might be raining’

Here, in the presence of the intervener desu between the Force element ka and the Fin element no, Force and Fin cannot form an amalgamated Force-Fin pair, and thus the SFP ne is too far
away from the Fin element no to select it. For this reason, the SFP ne cannot be the expressive SFP [+E], but is a normal SFP without any gender specification, and thus the sentence in (31) can be prefixed by the gender interjection ara yada ‘oh, no’ or kyaa ‘augh’.

One might think that the operation “concord” simply expresses certain semantic compatibility conditions between two elements. However, this operation is also syntactic in nature, because we see the point of attachment of an adverbal clause or a vocative element is determined by concord with a functional element in the matrix. The syntactic nature of “concord” relation can be seen in the case where two adverbial clauses appear at the same time. If, as suggested above, the matrix functional head signals the timing of the merger of the adverbal clause, this would lead to the prediction that the higher an adverbal clause is situated in the clause, the later it will be merged with the associated clause, where a higher adverbial clauses always precedes a lower adverbial clause. See Endo and Haegeman (2019), where this prediction is shown to be borne out with various types of adverbial clauses. The same prediction can be made with the case at hand, i.e. the concord relation between a vocative element like ara yada ‘oh, no’ and the associated expressive particles ka-sira. Because the vocative element ara yada ‘oh, no’ is expressed at the CP level or above, it always precedes an element in the IP domain, for instance, a temporal expression such as gogatu-ni ‘in May’, as shown below in (32):

(32)  
   oh, no May-in snow-Nom fall Asp Fin Q-SFP
   ‘Oh, no, I wonder if it is snowing in May’

b. ??Gogatu-ni ara, yada, yuki-ga fut teiru no ka-sira.
   May-in oh, no snow-Nom fall Asp Fin Q-SFP
   ‘Oh, no, I wonder if it is snowing in May’

To summarize, we have seen that gender orientation seen with expressive SFPs can be confirmed by the mechanism of concord.

6 Implication 2: how come questions

In the previous section, we have seen that expressive meanings are created by combing the Q particle ka and SFPs such as ne, yo, etc. There are other SFPs that seem to be relevant to expressive meanings, especially in questions asking for reason. To see the point, recall the fact that how come questions may express (i) surprise by the speaker that a particular state of affairs should hold or (ii) a mild curiosity by the speaker. Where do these expressive meanings of how come questions come from? To answer this question, it would be helpful to consider the following sentences that are borrowed from Schultz (2008):
(33) a. L: How come you never send me flowers? (Schultz 2008)
S: Because I don’t love you.

b. L: Do-site watasi-ni itido.mo hana-o okutte kurenai no?
S: Kimi-ga kirai dakara 13

You-Nom do.not.like because

In (33), Lucy uses a how come question to Linus and receives a response with the sentence prefixed by because. Here, the corresponding Japanese how come question sounds like a standard question suffixed by the SFP no, where a mild curiosity by Lucy is felt. In contrast, in (34), although Linus uses a how come question to Lucy, he does not receive response prefixed by because but only a comment from her. The corresponding Japanese how come question

13 The direct object watasi ‘me’ is suffixed by the nominative Case particle ga. This is because stative predicates in Japanese requires the nominative Case particle ga for the direct object.
sounds like a non-standard question or a rhetorical question with strong irritation by the speaker where the sentence is suffixed by the SFP sa. Note that Linus’s face and gesture in (34) show his strong emotion than Lucy’s face we saw in (33). Based on the fact the difference between the meaning of curiosity and surprise in how come questions is marked by different SFPs no and no-sa in Japanese, I suggest that the two meanings of how come questions in English arise by activating different covert functional heads that corresponds to the SFPs no and no-sa. Here, the minimal difference between the two cases is the presence and absence of the SFP sa. What is the SFP sa? The SFP sa is used to report familiarity by the speaker about the proposition, implying that the sentence suffixed by this SFP should be taken as a matter of course. According to Uyeno (1971), the meaning of the particle sa is contrasted with the meaning of the particles yoo ‘appear,’ rasii ‘seem,’ and soo ‘hear,’ which are used when the speaker’s judgment is made based on appearance. The SFP sa, in contrast, is used when the speaker’s judgment is already made based on his own supposition or inner feeling. Because the speaker’s judgment is already made in uttering the SFP sa, the speaker’s supposition is taken to be discourse-familiar, and thus, we cannot start a discourse with a sentence with the SFP sa, as illustrated by the following contrast. (See also Hasunuma (2015) on this point.)

(35) a Kore nani? / ??Kore nani sa?
   this what this what SFP
   ‘What is this?’

b. * Doo suru sa?
   how do SFP
   ‘How are you going to do?’

Based on these facts, I suggest that the SFP sa is related to old information or discourse-familiarity. Based on the work by Fitzpatric (2005), Radford (2018) emphasizes that how come questions are factive in nature and proposes that the completer that following how come in English is FactP. I suggest that the SFP sa might be an overt realization of Radford’s FactP. When this functional head is activated in English, it might give rise to the expressive meaning of surprise, while it is not activated, mild curiosity interpretation might arise. More research is required in this area.

7 Conclusion

To summarize the whole story, I have discussed the nature of SFPs in Japanese with special attention to non-standard questions by making some reference to the corresponding German cases. One of the differences between Japanese and German is that non-standard question in German is expressed by placing DiPs in the middle of a clause while the non-standard question in Japanese is expressed by SFPs in the clause final position, where formation of non-standard questions is constrained by a theory of locality. I have also suggested that the presence and the absence of the expressive meaning in English how come questions can be attributed to the activation and deactivation of the functional category that is response for the expressive meaning that is silent in English and is overt in Japanese.
References


Yoshio Endo


Adverb-predicate agreement in Japanese and structural reduction

Norio Nasu
nnasu@inst.kobe-cufs.ac.jp

1 Introduction*

Two principal ideas underlie the cartographic approach to the study of clause structure. One is that there is a universal template that determines the arrangement of functional projections. Each functional projection is headed by a category that has a different interpretive import such as illocutionary force, finiteness, topic, focus, and so on. They are assumed to be arranged in a fixed order. The second is that each functional projection may contain at most one specifier. The specifier and the functional head enter into a one-to-one relation within the same projection by sharing a feature relevant to scope-discourse properties. When coupled together, these assumptions lead to a generalization, whereby the distribution of phrasal constituents on the left periphery of a clause reflects a universal and fixed hierarchical ordering of functional heads.

This paper tests the generalization from a critical point of view. In particular, it aims at identifying factors that constrain the distribution of constituents on the left periphery of a clause. While retaining the idea that an element in the specifier is licensed by a functional head, we argue that this relation need not be formed within a single projection. The licensing can be done even if the elements involved are located in separate projections.

We provide support for this view by examining the distribution of sentential adverbs (S-adverbs) in Japanese. As illustrated below, they behave differently with respect to their occurrence in a subordinate clause.

(1) a. *[Mosi kitto sigoto-ga hayaku aware-ba] paatii-ni maniau.
   Lit. 'If we certainly finish the job early, we will be in time for the party.'
   if certainly job-Nom early is.over-if party-Dat in.time

b. *[Mosi dooyara sigoto-ga hayaku aware-ba] paatii-ni maniau.
   Lit. 'If we apparently finish the job early, we will be in time for the party.'
   if apparently job-Nom early is.over-if party-Dat in.time

c. *[Mosi un'yoku sigoto-ga hayaku aware-ba] paatii-ni maniau.
   Lit. 'If we luckily finish the job early, we will be in time for the party.'
   if luckily job-Nom early is.over-if party-Dat in.time

* An earlier version of this paper was presented at the HFL 2018 conference (Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages) held in Konstanz, Germany May 14-17, 2018. I would like to thank the audience for their questions and comments. In particular, I thank Yoshio Endo, Sergio Monforte, and Andrew Simpson. Thanks are also due to Mark Campana for suggesting stylistic improvements. Needless to say, all remaining errors and inadequacies are mine. This research was supported by the Grant-in-Aid for Scientific Research (C) (#18K00578) from the Japan Society for the Promotion of Science.
While epistemic and evidential adverbs (*kitto* ‘certainly’ and *dooyara* ‘apparently’) cannot occur in a conditional clause, an evaluative adverb such as *un’yoku* ‘luckily’ can. An initial question to be addressed here is what this difference is rooted in.

The analysis proposed in this paper consists of two assumptions. First, the distribution of Japanese *S*-adverbs is dependent on the type of predicate and its inflectional morphology: Each adverb is required to occur with a particular inflectional form of a predicate. This phenomenon is dubbed Adverb-Predicate Agreement (APA). Second, each inflectional form is associated with a different functional head in the clausal spine. This entails that the inflectional form of a predicate indicates the position where the predicate is realized. These assumptions lead to the following generalization: Each *S*-adverb is linked with a particular functional head. This in itself is a garden-variety phenomenon in the cartography theory known as “Spec-Head Agreement” or “Criterion” (Rizzi 1996, 1997, 2004, 2006, 2011, 2015), whereby a constituent in the specifier position of a projection is licensed by a functional head. Our analysis departs from the standard assumption, however, in that it does not restrict the licensing relation to a single projection. Instead, it argues that APA is a reflection of a potentially long-distance probe-goal (or Agree) relation.

The paper is organized as follows. Section 2 provides a brief overview of the inflectional system in Japanese in order to facilitate the readers’ understanding of data and discussion in subsequent sections. It then lays out the primary facts concerning APA in Japanese. Section 3 advances a novel analysis of APA. The gist is that an inflectional form can be associated with different functional heads, each one of which carries interpretable formal features that serve to characterize its scope-discourse functions. *S*-adverbs, on the other hand, have uninterpretable features. They enter into an Agree relation with an appropriate functional head. Subsequent sections discuss consequences and implications of the proposed analysis. Section 4 compares our analysis with Haegeman’s (2010, 2012). She argues that the distribution of phrasal categories on the left periphery of an adverbial clause is subject to a locality condition on operator movement. However, her analysis cannot be extended to Japanese, because there are ample pieces of evidence that argues against involvement of an operator in adverbial clauses. An alternative account is shown whereby the head plays a significant role in languages like Japanese. Section 5 examines the interaction between *S*-adverbs inside a single clause. There it is argued that the distribution of *S*-adverbs in Japanese cannot be captured by a template-based approach to clause structure. Section 6 concludes the paper.

## 2 An overview of core data

### 2.1 Inflection in Japanese

Predicates in Japanese are realized in various inflectional forms, depending on the syntactic environments in which they occur. Among the various inflectional forms, we will focus on the following four: conclusive, prenominal, infinitival, and preverbal forms. They will be of direct relevance to the discussion in subsequent sections.

The conclusive form canonically appears in either the root or the complement clause of a verb of saying or thinking. In the verbal paradigm, it is formed by attaching either the past tense morpheme *-ta* or the present tense morpheme *-(r)u* to the stem.\(^1\)

---

\(^1\) The following abbreviations are used in this paper: *Acc* = accusative, *Asp* = aspect, *Cl* = classifier, *Concl* = conclusive, *Cop* = copular, *Dat* = dative, *Fin* = finiteness, *Foc* = focus, *Gen* = genitive, *Imp* = imperative,
(2) a. *John-wa ohiru-ni soba-o *tabe-{ru / ta}.
    John-Top lunch-for noodle-Acc eat-{ Pres.Concl / Past.Concl }
‘John { eats / ate } noodles for lunch.’

b. *Mary-wa [John-ga ohiru-ni soba-o *tabe-{ru / ta}
    Mary-Top John-Nom lunch-for noodle-Acc eat-{ Pres.Concl / Past.Concl }
to ] oomotteiru.
C think
‘Mary thinks that John { eats / ate } noodles for lunch.’

The prenominal form most typically shows up when the predicate immediately precedes a
nominal category. For example, the verb inside a relative clause is realized in this form.

(3) [Ohiru-ni soba-o *tabe-{ru / ta}] 
    lunch-for noodle-Acc eat-{ Pres.Prenom / Past.Prenom } person
‘the person who { eats / ate } noodles for lunch.’

One may notice that the prenominal forms *tabe-ru* ‘eat’ and *tabe-ta* ‘ate’ in (3) are
morphologically identical with the conclusive forms in (2a-b). Such morphological identity is
often found in verbal and adjectival paradigms; because of this, it is virtually impossible to tell
these inflectional classes apart. Nevertheless, there is a clear distinction between them in the
present tense copula paradigm. Consider the contrast below.

(4) a. *Sono gizyutu-wa iryoo-ni ooyookanoo-{da / *na}.
    the technology-Top medicine-to applicable-{ is.Concl / is.Prenom }
‘The technology is applicable to medicine.’

b. [iryoo-ni ooyookanoo-{ *da / na } ] gizyutu
    medicine-to applicable-{ is.Concl / is.Prenom } technology
‘the technology which is applicable to medicine’

*Da* is the present tense conclusive form of the copula, whereas its prenominal counterpart is *na*. The root clause in (4a) only allows the former. Conversely, the relative clause in (4b) is only
compatible with the latter.\(^2\)

The term ‘infinitival’ is used for an inflectional pattern that does not exhibit a past-present
alternation.\(^3\) In certain types of subordinate clause, the predicate is only realized either in the
past or the present tense form, depending on clause types. In (5a), for instance, the temporal

\(^2\) This distinction disappears in the past tense paradigm, where both conclusive and prenominal forms are
realized in the same form *datta*.

\(^3\) Although opinions vary among scholars as to whether the infinitival form constitutes an independent
inflectional class, we assume that it does, following Mihara (2011, 2012, 2015) (see also Mikami 1959).
clause headed by the conjunction to 'when' allows only a present tense predicate. On the other hand, the clause headed by *mama* 'with' in (5b) is compatible only with a past tense predicate.\(^4\)

   John-Top oyster-Junct eat-{ Pres / Past } -when stomach-Junct break
   'John has stomach trouble when he eats oysters.'

   John-Top on.his.back lie-{ Pres / Past } -with move-not
   'Lying on his back, John doesn’t move.'

The preverbal form appears when the predicate immediately precedes another (see (6a)). It also appears with certain types of subordinating conjunct (see (6b)). The most conspicuous trait of this form resides in its lack of the tense affix.

(6) a. Karera-wa ohiru-o [ t0 / 'ru' / 'ta'] -hazimeta.
   they-Top lunch-Junct eat-{ Prev / Pres / Past } -began
   'They began eating lunch.'

   b. Karera-wa [ ohiru-o *t0 / 'ru' / 'ta'] -nagara hanasiasi-o sita.
   they-Top lunch-Junct eat-{ Prev / Pres / Past } -while discussion-Junct did
   'They had a discussion while eating lunch.'

2.2 Adverb-predicate agreement

Japanese S-adverbs are required to co-occur with particular types and forms of predicates (Yamada 1936, Hashimoto 1959, Watanabe 1971, Sawada 1978, Kudo 2000, Sugimura 2009, Morimoto 2011, Larm 2012). This phenomenon is called Adverb-Predicate Agreement (APA) in this paper. Here we offer a bird-eye overview of how it works, with special attention to the distribution of evidential, epistemic, and evaluative adverbs. Two factors are taken into account: type of predicate and type of inflectional form.

Note first that each class of S-adverb occurs with a different type of predicate.

(7) Evidential

   Dooyara ame-ga yanda { yooda / rasii / *hazuda / *nitigainai / *0 }.
   apparently rain-Nom stopped seem seem should must
   'Apparently it stopped raining.'

---

\(^4\) This restriction applies regardless of the tense of the matrix predicate. Compare (5a-b) with (ia-b), where the matrix predicate is in the past tense:

(i) a. John-wa [ kaki-o *t0] onaka-o kowasita.
   John-Top oyster-Junct eat-{ Pres / Past } -when stomach-Junct broke
   'John had stomach trouble when he ate oysters.'

   John-Top on.his.back lie-{ Pres / Past } -with move-not-Past
   'Lying on his back, John didn’t move.'

Notice that the tense shift does not affect the form of the subordinate predicate.
(8) Epistemic

*Kitto* ame-ga yamu { *yooda / *rasii / hazuda / nitigainai / *Ø }.
certainly rain-Nom stop seem seem should must
‘Certainly it will stop raining.’

(9) Evaluative

*Saiwainimo* John-wa sono basu-ni notta { yooda / rasii / hazuda / nitigainai / Ø }.
luckily John-Top the bus-Dat took seem seem should must
Lit. ‘(It seems that / I’m sure that) luckily John took the bus.’

An evidential adverb must occur with an evidential modal. It cannot occur with a modal of different semantic class (7). Likewise, an epistemic adverb requires an epistemic modal (8). Neither class of adverb can be used without a modal. By contrast, an evaluative adverb places no such restriction on the type of predicate it may occur with. As illustrated in (9), it is compatible with both evidential and epistemic modals; it can even show up without a modal.

Second, as shown in the tables below, an S-adverb must enter into an agreement relation with a particular inflectional form (or forms). Thus, depending on the type of adverb, the predicate can be realized in different inflectional forms.

(10) Inflectional forms that S-adverbs agree with

<table>
<thead>
<tr>
<th>modal predicate</th>
<th>preverbal</th>
<th>infinitival</th>
<th>prenominal</th>
<th>conclusive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidential adverb</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Epistemic adverb</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>verbal /adjectival predicate</th>
<th>preverbal</th>
<th>infinitival</th>
<th>prenominal</th>
<th>conclusive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluative adverb</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

Evidential adverbs are subject to the tightest restriction. They are compatible only with the conclusive form of an evidential modal, i.e., they cannot occur with a modal realized in another form. This is illustrated below.

apparently rain-Nom fall seem-Pres.Concl  
Lit. ‘Apparently it seems to rain.’

b. [ (*dooyara) ame-ga huru yoo-na ] tenki-wa kiraida.  
apparently rain-Nom fall seem-Prem weather-Top hate  
Lit. ‘I hate the weather such that apparently it seems to rain.’

apparently rain-Nom fall seem-Inf -when everyone-Nom in.trouble  
Lit. ‘Everyone is in trouble when apparently it seems to rain.’

John-Top apparently kind seem-Prev -though in.reality unkind-is  
Lit. ‘John is unkind in reality, though apparently seeming to be kind.’
Epistemic adverbs, on the other hand, can occur with either conclusive or prenominal forms of an epistemic modal. Still, they are incompatible with infinitival or preverbal forms.

(12) a. **Nimotu-wa kitto todoku hazu-da.**
    parcel-Top certainly arrive should-Pres.Concl
    ‘The parcel certainly should arrive.’

b. [ **Kitto todoku hazu-no** ] nimotu-ga mada todoite inai.
    certainly arrive should-Pres.Prenom parcel yet arrive hasn’t
    ‘The parcel which certainly should have arrived has not arrived yet.’

    certainly parcel-Nom on.time arrive should-Inf -when glad
    Lit. ‘I’ll be glad if certainly the parcel should arrive on time.’

    certainly TV-Acc watch should-Preset is -while John-Top studying-was
    Lit. ‘John was studying, while certainly being bound to watch the TV.’

Evaluative adverbs exhibit the most flexible agreement pattern of all. They can occur with any inflectional form except the preverbal (13a-d).

(13) a. **Saiwainimo John-wa ie-ni i-{ru / ta}.**
    luckily John-Top home-at be-Inf {Pres.Concl / Past.Concl}
    ‘Luckily John { is / was } at home.’

b. [ **Saiwainimo sii-ni kat-ta** ] sensyutati-wa yorokonde-ita.
    luckily game-Dat win-Past.Prenom players-Top delighted-were
    ‘The players who luckily won the game were delighted.’

c. [ **Saiwainimo John-ga subayaku kagi-o mituke-ru -to** ]
    luckily John-Nom quickly key-Acc find-Inf -when
    mina hottosita
    everyone was.relieved
    Lit. ‘Everyone was relieved when luckily John found the key quickly.’

d. [ *(Saiwainimo) kasa-o mot-te** ] dekake-nasai.
    luckily umbrella-Acc carry-Prev go.out-Imp

Lit. ‘Go out, luckily carrying an umbrella!’

3 Analysis

3.1 Outline

As discussed above, an S-adverb enters into an agreement relation with a particular inflectional form. Based on this observation, this section shows how the association is implemented. The proposed analysis consists of two parts. First, following Mihara (2011, 2012, 2015), we assume that there is a close correlation between inflectional forms and clause structure. A predicate undergoes head movement and is realized in various inflectional forms depending on its final landing site. This is schematically illustrated below.

(14) preverbal infinitival prenominal conclusive
    \[
    V \quad \text{\textit{VP}} \quad V \quad \text{\textit{IP}} \quad \text{T} \quad \text{\textit{TP}} \quad \text{\textit{Fin}} \quad \text{\textit{FinP}} \quad \text{\textit{Force}} \quad \text{\textit{ForceP}}
    \]
When a verb moves to $v$ and stops there, it is realized in the preverbal form. When it moves up to $T$, it takes the infinitival form. Likewise, $\text{Fin}$ is the final landing site for a predicate in the prenominal form, and $\text{Force}$ is the position where a predicate takes the conclusive form.

Second, recall that an $S$-adverb can only occur with a particular inflectional form of a predicate. Coupled with the idea that each inflectional form is associated with a functional head, it follows that $\text{APA}$ represents agreement between an $S$-adverb and a functional head. We propose that this agreement relation is a manifestation of the probe-goal relation, as summarized below.

(15) \begin{array}{|l|l|}
\hline
\text{probe (adverb)} & \text{goal (functional head)} \\
\hline
\text{evidential:} & \text{Force:} \\
[u\text{Force}] & [i\text{Force}] \\
[u\text{Modal}] & [i\text{Modal}] \\
\text{epistemic:} & \text{Fin:} \\
[u\text{Fin}] & [i\text{Fin}] \\
[u\text{Modal}] & [i\text{Modal}] \\
\text{epistemic:} & \text{T:} \\
[u\text{T}] & [i\text{T}] \\
\hline
\end{array}

An $S$-adverb carries an uninterpretable formal feature (or features) and enters into an Agree relation with a functional head that carries interpretable features of the same type. Two kinds of features are postulated. For ease of exposition, let us call them “modal feature” (represented as [Modal]) and “inflectional feature”. The former specifies the modal with which an $S$-adverb is associated. The latter is a cover term for features specifying individual functional heads. It subsumes features like [Force], [Fin], and [T]. As summarized in (15), $S$-adverbs carry uninterpretable tokens of these features, whereas modals and functional heads carry their interpretable counterparts.

Agreement is an asymmetric relation in that one element affects the other. In a subject-verb agreement language like English, the verb’s inflectional form is determined in accordance with the $\Phi$-properties of the subject (person, number, and gender). In this case, the subject affects the verb. The $\text{adverb}$-$\text{predicate}$ relation in Japanese is also asymmetric. An $S$-adverb is sensitive to the type and inflectional form of the predicate it occurs with (see (7)-(9) and (11)-(13)). The former is the affected element and the latter is the affecter. In the Agree-based model (Chomsky 2000, 2001, et seq.), the affected element carries an uninterpretable feature and acts as the probe, whereas the affecting element acts as the goal. An $S$-adverb (the “affectee”) thus carries the uninterpretable feature(s) in (15).

\text{Note that the probe is located in the non-head position in the scheme summarized in (15). The adverb is located on the left edge, presumably the specifier position, of a functional projection. This position is usually occupied by a phrasal category. The proposed analysis is therefore incompatible with Chomsky’s (2004, 2007) that only heads can be a probe or a goal. However, as shown by Bošković (2007), if a probe is defined as an item that carries an uninterpretable feature, and if a probe must find its goal in its c-command domain, any item with an uninterpretable feature qualifies as a probe, i.e. irrespective of its structural position. See also Kobayashi (2014) for a similar view.}
3.2 The agree-based account of APA

An evidential adverb carries two uninterpretable features. This reflects the fact that it must occur with an evidential modal realized in the conclusive form.

(16) a. Dooyara ame-ga huri soo-da.
    apparently rain-Nom fall likely-Concl
    'Apparently it is likely to rain.'

b. *Dooyara ame-ga hur-u.
    apparently rain-Nom fall-Concl
    Lit. 'Apparently it rains.'

    John-Top apparently rain=nom fall likely-Prenom weather-Nom hate
    Lit. 'John hates the weather such that apparently it is likely to rain.'

The diagram (17) indicates the structure of (16a). The modal head Mod undergoes successive-cyclic movement up to Force. The head-cluster is realized as the conclusive form soo-da. In order for the evidential adverb dooyara 'apparently' to have its uninterpretable features deleted, it must then be merged at the left edge of ForceP. From this position, it is able to c-command the functional heads that carry the corresponding interpretable features.

(17)

It also follows that (16b), which lacks a modal, is ungrammatical because the adverb is not able to have its [uModal] feature checked off. The noun-modifying clause in (16c) contains a modal in prenominal form. The structure of the relevant part is shown below.

---

6 We postulate that the head Mod does not carry an inflected modal. Rather, the cluster of functional heads in the dotted circle of (17-18) is converted into a single inflected form. This mechanism is reminiscent of post-syntactic lexical insertion in the Distributed Morphology paradigm (Halle and Marantz 1993, Marantz 1997, Harley and Noyer 1999, Embick and Noyer 2006, among others). There, terminal nodes in syntax consist of morphosyntactic features rather than lexical items.
In this structure, the adverb is able to c-command the modal head. Its [uModal] feature is deleted in accordance with Agree. The modal in (16c) is realized in the prenominal form. Since this form is associated with the Fin head, the whole cluster of functional heads in the dotted circle in (18) is converted to prenominal inflection. Since (18) lacks Force(P), however, the [uForce] feature carried by the adverb fails to find an appropriate goal and cannot be deleted.

Our analysis assumes that the noun-modifying clause in (16c) has a reduced (or truncated) structure, lacking ForceP. If it did contain ForceP, an adverb attached to the left edge of ForceP would be able to c-command the Force head, and the adverb’s [uForce] feature could be deleted. Obviously, this is not the case. Therefore, the noun-modifying clause in (16c) must be no larger than FinP. It may be possible to say, then, that an inflectional form indicates clause size. Only a clause with a conclusive predicate has a full CP structure. A clause with a prenominal predicate only projects up to FinP; one with an infinitival predicate only goes as high as TP and one with a preverbal predicate is simply a vP. This raises a question as to why structural reduction takes place. We return to this issue in section 4.1.

An evidential adverb can occur with only one type of inflection (i.e. conclusive form), but an epistemic adverb is able to occur with either a conclusive or a prenominal predicate as illustrated in (12a-b). As summarized in (15), an epistemic adverb carries the uninterpretable feature [uFin]. This means that it is able to occur in a clause of any size so long as the clause contains Fin(P). Bearing this in mind, let us consider the derivation of a noun-modifying clause.

\[(19) \quad [\text{FinP} \quad \text{Kitto}[^{u\text{Modal}}][^{u\text{Fin}}]] \quad [\text{FinP} \quad [\text{TP} \quad t \quad \text{todoku} \quad \{ [\text{Mod}[^{i\text{Modal}}]^{+v+T}]^{+\text{Fin}[^{i\text{Fin}}]} \}] \quad \text{nimotu}] \quad \rightarrow \text{hazu-no} \quad \text{should-Prenom} \]

In (19), the predicate cluster Mod+v+T moves out of TP and adjoins to Fin. The combined cluster is realized as hazu-no, the prenominal form of the modal. The epistemic adverb can then be attached to the left edge of FinP. From this position, it is able to c-command both Mod and Fin heads, having all of its uninterpretable features deleted.

In contrast, the derivation of a root clause (20a) follows the steps shown in (20b-c).

\[(20) \quad \text{a. Kitto} \quad \text{ame-ga} \quad \text{huru} \quad \text{hazuda.} \quad \rightarrow \text{Certainly, it will rain.'} \]
Adverb-predicate agreement in Japanese and structural reduction

The intermediate step (20b) is more or less similar to (19). The predicate cluster Mod+v+T+Fin is not given a concrete form at this stage but moves further up to Force, resulting in the final structure (20c). At this stage, the modal is realized as conclusive. The parentheses containing the epistemic adverb in (20b-c) indicate that the merger of the adverb may take place at either stage. In both structures, the adverb is able to c-command both Mod and Fin heads.7

The occurrence of an evaluative adverb in multiple positions in (13a-c) is accounted for in the same manner. (21b) represents the structure of the temporal clause in (21a).

(21) a. [Saiwainimo John-ga subayaku kagi-o mituke-ru -to ]
   luckily John-Nom quickly key-Acc find-Inf -when
   mina hottosita.
   everyone was relieved
   Lit. ‘Everyone was relieved when luckily John found the key quickly.’

b. [TP saiwainimo] [TP John-ga [vP subayaku kagi-o t_i] [V+v[T] ] ] -to
   luckily John-Nom quickly key-Acc
   → mituke-ru -when
   find-Inf

Here, the verbal cluster moves out of vP and adjoins to T, which carries the interpretable feature [iT]. Since the evaluative adverb located at the left edge of TP is able to c-command T, Agree takes place successfully, licensing the adverb.

An evaluative adverb is not limited to TP, however. It can be merged with any projection so long as it can c-command T carrying the [iT] feature. Consider the following example. (22a) is a root clause. Its structure is given in (22b).

---

7 As pointed out by a reviewer, (20b-c) yield a linearly identical structure. They can be distinguished by adding an extra constituent, however, as in the following examples.

(i) a. Kitto John-wa Mary-ga syootaisuru hazuda.
   certainly John-Top Mary-Nom invite should
   ‘Certainly Mary should invite John.’

b. John-wa kitto Mary-ga syootaisuru hazuda.
   John-Top certainly Mary-Nom invite should
   ‘Certainly Mary should invite John.’

On the assumption that the topic phrase is located in Spec-TopP, the adverb-topic order in (i) indicates that the adverb is positioned above TopP, whereas the topic-adverb order in (ib) indicates that it occurs below TopP.
Since the predicate is realized in the conclusive form, the sentence has a full CP structure. In (22b), the evaluative adverb merges with ForceP when the verbal cluster moves to Force. The adverb probes the structure and successfully enters into an Agree relation with T contained in the cluster of functional heads.

Notice that (22b) is not the only possible structure for (22a). The evaluative adverb can be merged with either FinP or TP.

(23a-b) represent intermediate stages in the derivation of (22a).\(^8\) (23a) illustrates a structure where the derivation reaches the TP stage. The evaluative adverb can be introduced into it at this point because it is able to c-command T from the left periphery of TP. Similarly, the adverb may also be merged when the derivation proceeds one step further, up to FinP. As illustrated in (23b), the adverb located on the periphery of FinP c-commands T contained within the cluster of functional heads adjoined to Fin. In summary, (22a) has three possible underlying structures, depending on when (and where) the evaluative adverb is merged. Further evidence for this observation will be provided in section 5.

### 3.3 Predictions and implications

The analysis presented above has various consequences. First, it predicts that an evaluative adverb cannot occur inside vP, because it is not able to c-command T from that position. This prediction is borne out.

The vP adverb subayaku ‘quickly’ marks the vP edge.\(^9\) The contrast in (24a-b) thus indicates that an evaluative adverb cannot occur in vP. The structure below illustrates an intermediate

---

\(^8\) The cluster of functional heads has not been given concrete form in (23a-b), as the structures only represent intermediate stages of the derivation.

\(^9\) The adverb subayaku ‘quickly’ forms a constituent with vP. When vP-preposing takes place, it must be moved along with the preposed constituent.
stage of the derivation of (24b). The evaluative adverb fails to c-command T in this structure, resulting in the ungrammaticality.

\[(TP \ John-ga [\varphi \ subayaku \ saiwainimo_{oT}] \ kagi-o \ t_i \ [V+v]+T_{[T]} ]\]

John-Nom quickly luckily key-Acc

Second, the proposed analysis casts doubt on the idea that a phrasal constituent on the left periphery is licensed by a functional head in the same projection. The representation constructed in the syntactic component is read off by an outer interpretive component. The left periphery of a clause is dedicated to representing scope-discourse properties, and the central tenet of the cartographic approach is that scopal properties are determined by local Spec-Head relations (Rizzi 1997, 2006, 2011, 2015). A phrasal category in this domain with a scope-discourse property (such as topic and focus) enters into a biunique Spec-Head configuration with a head bearing a matching feature. It is thus assumed that the linking of a functional head with the specifier in a separate projection is not possible.

In this vein, Cinque (1999) attempts to account for the distribution of adverbs in terms of one-to-one Spec-Head relations. He argues that each class of adverb is located in a distinct specifier position that is associated with a matching functional head. Consequently, the order of adverbs reflects that of the functional heads that enter into agreement relations.

However, APA in Japanese is free from this restriction: It can exhibit a one-to-many relation. An epistemic adverb placed in either ForceP or FinP can thus be linked with Fin (see (19)-(20)). Likewise, an evaluative adverb placed in different projections such as ForceP, FinP, or TP can always enter into an Agree relation with T (see (21)-(23)). The availability of different positions for one adverb can be confirmed by the permutation of word order between the adverb and an adjacent constituent.

(26) a. **Kitto** John-wa yorokondeiru hazuda.
certainly John-Top be.delighted must

b. John-wa **kitto** yorokondeiru hazuda.
John-Top certainly be.delighted must

‘John must be delighted.’

(27) a. **Kusikumo** John to Bill-wa onazi yume-o mita.
strangely John and Bill-Top same dream-Acc dreamed

(i) a. [Subayaku kagi-o mituke-sae ]; John-ga t_i sita.  
quickly key-Acc find-even John-Nom did

‘Quickly find the key, John did.’

b. *[Kagi-o mituke-sae ]; John-ga subayaku t_i sita.  
key-Acc find-even John-Nom quickly did

Lit. ‘Find the key. John did quickly.’

Additionally, when it occurs with a vP-related adverbial phrase like *heya-de* 'in the room', they can occur in either of the following orders.

(ii) a. John-ga subayaku heya-de kagi-o mituketa.  
John-Nom quickly room-in key-Acc found

‘John found the key in the room quickly.’

b. John-ga heya-de subayaku kagi-o mituketa.  
John-Nom room-in quickly key-Acc found

‘John found the key quickly in the room.’

This adverb order does not occur between subayaku ‘quickly’ and saiwainimo ‘luckily’ in (24a-b), which indicates that the latter is not a constituent of vP.
b. John to Bill-wa kusikumo onazi yume-o mita.
   John and Bill-Top strangely same dream-Acc dreamed
   ‘Strangely, John and Bill dreamed the same dream.’

At first glance, this would seem to be unpredictable in Cinque’s (1999) analysis, where each adverb is allotted just one position. His claim is that when one and the same adverb occurs in two different positions, each instance of the adverb occupies the specifier position of a different functional head. He further argues that the association of different functional heads can give rise to interpretive differences.

As far as English is concerned, his account makes the right predictions. The adverb cleverly is interpreted differently depending on the position which it occupies. When it is placed in the clause-final position as in (28a), it has a pure manner interpretation. On the other hand, it takes on a subject-oriented interpretation when it occurs in the position immediately after the subject in (28b).

(28) a. John has answered their questions cleverly.
   (the way John answered their questions was clever.)
   b. John cleverly has answered their questions.
   (It is clever of John to have answered their questions.)  (Cinque 1999: 19)

If interpretive differences are associated with locational ones, it is further predicted that the same adverb can occur twice in a single clause. This prediction is borne out.

(29) John cleverly has answered their questions cleverly.  (Cinque 1999: 19)

Cinque’s (1999) argument, however, cannot be carried over to the word order variations in (26-27), since the a- and b-examples do not exhibit semantic differences. Furthermore, the simultaneous occurrence of the same adverb in two different positions results in ungrammaticality in Japanese.

   certainly John-Top certainly be.delighted must
   ‘John must be delighted.’
   b. *Kusikumo John to Bill-wa kusikumo onazi yume-o mita.
   strangely John and Bill-Top strangely same dream-Acc dreamed
   ‘Strangely, John and Bill dreamed the same dream.’

Given these facts, the word order variations in (26-27) suggest that one and the same adverb can be placed in different positions in the phrase structure in accordance with the analysis proposed here. An S-adverb can be licensed in different positions so long as it can find its goal in its c-command domain.

Needless to say, the mere existence of one-to-many relations does not immediately force the abandonment of the Criterion-based approach pursued by Cinque (1999). A more constructive and interesting question would be what factors distinguish languages that employ one-to-one licensing and those that tolerate one-to-many patterns. This issue will not be pursued here, but it certainly deserves further in-depth investigation.

4 Truncation (aka structural reduction)

It was suggested in section 3.2 that clauses may vary in size. We pointed out the possibility that while some clauses have full structure and are equipped with all the functional projections
forming the CP domain, others have a smaller structure, without some CP-related projections. This section offers further evidence that this is indeed the case. Here we argue that truncation (aka structural reduction) is not something to be stipulated but can instead be derived from a more general principle of grammar.

### 4.1 Complement selection and head movement

Complementizers do specify clause types. The English complementizer that appears in either declarative or factive clauses. Another complementizer while indicates that the clause it subordinates is a temporal one. Japanese makes a finer distinction between declarative and factive clauses in the choice of complementizer. The former is headed by to, whereas the latter by koto. Another conspicuous trait of Japanese complementizers is that each item may select a different inflectional form. Consider the following examples.

\[(31)\]
\[
\begin{align*}
   & a. \text{John-wa [ Mary-ga totemo sinsetu-}{\text{da / *na}} \text{] } \text{omotteiru.} \\
   & \quad \text{John-Top Mary-Nom very kind-is.\{ Concl / Prenom \} that think} \\
   & \quad \text{‘John thinks that Mary is very kind.’} \\
   & b. \text{John-wa [ Mary-ga totemo sinsetu-}{\text{na / *da}} \text{] } \text{koto } \text{-o sitteiru.} \\
   & \quad \text{John-Top Mary-Nom very kind-is.\{ Prenom / Concl \} that \text{-Acc know} \\
   & \quad \text{‘John knows that Mary is very kind.’} \\
   & c. \text{John-wa [ aomukeni ne-}{\text{ru / ta}} \text{] } \text{-mama } \text{ugoka-naka-tta.} \\
   & \quad \text{John-Top on.his.back lie-\{ Pres / Past \} \text{-with move-not-Past} \\
   & \quad \text{‘Lying on his back, he didn’t move.’} \\
   & d. \text{Karera-wa [ ohiru-o tabe-}{\text{Ø / *ru / *ta}} \text{] } \text{-nagara } \text{hanasiai-o sita.} \\
   & \quad \text{they-Top lunch-Acc eat-\{ Prev / Pres / Past \} \text{-while discussion-Acc did} \\
   & \quad \text{‘They had a discussion while eating lunch.’}
\end{align*}
\]

While the declarative to selects a conclusive predicate in (31a), the factive koto selects a prenominal one in (31b). The item mama in (31c) occurs in an adverbial clause that refers to a circumstance attendant on the event expressed by the matrix clause. It also selects an infinitival predicate, which can only take a present or a past tense morpheme. Finally, the temporal complementizer nagara in (31d) is compatible only with a preverbal predicate.

Predicate selection by a complementizer is not simply a syntactic process. It also has a morphological effect: predicate and complementizer form a morphological unit. As standardly assumed, items forming a morphological unit cannot be separated by a constituent (Kageyama 1993, etc.). The combination of a stem and affix (for instance, bake-d) does not allow the intervention of a modifier (cf. *bake-skilfully-d the cake vs. bake-ed the cake skillfully*). The predicate-complementizer combination in Japanese also exhibits this sort of morphological integrity.

\[(32)\]
\[
\begin{align*}
   & a. \text{*... [ Mary-ga sinsetu-da totemo to ] ...} \\
   & \quad \text{Mary-Nom kind-is.Concl very that} \\
   & b. \text{*... [ Mary-ga sinsetu-na totemo koto ] -o ...} \\
   & \quad \text{Mary-Nom kind-is.Prenom very that \text{-Acc} } \\
   & c. \text{*... [ ne-ta aomukeni -mama ] ...} \\
   & \quad \text{lie-inf on.his.back \text{-with}} \\
   & d. \text{*... [ ohiru-o tabe-Ø yukkuri -nagara ] ...} \\
   & \quad \text{lunch-Acc eat-Prev leisurely \text{-while}}
\end{align*}
\]
This requirement does not apply to complementizer-predicate relations in English. Although a complementizer must co-occur with a predicate in the appropriate form, they can be syntactically separated.

(33) a. \textit{It is believed [that Mary \{ is / *to be \} kind].}

b. \textit{It is desirable [for Mary \{ to stay / *stays \} here].}

\textit{That} and \textit{for} select finite and infinitival predicates respectively. Still, the relation is purely syntactic. They can be separated from predicates by the subject.

Morphological integrity is typically observed among various lexical items. The fact that predicate-complementizer combinations in Japanese exhibit this property may lead one to conclude that they are stored in the lexicon as single lexical items. However, a closer examination reveals that they exhibit morphological integrity only partially. For instance, a lexical item that exhibits strong morphological integrity does not allow a pro-form to occur inside it. Thus, it is not possible to replace \textit{yama} ‘mountain’ in the compound \textit{yama-nobori} ‘mountain-climbing’ with a pronoun \textit{sore} ‘it’ (i.e. \textit{*sore-nobori}). By contrast, a predicate-complementizer combination allows the predicate to be replaced with a verbal pro-form \textit{soo suru} ‘do so’.

(34) a. \textit{John-wa Mary-o sikatta. Bill-wa [ Ken-mo soo suru to ] omotta.}

John-Top Mary-Acc scolded Bill-Top Ken-also so do that thought ‘John scolded Mary. Bill thought that Ken would do so too.’

b. \textit{John-wa Mary-o sikatta. Bill-wa [ Ken-mo soo sita koto ]-o sitteiru.}

John-Top Mary-Acc scolded Bill-Top Ken-also so did that -Acc know ‘John scolded Mary. Bill knows that Ken did so too.’

c. \textit{John-wa me-o tozita -mama suwatte-ita. Bill-mo [ soo sita -mama ]}

John-Top eye-Acc closed -with sitting-was Bill-also so did -with sitting-was ‘John was sitting with his eyes closed. Bill was also sitting doing so.’

d. \textit{John-wa koohii-o nomi -nagara hon-o yonde-ita.}

John-Top coffee-Acc drinking -while book-Acc reading-was \textit{Bill-wa [ soo si -nagara ] zassi-o yonde-ita.}

Bill-Top so doing -while magazine-Acc reading-was ‘John was reading a book while drinking coffee. Bill was reading a magazine while doing so.’

Kageyama (1993) argues that items exhibiting partial morphological integrity form units by syntactic head movement. Adopting this idea, we propose that the morphological integrity observed in Japanese predicate-complementizer combinations is also attained as a result of head movement when the predicate complex adjoins to the complementizer.

Keeping this in mind, let us now consider why some subordinate clauses are structurally reduced. Suppose, for the sake of argument, that the factive complement in (31b) has a full CP structure. The predicate appears in prenominal form, which means that it moves to Fin and is realized there. Recall that it must adjoin to the complementizer \textit{koto} in order to satisfy the morphological requirement imposed on predicate-complementizer combinations. However, such movement would skip the Force head, giving rise to a violation of the Head Movement Constraint (HMC).
One could say that a HMC violation may be circumvented if the predicate stopped at Force and then moved on. Still, recall that a cluster of heads containing Force is realized in the conclusive, rather than prenominal form. This would then give rise to an ill-formed sentence (36b).

(35) ... Mary-ga { totemo t; [TP [A+Cop+v+T]+Fin] FinP Force FinP -koto Mary-Nom very } → sinsetu-na 'kind-is.Prenom' \\


It follows then that Force(P) must be absent so that the complementizer -koto can successfully combine with a prenominal predicate without violating the HMC.

The same conclusion can be drawn from predicate-modal combinations. Like complementizers, modals also select particular inflectional forms they combine with.

(36) a. ... Mary-ga { totemo } [TP t; [FinP [A+Cop+v+T]+Fin] Force FinP -koto Mary-Nom very } → sinsetu-da 'kind-is.Concl'


Yooda ‘seem’ and hazuda ‘should’ both select a prenominal predicate. Sooda ‘likely’ selects a preverbal one.

As with complementizers, modals also form a morphological unit with the predicate they select for. Moreover, a predicate-modal sequence does not allow the intervention of another constituent.

(37) a. John-wa Mary-ga { kirai- } [TP t; { na / *da } yooda. Mary-Nom { hate- } [Prenom / Concl ] seem ‘It seems that John hates Mary.’

b. John-wa Mary-ga { kirai- } [TP t; { na / *da } hazuda. Mary-Nom { hate- } [Prenom / Concl ] should ‘John should hate Mary.’

c. Ame-ga { hur- } [TP t; { i / *u } sooda. rain-Nom { fall- } [Prev / Concl ] likely ‘It is likely to rain.’

Similar to predicate-complement combinations, those involving predicate-modal combinations exhibit partial morphological integrity. Below, a predicate selected by a modal may involve a pro-form.
   John-Top Mary-Acc praised Bill also so. Prenom seem
   ‘John praised Mary. Bill seems to have done so too.’

   John-Top Mary-Acc praised Bill also so do. Prenom should
   ‘John praised Mary. Bill should do so too.’

   John-Top Mary-Acc praised Bill also so do. Prev likely
   ‘John praised Mary. Bill is likely to do so too.’

The partial integrity of a predicate-modal combination receives the same explanation as that of a predicate-complementizer combination: the predicate forms a single unit with the modal via head movement. Since pronominal and preverbal inflections are associated with Fin and v respectively, it follows that the complement of the modal in (39a-c) must have a reduced structure, in compliance with the HMC. More specifically, the modals yooda and hazuda select for FinP, and sooda for vP.

Unlike predicate-modal combinations, the combination of a verb and its clausal complement does not exhibit morphological integrity. Here an extra element can intervene between the complementizer and the matrix verb that selects it.

(40) [Mary-ga sinsetu-da to ] John-wa omotteiru.
   Mary-Nom kind-is Concl that John-Top think
   ‘John thinks that Mary is kind.’

A full CP complement differs from a reduced one in terms of clausehood. While the former has a clause boundary, the latter does not. One diagnostic phenomenon that can detect a clause boundary is scrambling. As is well-known, scrambling can take place either inside a single clause or across a clause boundary. The former is known as clause-internal scrambling (41a) and the latter long-distance scrambling (41b).

   John-Nom the apple-Acc Mary-Nom ate that said
   ‘John said that the apple, Mary had eaten.’

   the apple-Acc John-Nom Mary-Nom ate that said
   ‘The apple, John said that Mary had eaten.’

While clause-internal scrambling can remedy a violation of binding Condition A, long-distance scrambling cannot (Saito 1992, Nemoto 1993, Tada 1993). Consider the following examples.

   John-Nom each other-Gen teacher-Nom them-ACC criticized that said
   Lit. ‘John said that each other’s teachers had criticized them.’

   John-Nom them-ACC each other-Gen teacher-Nom criticized that said
   Lit. ‘John said that each other’s teachers had criticized them.’

   each other-Gen teacher-Nom John-Nom them-ACC criticized that said
   Lit. ‘Each other’s teachers said that John had criticized them.’
The a-examples of each pair represent violations of Binding Condition A. In this configuration, the anaphor *otagai-no* 'each other’s' fails to be bound by its antecedent *karera-o* 'them'. (42b) shows that Condition A is not violated by clause-internal scrambling. Long-distance scrambling in (43b), on the other hand, does not bring about this mitigation effect. There, the clausal complement headed by to ‘that’ is a full ForceP clause since the predicate is realized in the conclusive form. The contrast between (42b) and (43b) can thus be used for determining whether scrambling crosses a clausal boundary.

Let us now consider how scrambling affects binding relations in sentences involving a predicate-modal sequence.

(44) a. *Otagai-no*, sensei-ga karera-o, hihansita yooda.
    each.other-Gen teacher-Nom them-Acc criticized.Prenom seem
b. *Karera-o*, otagai-no, sensei-ga ___ hihansita yooda.
    them-Acc each.other-Gen teacher-Nom ___ criticized.Prenom seem
Lit. ‘Each other’s teachers seem to have criticized them.’

(45) a. *Otagai-no*, sensei-ga karera-o, hihansuru hazuda.
    each.other-Gen teacher-Nom them-Acc criticize.Prenom should
b. *Karera-o*, otagai-no, sensei-ga ___ hihansuru hazuda.
    them-Acc each.other-Gen teacher-Nom ___ criticize.Prenom should
Lit. ‘Each other’s teachers should criticize them.’

    each.other-Gen teacher-Nom them-Acc criticize.Prev likely
b. *Karera-o*, otagai-no, sensei-ga ___ hihansi sooda.
    them-Acc each.other-Gen teacher-Nom ___ criticize.Prev likely
Lit. ‘Each other’s teachers are likely to criticize them.’

The b-examples in each pair are all grammatical, indicating that they involve clause-internal scrambling. This entails that there is no clausal boundary between the modal and its complement. In other words, the latter has a reduced structure.

4.2 Deriving truncation

The present analysis lends support to the observation made by Endo (2012) and Endo and Haegeman (2015) to the effect that truncation does not have to be stipulated, but can be derived from more general syntactic principles, namely, locality of movement. According to these authors, the locality principle at work in Japanese is the HMC.

In this connection, Haegeman (2010, 2012) proposes a slightly different account of the non-occurrence of S-adverbs than the one presented here. While ascribing it to an intervention effect, she observes that the relevant mechanism involves phrasal movement rather than head movement. Adopting Bhatt and Pancheva’s (2006) observation, she argues that the adverbial clauses in (47a-c) involve the movement of a null operator to the left periphery, which is then blocked by an intervening S-adverb. This is schematically illustrated in (48).

(47) a. "[If the students *apparently* can’t follow the discussion in the third chapter], we’ll do the second chapter.

b. "[If George *probably* comes], the party will be a disaster.

c. "[If they *luckily* arrived on time], we will be saved. (Haegeman 2010: 603)
(48) *[ If ... apparently / probably / luckily ... OP ... ], ... 

However, Haegeman’s (2010, 2012) account does not carry over to Japanese. For one thing, an adverbial clause in English does not tolerate an evaluative, evidential, or epistemic adverb. In Japanese, however, an evaluative adverb may occur in the same context.

(49) [ Mosi un’yoku sigoto-ga hayaku aware-ba ] paatii-ni maniau. 
if luckily job-Nom early is,over-if party-Dat in.time

Lit. ‘If we luckily finish the job early, we will be in time for the party.’

This suggests that the distribution of Japanese S-adverbs is subject to different conditions. Alternatively, we propose an Agree-based account where the distribution of Japanese S-adverbs is determined by agreement with predicates (see section 3).

There are in fact good reasons for doubting the involvement of a null operator in Japanese adverbial clauses. Consider, for example, the weak island effect. A weak island is a domain involving an operator which blocks extraction of a non-referential expression (see Cinque 1990, Rizzi 1990 among others). The following examples involve wh-movement.

(50) a. ?Which problem do you wonder [ how to solve t ]?

b. *How do you wonder [ which problem to solve t ]?

c. How do you think [ that Bill could solve the problem t ]? (Rizzi 1990: 73)

In (50a), the referential wh-phrase which problem is extracted from an embedded clause that contains another wh-phrase. By contrast, wh-extraction is banned in (50b): how is a non-referential expression and its movement is blocked by an intervening wh-phrase. No such intervention effect arises in (50c), due to the absence of an intervening operator.

Given this property of weak islands, it is predicted that if a Japanese adverbial clause involves a null operator, it will block extraction of a non-referential expression. If no such operator is present, however, a non-referential expression should be able to undergo extraction. As we see below, there are abundant pieces of evidence in favor of the second possibility.

As pointed out by Yoshida (2006), a numeral quantifier is a non-referential expression. It can be scrambled as in (51a), but scrambling is blocked in (51b) by the intervening focus particle -sika ‘only’ — an island-inducing operator. Crucially, scrambling of a numeral quantifier from an adverbial clause is possible, as indicated by (51c). This suggests the absence of an operator in the adverbial clause.

(51) a. Huta-tu, John-ga ringo-o t tabe-ta.
2-Cl John-Nom apple-Acc eat-Past. ‘John ate two apples.’

b. ‘Huta-tu, [ISLAND John-sika ringo-o t tabe-naka-tta ].
2-Cl John-only apple-Acc eat-not-Past ‘Only John ate two apples.’

2-Cl John-Top Mary-Nom apple-Acc eat-if be.surprised will ‘John will be surprised if Mary eats two apples.’

The same conclusion can be drawn from the following sets of data.

(52) a. Nanika, John-ga Mary-ni purezento-o t ageta (koto) some John-Nom Mary-Dat present-Acc gave fact ‘John gave Mary some present.’
b. *Nanika* [ISLAND John-sika Mary-ni prezento-o t. age-naka-tta (koto) some John-only Mary-Dat present-Acc give-not-Past fact 'Only John gave Mary some present.'

c. Nanika. Mary-wa [ John-ga prezento-o t. kureta-ra ] some Mary-Top John-Nom present-Acc give-if kare-to tukiau tumorida. him-with go.out intend 'Mary intends to go out with John if he gives her some present.'

(53) a. Dare-kara-mo. John-wa [ Mary-ga t. okane-o morawa-naka-tta to ] itta. who-from-Foc John-Top Mary-Nom money-Acc receive-not-Past that said 'John said that Mary hadn't received money from anyone.'

b. ?*Dare-kara-mo. John-wa [ISLAND Mary-ga t. okane-o morawa-naka-tta who-from-Foc John-Top Mary-Nom money-Acc receive-not-Past kadooka ] tazuneta. whether asked 'John asked whether Mary hadn't received money from anyone.'

c. Dare-kara-mo. John-wa [ Mary-ga t. okane-o moraw-anakere-ba ] who-from-Foc John-Top Mary-Nom money-Acc receive-not-if kanozyo-o homeru tumorida. her-Acc praise intend 'John intends to praise Mary if she doesn't receive money from anyone.'

(Nasu 2015:359)

(54) a. John-wa [ Mary-ga ittai nani-o katta to ] omotteiru no? John-Top Mary-Nom the.hell what-Acc bought that think Q 'What the hell does John think Mary bought?'

b. ?*John-wa [ISLAND Mary-ga ittai nani-o katta kadooka ] John-Top Mary-Nom the.hell what-Acc bought whether sitteiru no? know Q Lit. 'What the hell does John know whether Mary bought?'

c. John-wa [ Mary-ga ittai nani-o kae-ba ] yorokobu no? John-Top Mary-Nom the.hell what-Acc buy-if be.delighted Q Lit. 'What the hell will John be delighted if Mary buys?'

(55) a. John-wa [ Mary-ga naze sore-o katta to ] omotta no? John-Top Mary-Nom why it-Acc bought that think Q Lit. 'Why, does John think [ that Mary bought t. ]?'

b. *John-wa [ISLAND Mary-ga naze sore-o kau kadooka ] John-Top Mary-Nom why it-Acc buy whether sintagatteiru no? want.to.know Q Lit. 'Why, does John want to know [ whether Mary buys t. ]?'

(Lasnik and Saito 1992: 15)

c. ?Kimi-wa [ Taro-ga naze kaisya-o yamere-ba ] manzokusuru no? you-Top Taro-Nom why company-Acc resign-if be.satisfied Q Lit. 'Why, will you be satisfied [ if Taro resigns from the company t. ]?'

(Fukui 1988: 520, n. 13)
Each of the a-examples in (52-55) contains a non-referential expression: an indefinite quantifier in (52a), a negative polarity item in (53a), an aggressively non-D-linked expression in (54a), and an adjunct wh-expression naze ‘why’ in (55a). The b-sentences instantiate a weak island effect. There, an intervening operator prevents the linking of the non-referential expression to the periphery of the matrix clause. The c-sentences involve conditional clauses. They are all grammatical, indicating that an adverbial clause does not constitute a weak island, and hence lacks a null operator.

It appears then that while operator movement plays an important role in determining the distribution of S-adverbs in English, it does not do so in Japanese. Instead, Agree(ment) with an appropriate head is the primary factor governing the distribution of S-adverbs. This cross-linguistic difference suggests that the (non-)occurrence of S-adverbs is caused by different mechanisms.

4.3 Some speculations on the cause(s) of truncation “effects”

Generally speaking, the terms “truncation” or “structural reduction” are used when a given clause exhibits properties that imply that some projections are missing. Still, it is not necessarily the case that a clause resisting a periphery-related phenomenon has a literally truncated structure. Bošković (1997) argues that a complement clause lacking an overt complementizer is smaller in size: It does not contain the CP layer but simply consists of a TP as in (56b).

(56) a. John believes [CP that [TP Bill is kind]].
   b. John believes [TP Bill is kind].

Kishimoto (2006) argues, however, that a clause without an overt complementizer can still be a full CP headed by a phonetically null complementizer, as illustrated in (56c).

(56) c. John believes [CP Ø [TP Bill is kind]].

This suggests that the absence of a periphery-related phenomenon is not sufficient to establish actual truncation, namely, the absence of a projection that could host it.\(^\text{10}\)

To avoid confusion, we use the term “truncation effect” to refer to the absence of periphery-related phenomena, while keeping the term “truncation” for the genuine absence of the upper

\(^{10}\) A reviewer asks why a clause without an overt Force marker cannot be analyzed as containing a null Force head in Japanese. One candidate for such a clause is the so-called factive clause.

(i) John-wa [Mary-ga sinsetu-na -koto ]-o sitteiru.
   John-Top Mary-Nom kind-is.Prenom -C-Acc know
   ‘John knows that Mary is kind.’

This clause lacks the Force-marking complementizer -to. The null Force analysis would postulate the following structure for such a clause.

(ii) ... [Mary-ga sinsetu-na -koto Ø_{Force} ]-o
    Mary-Nom kind-is.Prenom -C -Acc

This analysis is untenable, however. The null complementizer analysis presupposes that the null complementizer can alternate with an overt one. When applied to (ii), it predicts incorrectly that the null Force head in it can be replaced with the overt Force marker to. The co-occurrence of -koto and -to is not possible (iiia). Nor is it possible to totally replace the factive complementizer -koto with -to (iiib).

(iii) a. *John-wa [Mary-ga sinsetu-na -koto -to ]-o sitteiru.
    John-Top Mary-Nom kind-is.Prenom -C -C-Acc know
    b. *John-wa [Mary-ga sinsetu-na -to ]-o sitteiru.
    John-Top Mary-Nom kind-is.Prenom -C-Acc know
    ‘John knows that Mary is kind.’

We thus conclude that a factive clause does not contain a null Force head. See also the discussion concerning prenominal inflection in factive clauses (35)-(36).
part of a clause. Again, it is worth emphasizing that the presence of a truncation effect is not itself a deciding factor in determining whether a given clause has a full or truncated structure.

In connection with the treatment of truncation effects, Haegeman’s (2010, 2012) operator movement analysis is worth special attention. Haegeman argues that although certain adverbial clauses exhibit a truncation effect (such as the non-occurrence of an S-adverb), they still may have a full CP structure. Thus although the adverbial clauses in (47a-c) exhibit a truncation effect, they may be equipped with all the projections that constitute the CP domain. The non-occurrence of an S-adverb there is not due to the lack of a position but rather due to an intervention effect. Truncation in (47a-c), therefore, is only illusory.

Another point to bear in mind is that even if the truncation effect results from actual truncation (i.e. absence of upper CP-related projections), simply pointing it out is merely a description rather than an explanation. The fundamental question is what causes (upper) part of a clause to be absent. To elaborate, consider the following data.

(57) a. *[If probably George comes], the party will be a disaster.

   b. *[Tabun George-ga kuru-ba], paatii-wa dainasi daroo.

   Lit: ’If probably George comes, the party will be a disaster.’

Both sentences exhibit a truncation “effect” in that the conditional clause does not license an S-adverb to occur. However, this effect may be caused by different factors. Recall Haegeman’s (2010, 2012) claim that an English conditional clause has a full CP structure and that the truncation effect results from the intervention of operator movement. On the other hand, a Japanese conditional clause does not involve an operator, but does have reduced structure (see section 4.2). The complementizer ba ‘if’ does not select for conclusive or prenominal verbs (hence, “kur-u-ba ‘come-Prenom/Concl-if’”). This means that a conditional clause lacks ForceP and FinP (see section 4.1). The exclusion of an S-adverb must therefore result from genuine truncation: there is no position for a functional head to license the adverb in a conditional clause.

One remaining question is what the difference between English and Japanese is rooted in. It may reduce to the type of movement: while A-bar movement (i.e. operator movement) causes the truncation effect in English, head movement does so in Japanese. This is probably too simple a generalization, however. As pointed out by Endo and Haegeman (2015), head movement can play a role in yielding a truncation effect in English. Consider the following data.

(58) a. If you see her tomorrow, call me.

   b. If tomorrow you see her, call me.

   c. Should you see her tomorrow, call me.

   d. *Should tomorrow you see her, call me. (Endo and Haegeman 2015: 32)

Topicalization of an adjunct inside a conditional clause is usually tolerated as in (58b), but it is forbidden if the conditional clause is formed by head movement of the modal should (58d). Endo and Haegeman (2015) ascribe the ungrammaticality of (58d) to a violation of the HMC. Following Rizzi (1997), they assume that the adjunct tomorrow is located in the specifier position of some functional projection, say TopP, and that T-to-C movement of should is blocked by the head of this projection.

(59) *CP  [TopP tomorrow [ Top [ … you should see her ][ ]]], call me.

If their analysis is on the right track, it would appear that head movement does indeed play a role in producing this effect in English as well as in Japanese.
Likewise, it is not necessarily the case that operator movement does not produce a truncation effect in Japanese. Consider the following contrast noted by Tomioka (2015: 279).

(60) a. \[ \text{OP: } \text{Mari-} \{ \text{ga} / \text{wa} \} \text{ tito } \text{hutta to } \text{Ken-ga omotteiru } \text{ hito} \\
             \text{Mari-} \{ \text{Nom / Top} \} \text{ dumped that Ken-Nom think person} \\
             \text{‘the person that Ken thinks that Mari dumped.’} \]

b. \[ \text{OP: } \text{Kei-o } \text{hutta to } \text{omotteiru } \text{ hito} \\
             \text{Mari-} \{ \text{Nom / Top} \} \text{ Kei-Acc dumped that think person} \\
             \text{‘the person who thinks that Mari dumped Kei’} \]

(60a) exhibits a truncation effect: topicalization is banned. Still, it does not entail the actual truncation of the to-clause. As shown in (60b), topicalization is possible there as well. Moreover, the to-clause is a full CP containing ForceP since its predicate is in the conclusive form. Nevertheless, these sentences differ in the base position of the relative operator. In (60a), it originates inside the to-clause. On its way to the surface position, it crosses the topicalized constituent. In contrast, no such intervention occurs in (60b), as the operator is base-generated in the higher clause. The truncation effect (i.e. the illegitimacy of topicalization in (60a)), thus, is attributable to an intervention effect.

To sum up, it may not be appropriate to capture cross-linguistic differences in the cause of truncation effects only in terms of the type of movement. Still, English and Japanese do differ in the type of movement they employ for the derivation of adverbial clauses. Further discussion of this issue is beyond the scope of the present paper; we leave it for future research.

5 Interactions between S-adverbs

This section discusses the advantages of an Agree-based analysis of APA such as the one proposed in this paper. It demonstrates how our analysis can successfully account for certain restrictions imposed on the alignment of S-adverbs in a single clause.

According to Cinque (1999), different classes of adverbs occupy the specifier position of different functional heads. Since functional projections are aligned in a fixed order, adverbs should also follow the same ordering pattern. The S-adverbs discussed in this paper would be arranged in the following order on Cinque’s analysis.

(61) \[ \text{MoodPevaluative Evaluative } \ldots \text{MoodPevidential Evidential } \ldots \text{ModPepistemic Epistemic } \ldots \]

One consequence of this template is that when two adverbs belonging to different classes co-occur in a clause, they will obey a strict ordering constraint. It is predicted that among the six logically possible combinations listed below, (62a, b, d) are legitimate while the others are not.

---

11 A tacit premise underlying the analysis in (60) is that relativization in Japanese involves operator movement. As noted by Ishii (1991) and Murasugi (1991), a Japanese restrictive relative clause exhibits a subjacency violation.

John-Nom Mary-Nom gave child -Acc scolded boy

Lit. ‘the toy, [which John scolded the child, [to whom Mary gave e],].’

It is not possible to associate the head noun omotya ‘toy’ with the gap inside the complex NP island. Given that an island effect is a hallmark of movement, it may be concluded that operator movement takes place inside relative clauses in Japanese as well as other languages.
The distribution of S-adverbs in Japanese does not conform perfectly to the predicted patterns. Combinations corresponding to (62a, b, d, f) are attested (see (63, 64, 66, 68)), but (62c, e) are not. Contrary to prediction, sentences in (65) and (67) are grammatical. Apparently, the distribution of S-adverbs in Japanese is not governed by the template (61).

(62) a. Evaluative > Evidential √
b. Evaluative > Epistemic √
c. Evidential > Evaluative ×
d. Evidential > Epistemic √
e. Epistemic > Evaluative ×
f. Epistemic > Evidential ×

The distribution of S-adverbs in Japanese does not conform perfectly to the predicted patterns. In English, this prediction is mostly borne out, with only one exception. (62a-e) exhibit the predicted patterns, but (62f) does not. Consider the following sentences.
(i) a. Unbelievably, he was seemingly won the game. (Amano 1999: 223)
b. Martin, fortunately, will probably come tomorrow. (Amano 1999: 224)
c. *Seemingly, he has, unbelievably, won the game. (Amano 1999: 224)
d. Well, I’m not surprised; she certainly has probably attracted a ton of other college guys. (Haumann 2007: 371)
e. *Martin, probably, will fortunately come tomorrow. (Amano 1999: 225)
f. I believe it would be disruptive for us to continue our program since this probably would obviously put it in the House of Representatives and be disruptive to the country. (COCA; 1992)

It remains unclear why (i,f) is possible. We leave this question for future research.
Norio Nasu
115

(68) Epistemic > Evidential (62f)

*Kitto* John-wa *dooyara* toosensuru *nitigainai* yooda.
surely John-Top apparently be.elected must seem
Lit. ‘Surely John apparently must be elected.’

As before, we propose that the sentences in (63)-(68) can be accounted for via a locality condition on Agree.

The relevance of locality to Agree has been repeatedly taken up by Chomsky (2000, 2001, 2004) ever since it was advocated as a mechanism of feature-checking. Hiraiwa (2001) shows convincingly that locality can only be evaluated derivationally. Specifically, in a configuration “X ... Y ... Z”, where X is a probe and Y and Z are goals, an intervention effect will arise in cases where the probe for Z and the probe for Y are “derivationally distinct” (p. 71). The phrase “derivationally distinct” refers to a situation where Y enters into Agree with a probe W which is distinct from X before Z enters into Agree with X. This is schematically illustrated in (69).

(69) a. W ... Y ... Z
   [uF] [iF] [iF]
   Agree

b. X ... W ... Y ... Z
   [uF] [uF] [iF] [iF]
   Agree

Suppose that W is introduced into the derivation before X. It carries an uninterpretable feature [uF], so it will act as the probe and identify Y as its goal, which carries the feature [iF]. This is shown in (69a). X is merged at the next stage (69b). According to Hiraiwa’s derivationally defined locality condition, Agree between X and Z will then be blocked.

In consideration of Hiraiwa’s observations, Chomsky (2004) also argues that intervention matters for Agree in configurations akin to (69b). He remarks that “intervention effects will hold only if the intervening element is not rendered inactive by P[robe] itself” (p. 115). Although W has its [uF] checked off in (69b) (and has become inactive), deletion of [uF] cannot be brought about via Agree with X. Rather it results from Agree with Y. In this sense then, W can be said to intervene between X and Z.

One common aspect of Hiraiwa’s (2001) and Chomsky’s (2004) approaches is that locality does matter when there are two instances of Agree. This is exactly what is found in the sentences (63)-(68). Each of them contains two S-adverbs which are licensed via Agree. This means that two instances of Agree are involved in these sentences. It is predicted then that a derivationally defined intervention effect arises in the ungrammatical (68), but not in the well-formed sentences (63)-(67).

Let us examine whether this prediction is borne out. In (66), the epistemic adverb *kitto* ‘surely’ enters into a probe-goal relation with the epistemic modal *nitigainai* ‘must’. This is indicated as Agree 1 in (70) below. The evidential adverb *dooyara* ‘apparently’ enters into a probe-goal relation with the evidential modal *yooda* ‘seem’ (indicated as Agree 2).

(70) Evidential > Epistemic (= (66))

```
dooyara ... [FinP kitto ... nitigainai+v+T+Fin ] yooda+v+T+Fin+Force
[uModal][uForce] [uModal][uFin] [iModal][iFin] [iModal][iForce]

+-------------------------+ Agree 1
|                         |
|                         |
+-------------------------+ Agree 2
```
The two instances of Agree take place in separate projections: Agree 1 occurs in FinP, whereas Agree 2 takes place in ForceP. Given that locality is calculated in terms of “closest c-command” (Chomsky 2000: 122), Agree 1 and Agree 2 do not interfere with each other.

In contrast, the ungrammatical sentence (68) involves crossing links. Consider the schematic representation below.

(71) Epistemic > Evidential (= (68))

\[
\begin{array}{c}
\text{kitto} \ldots [\text{dooyara} [\text{FinP} \ldots [\text{nitigainai}+\text{v}+\text{T}+\text{Fin}]] \text{ yooda}+\text{v}+\text{T}+\text{Fin}+\text{Force} ] \\
[u\text{Modal}][u\text{Fin}] [u\text{Modal}][u\text{Force}] [i\text{Modal}][i\text{Fin}] [i\text{Modal}][i\text{Force}] \\
\text{Agree 2}
\end{array}
\]

When the adverb dooyara is merged, it probes the structure and Agree 1 takes place. At this stage, the adverb kitto has not been introduced into the structure. This means that it will count as a derivationally distinct probe from dooyara, because dooyara intervenes between them.

An intervention-based analysis, however, runs into problems when (63) and (64) are taken into consideration. Two probe-goal links cross in these sentences, similar to the case in (68/71), yet they remain grammatical.

(72) Evaluative > Evidential (= (63))

\[
\begin{array}{c}
\text{saiwainimo} \ldots [\text{dooyara} [\text{FinP} \text{ toosensita}+\text{v}+\text{T}+\text{Fin}]] \text{ yooda}+\text{v}+\text{T}+\text{Fin}+\text{Force} ] \\
[u\text{T}] [u\text{Modal}][u\text{Force}] [i\text{T}] [i\text{Modal}][i\text{Force}]
\end{array}
\]

(73) Evaluative > Epistemic (= (64))

\[
\begin{array}{c}
\text{saiwainimo} \ldots [\text{hyottositara} [\text{FinP} \text{ toosensuru}+\text{v}+\text{T}+\text{Fin}]] \text{ kamosirenai}+\text{v}+\text{T}+\text{Fin}+\text{Force} ] \\
[u\text{T}] [u\text{Modal}][u\text{Fin}] [i\text{T}] [i\text{Modal}][i\text{Fin}]
\end{array}
\]

This apparent quandary can be solved by appealing to the concept of feature-based relativized minimality widely adopted in the literature (cf. Starke 2001, Rizzi 2004, Endo 2007, Abels 2012, Haegeman 2012, among others). Rizzi (2004) observes that relativized minimality effects are found “within the same feature class but not across classes” (p.243). An intervention effect, in other words, is sensitive to feature classes rather than to types of movement. In the structure “X ... Y ... Z”, Y does not count as an intervening element between X and Z if its feature set consists of features that are distinct from those carried by X and Z.

The absence of an intervention effect in (72, 73) can thus be accounted for on this approach. An evaluative adverb does not share any feature with evidential or epistemic adverbs. Consequently, a probe-goal relation between an evaluative adverb and its corresponding predicate is not interrupted by other types of S-adverb.

The proposed analysis can account for (65, 67) as well. Recall that Cinque’s (1999) template (61) would rule out the evidential > evaluative order in (65) and the epistemic > evaluative order in (67). These are both grammatical sentences, however. Their grammaticality is successfully captured in the present analysis. The derivation of (65) takes place along the following lines.
(74) **Evidential > Evaluative (= (65))**

a. \[
\begin{array}{c}
\text{[FinP } \text{saiwainimo } \text{[FinV } \text{toosensita+}v+T+\text{Fin } ]} \\
\text{[uT]} \\
\text{[iT]}
\end{array}
\]

b. \[
\begin{array}{c}
\text{[ModalP } \text{[FinP } \text{saiwainimo } \text{toosensita+}v+T+\text{Fin } ] \text{ } \text{yooda }] \\
\end{array}
\]

c. \[
\begin{array}{c}
\text{[ForceP } \text{dooyara } \ldots \text{[FinP } \ldots \text{[ModalP } \text{[FinP } \text{saiwainimo } \text{toosensita } ]} \ldots ] \\
\text{[uModal]}[uForce] \\
\text{[voda+v+T+Fin+Force ]} \\
\text{[uModal] [iForce]}
\end{array}
\]

The evaluative adverb *saiwainimo* 'luckily' is merged at (74a). Agree occurs because the cluster of functional heads corresponding to the main verb *toosensita* 'was-elected.' The evidential modal *yooda* 'seem' is merged in the next step (74b). The evidential adverb *dooyara* 'apparently' is merged when ForceP is constructed, and it Agrees with the cluster corresponding to the modal (74c). No crossing occurs between the two probe-goal links. Likewise, probe-goal links do not cross in the derivation of (67).

(75) **Epistemic > Evaluative (= (67))**

a. \[
\begin{array}{c}
\text{[FinP } \text{saiwainimo } \text{[FinV } \text{toosensuru+}v+T+\text{Fin } ]} \\
\text{[uT]} \\
\text{[iT]}
\end{array}
\]

b. \[
\begin{array}{c}
\text{[ModalP } \text{[FinP } \text{saiwainimo } \text{toosensuru+}v+T+\text{Fin } ] \text{ } \text{kamosirenai }] \\
\end{array}
\]

c. \[
\begin{array}{c}
\text{[ForceP } \text{hyottositara } \ldots \text{[FinP } \ldots \text{[ModalP } \text{[FinP } \text{saiwainimo } \text{toosensuru } ]} \ldots ] \\
\text{[uModal]}[uFin] \\
\text{[voda+v+T+Fin+Force ]} \\
\text{[uModal] [iFin]}
\end{array}
\]

Agree occurs in (75a) when the evaluative adverb is merged. The epistemic modal *kamosirenai* 'may' then merges with FinP in the higher cycle (75b). Finally, the epistemic adverb *hyottositara* 'possibly' merges at the left edge of ForceP and enters into an Agree relation with the functional head cluster corresponding to the modal (75c). Here again, the two probe-goal links do not intervene with each other.

6 Conclusion

In this paper, we have offered a critical evaluation of some basic tenets commonly assumed in cartographic studies of clause structure: the universal arrangement of functional projections, and the one-to-one relation between a head and its specifier. The distribution of S-adverbs in Japanese has been shown to be a useful testing ground.

The distribution of English S-adverbs conforms to the template of functional projections proposed by Cinque (1999), but Japanese S-adverbs exhibit different patterns. Another trait not found in English is the obligatory co-occurrence of S-adverbs with a particular type of predicate and inflectional form. The phenomenon of Adverb-Predicate Agreement (APA) was discussed in this paper. The APA is asymmetric in that the occurrence of an S-adverb depends on the type of predicate and its inflectional form, rather than the converse. A similar asymmetry is also found in probe-goal relations. The APA can thus be captured by the same mechanism, Agree.

The analysis put forward here correctly predicts that some S-adverbs in Japanese can occur at the edge of more than one functional projection. So long as they enter into an Agree relation with the appropriate functional head, they can occur in any position. In this respect, the distribution of Japanese S-adverbs represents a departure from a principal assumption of the
In another novel aspect of the present analysis is the assumed correspondence between clause structure and the inflectional form of the predicate, each of which is associated with a particular functional head. Structural reduction (or truncation) then derives from a morphologically determined selectional relationship between the predicate and a complementizer. In other words, a complementizer selects a predicate with a particular inflectional form. The selectional requirement is met when the predicate undergoes head movement and adjoins to the complementizer. Any clause whose predicate takes a form other than the conclusive has a reduced structure without higher projections in the CP domain.

Lastly, this paper examined the ordering restrictions imposed on co-occurring S-adverbs. It was shown that an analysis which presupposes a one-to-one relation between an S-adverb and a functional head is incapable of accounting for the facts. Among the several logical combinations of evidential, epistemic, and evaluative S-adverbs in Japanese, any order except the epistemic-evidential is possible. The unattested combination is due to a blocking of Agree between the epistemic adverb and the predicate by the intervening evidential adverb. The other combinations are legitimate either because no intervention occurs, or because the two adverbs do not share any common features and feature-based relativized minimalism does not play any role.

Two important questions remain. First, although Japanese S-adverbs can adjoin to the left edge of more than one functional projection, it does not entail the immediate rejection of an approach which presupposes a one-to-one Spec-Head relation. In some languages, the distribution of S-adverbs does exhibit patterns in compliance with this restriction. An interesting question is what factors distinguish languages that employ one-to-one licensing and those that tolerate one-to-many patterns. Second, we have argued that languages may not be parameterized with respect to truncation effects only in terms of the type of movement involved. The underlying factors of the cross-linguistic differences have yet to be discovered. These questions are certainly worth addressing, and should be addressed in the future.

7 References


SOV-X: Syntactic and pragmatic constraints of the postverbal domain in Turkish

Tamer Akan & Katharina Hartmann
akan.tamer@gmail.com
k.hartmann@lingua.uni-frankfurt.de

1 Introduction

Turkish is an SOV language in both, main and embedded clauses, as illustrated in (1).¹

(1) a. Terzi elbise-m-i bitir-miş.
dress maker dress-1SG.POSS-ACC finish-EV.3SG
‘Apparently the dressmaker has finished my dress.’ (Göksel and Kerslake 2005: 337)

b. Ben terzi-nin elbise-m-i bitir-dig-i-ni bil-yor-um.
I dress maker-GEN dress-1SG.POSS-ACC finish- NMLZ-POSS-ACC know-PRES-1SG
‘I know that the dressmaker has finished my dress.’

Although being verb final in its default form, constituents may appear in the postverbal domain. The postverbal domain may host sentential and non-sentential constituents. This is a well-known fact about Turkish, which has been frequently described in the literature (Erguvanlı 1984, Öztürk 2005, Şener 2010, Gürer 2015). In this paper, we will be concerned only with non-sentential constituents in the postverbal domain, thus with PPs and DPs, see (2ab) for examples.

(2) a. Çiçek al-acak-miş [ŋ bir kız için].
flower buy-FUT-EV.3SG a girl for
‘He will buy flowers for a girl.’ (Veld 1993: 275, (37b))

b. Ali sat-ti [ŋ ev-i].
Ali sell-PFV house-ACC
‘Ali sold the house.’ (Göksel and Kerslake 2005: 343, (46c))

The purpose of this paper is twofold. First, we discuss some peculiar categorial and information-structural properties of the postverbal domain in Turkish. We do this in a contrastive way by comparing Turkish to German, another SOV language. German also makes use of a postverbal domain, but with different syntactic and pragmatic restrictions. Starting with the categorial properties of the postverbal domain in section 2, we show that Turkish is much less restricted than German concerning type and complexity of the postverbal constituents. In sections 3 and 4 we consider

¹ We would like to thank Rahul Balusu for his comments on an earlier version of this paper. We also thank the audience of the HFL Workshop in May 2018 in Freudental. Special thanks are due to the organizers of this workshop and the editors of this volume, Yvonne Viesel and Josef Bayer. All remaining errors are ours.

¹ The final verb appears in bold letters throughout the article. In our examples, we are following the glosses as established by the Leipzig Glossing Rules, see https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf.
information-structural differences between the postverbal domains of the two languages. We show that whereas Turkish is pragmatically restricted in that it only allows non-focal constituents to occur postverbally, the German postverbal domain is insensitive to information-structure in that it allows focal as well as non-focal constituents to occur in the postverbal position (the “Nachfeld” in topological terminology). Second, we develop a syntactic structure for Turkish, which is superior to previous proposals in that it establishes a tight connection between the observed syntactic and information-structural properties. Concretely, we provide evidence for low and high information structural projections in the Turkish clausal spine. In the spirit of Rizzi (1997) we assume that these projections are above the TP, which represents the clausal core in Turkish. We assume different IS-projections for each information-structural status, hence, in this order, topic, focus and discourse anaphoric phrases. The verb is always attracted to the head of the focus phrase. Postverbal constituents are therefore derived by leftward movement of given constituents to discourse anaphoric projections, and verbal movement to the higher focus head. Section 4 develops this proposal. In section 5 we corroborate our theory by adding further data. Section 6 concludes.

2 Formal properties of the postverbal domain

In this section, we discuss some of the most important properties of the postverbal domain in Turkish. For a better illustration, we use a contrastive method comparing Turkish to German. Both languages have a SOV basic word order. Differing from Turkish, German has verb second order in main clauses (see, e.g. Thiersch 1978, Taraldsen 1986, Vikner 1995). We investigate whether structural complexity has an impact on word order variation with respect to the postverbal domain; we discuss categorial restrictions in the two languages and consider effects of the Right Roof Constraint.

The structural complexity of a constituent (syntactic and/or prosodic) may have an impact on word order. It has been claimed that the complexity of a constituent correlates with the frequency of the occurrence in the postverbal domain (Ross 1967, Arnold et al. 2000, Bader et al. 2013 Hawkins 2004, Francis 2010, among others). A plausible reason for this preference may be the avoidance of center-embedded structures and therefore a facilitation of language processing. In a self-paced reading task, Francis (2010), corroborating Hawkins (2004), argues that grammatical weight affects the processing efficiency of extraposed relative clauses modifying a subject-DP. Francis (2010) shows that reading time advantages increase with the grammatical weight of the relative clause. The examples in (3) illustrate light, medium and heavy extraposed relative clauses, respectively. The most complex postverbal relative clause in (3c) was produced fastest.

(3) a. Three people arrived here early yesterday morning [who were from Chicago].
   b. Three people arrived here early yesterday morning [who were from a northern suburb of Chicago].
   c. Three people arrived here early yesterday morning [who were originally from a far northern suburb of Chicago which is called Lake Forest]. (Francis 2010: 21)

Structural complexity has a strong impact also on extraposition in German. German does not allow extraposition of simple DPs. If, however, the DP is structurally complex, it may well appear postverbally. This correlation between complexity and late occurrence in the clause was first observed by Behaghel (1909) as the “Law of Increasing Terms” (= “Gesetz der wachsenden Glieder”), which is illustrated in (4). In (4a), the postverbal demonstrative pronoun dasjenige is

---

2 The German postverbal domain is categorically restricted in that it may not host non-complex DPs. This fact will be briefly addressed in section 3.
extended by a relative clause; in (4b), in contrast, the nominal expressions do not contain an attributive modifier and are therefore banned from occurring postverbally.

(4) a. Die Kuh hat Hörner, um in sich hineinzusenden [DP dasjenige, the cow has horns to in REFL send.inside whatever was astralisch-ätherisch gestalten soll].
what astro-etheral shape should
‘The cow has horns to send inside of her whatever should give astral and ethereal shape.’ (Rudolf Steiner, Süddeutsche Zeitung, 26.2.11)

b. *Die Kuh hat Hörner, um in sich hineinzusenden [DP dasjenige / den Äther].
the cow has horns to in REFL send.inside whatever the ether

In a corpus study on non-sentential constituents in the postverbal domain in German, Proske (2010) found that 6% of all postverbal constituents are DPs. The majority (77%) of these are instances of right dislocation, followed by discontinuous conjuncts (15%) and complex argument DPs (8%). There is not a single instance of a non-complex DP in the German postverbal domain. Most of the postverbal argument DPs are presentational foci, a representative example of which is given in (5).

(5) Auf Gleis eins fährt ein: [die verspätete Regionalbahn nach Wiesbaden].
on platform one arrives PRT DEF delayed regional.train to Wiesbaden
‘The delayed train to Wiesbaden is arriving on platform one.’

It is worth noting that the Law of Increasing Terms appears to restrict only DPs in German. PPs do not seem to fall under it. This is shown in (6a), where the extraposed PP consists of a preposition taking a pronominal DP as a complement. (6b) represents a minimal pair which is ungrammatical due to a categorial violation: although identical in syllable number, the relatively light DP is excluded in the postverbal position. This shows that complexity alone is not the only principle behind extraposition in German but that in addition, the category of the extraposed constituent matters as well.

(6) a. Ich weiß, dass ich mich nicht getäuscht habe [PP in ihm].
I know that I me NEG was.wrong have in him
‘I know that I wasn’t wrong about him.’

b. *Ich weiß, dass ich nicht getraut habe [DP der Frau].
I know that I NEG trusted have the woman

In Turkish, on the other hand, the structural complexity of a constituent does not influence its syntactic position in relation to the verb: the postverbal position is available for complex as well as for non-complex constituents (Kornfilt 2005). In (7), the postverbal object DP is structurally complex in that it contains a nominalized deverbal complement. (8) illustrates that elements as light as pronouns may occupy the postverbal position in Turkish as well.

(7) Ben bil-iyor-um [DP öğretmen-in kişisel bir problem-den dolay] I know-PRS-1SG teacher-GEN personal a problem-ABL due.to
yarnki sinav-in saat-i-ni değiştirecek-i-ni].
tomorrow exam-GEN time-POS-ACC change-NMLZ-POSS-ACC
‘I know that the teacher is going to change the time of the exam tomorrow due to a personal problem.’
One important difference between the Turkish and the German postverbal position is their ability to employ DPs in this position. As shown above, German is quite restrictive when it comes to DP extraposition. Turkish, on the other hand, allows DPs of all grammatical roles in the postverbal domain as shown in (9). (9a) illustrates a verbfinal sentence; (9bc) show that direct and indirect objects may occur postverbally; (9d) exemplifies a postverbal subject. Notice that all grammatical roles could also be expressed pronominally in postverbal position.

Turkish also allows the occurrence of multiple constituents in the postverbal domain without any categorial restrictions. In example (10), two DPs occur postverbally, in (11) a DP and a PP are placed after the verb.

A further notable difference between German and Turkish concerns the behavior of postverbal constituents wrt. The Right Roof Constraint (Ross 1967:185). In German it is known that the Right Roof Constraint prohibits extraposition across a clausal boundary, see (12) from Müller (1995:215). In (12a), PP-extraposition targets the complement clause; it does not cross a clausal boundary. In (12b), the PP extraposes across the adverbial clause, which violates the Right Roof Constraint.
The Right Roof Constraint does not hold in Turkish: constituents from within an embedded clause may also appear postverbally as seen in (13).

(13) Öğretmen [öğrenciler t. oku-du diye] bil-iyor [kitab-i],
    teacher student-PL read-PST C knowPRS book-ACC
    'The teacher knows that the students read the book.'

To sum up the discussion so far, the Turkish postnominal domain is not restricted by conditions having to do with structural complexity. It allows structurally heavy and light constituents to occur postverbally. Multiple constituents are possible after the final finite verb. Turkish is not subject to the Right Roof Constraint, which prohibits structural dependencies between a constituent's base position and its postverbal position across an embedded clause. Turkish contrasts with German, where extraposition is clearly conditioned by complexity.

The next two sections focus on pragmatic conditions of the postverbal domain. Section 3 shows that German does not exhibit any pragmatic restrictions with respect to the postverbal domain. Section 4 discusses Turkish, which contrasts with German in that it is discourse-configurational. We develop a syntactic account which is able to account for the observed patterns in Turkish.

3 No information-structural restrictions in the German postverbal domain

In this section we argue that the German postverbal domain (the "Nachfeld") is not constrained by information-structural restrictions. The postverbal domain thus pragmatically resembles what is traditionally called the "Vorfeld", i.e. the first position in the German main clause which precedes the finite verb in second position. The Vorfeld is pragmatically also unrestricted and may host a focal, a topical, or a given constituent (or an expletive pronoun). This indicates that none of the German peripheries is endowed with specific information-structural features. With respect to the final, postverbal domain, the following data from a broadcast interview with Dorothee Bär, by the time representative Secretary-General of the CSU, the Christian-Social Union in Bavaria (Deutschlandfunk, 16.11.2013)\(^3\). The content of the interview are the upcoming coalition negotiations. It can be seen that topic, given and new information may all appear in the postverbal domain in German. All postverbal constituents are PPs. (14a) represents the first question of the interviewer to Dorothee Bär; the postverbal PP is an aboutness topic which takes up the theme of the interview. In (14b), the PP refers back to the main theme and can be categorized as given. (14c) is an example with multiple postverbal PPs. The first PP gives new information, which has not been introduced before. The information given here cannot be considered as focal since we take focus to invoke alternatives (Rooth 1985, 1992). In (14c), however, the content of the first PP is rather marginal for the line of argumentation pursued here. There are no suitable alternatives available. The second PP represents the complement of the main verb. It paraphrases a statement of the interviewer's question and can be therefore considered as given. We take examples such as (14c) as evidence for the differentiation between focus and new information, as advocated for, e.g. in Katz and Selkirk (2011).

\(^3\) The interview is available in written form under: https://www.deutschlandfunk.de/ber-merkel-und-seehofer-muessen-in-so-einer-koalition-die.694.de.html?dram:article_id=269190.
(14) a. *Topic*

Dobovisek: “Es geht hoch her [bei Ihrem möglichen Koalitionspartner].”
Dobovisek: it goes high on with your possible coalition partner.
‘There is plenty going on with your possible coalition partner.’

b. *Given*

Bär: “Da muss eben klar sein, wer eben das Agenda-Setting macht
Bär: it must clear be who DEF agenda-setting makes
[in dieser Koalition].”
‘It has to be clear who is setting the agenda in this coalition.’

c. *New / Given*

Bär: „Ja, es ist zumindest so, dass ich schon mich gewundert habe
Bär: yes it is at least so that I myself wondered have
an der einen oder anderen Stelle in der letzten Woche [über das
in the one or other point in DEF last week about DET
Gebaren in den einzelnen Arbeitsgruppen auch einzelner SPD-
behavior in DEF individual task.forces also individual SPD-
Abgeordneter].”
‘Yes, it is at least the case that I was wondering here and there during the last week
about the behavior in the individual task forces also of individual SPD-deputies.’

It is often not trivial to determine the information-structural status of constituents in a text since
the typical strategies in elicitational settings are not available. This is especially the case for focus,
which even in question-answer sequences, which typically occur in interviews, cannot be easily
detected since the speakers often freely deviate from the Question Under Discussion (QUD).4
Therefore, Balbach and Hartmann (2012) carried out an acceptability study on post-verbal focus.
The aim of the study was to better control for the IS-status of the constituent in the postverbal
position by giving specific contexts. The method of the test was the following: B&H (2012) con-
structed 6 test sentences with PPs in preverbal and postverbal position. The test sentences were

4 A nice example where the addressee modulates the QUD set by the speaker is given in (i), an example is
from the same interview. Bär’s answer does not give the requested reason as to why her party wants
the grant coalition with the socialists. Instead she rejects that this is a volitional act but suggests that the party
*has to* enter the coalition. This is a case of verum focus. The given explanation (the voter’s will) elaborates
the implicit contrastive predicate of obligation.

(i) **Dobovisek:** Aber wenn Sie schon jetzt Zweifel haben an dem Zuspruch innerhalb der
Dobovisek: if you now doubts have at DEF acceptance within DEF
SPD-Basis [...] warum wollen Sie dann überhaupt eine Koalition mit der SPD
SPD-basis why want you then at.all INDEF coalition with DEF SPD
eingehen?”
‘But if you now have doubts wrt. the acceptance within the SPD basis, why do you want to enter in a
coalition with the SFB at all?’

**Bär:** Na ja gut, wollen das ist natürlich auch ein Auftrag, den wir haben.
Bär: PRT PRT okay want DEM is of.course also a mandate REL we have
So viele Optionen bleiben nicht.”
so many options rest NEG
‘Okay, want – this is of course also a mandate that we have. That many options do not remain.’
embedded in small contexts triggering different IS-interpretations on a PP, namely given, new, presentational and contrastive focus, see (15). In total, this gave 48 test sentences. 20 Filler sentences were added. The contexts and test sentences were read aloud and recorded. The acoustic data was presented to 16 test persons who were asked to rate their acceptability on a scale from 1 (very good) to 5 (very bad). In a second phase of the experiment, the data were presented in written form to 100 test persons in order to find out whether there is a modality effect.

(15) a. **Given**: mentioned before, inferable/accessible from the text or the situation  
b. **New**: not mentioned before, not inferable/accessible  
c. **Focus**:  
   (i) presentational focus: new information focus induced by a wh-question  
   (ii) contrastive focus

In the following, we present one test item. The PP was either in preverbal (“in situ”, (16a)) or in postverbal position (“ex situ”, (16b)).

(16) **Target sentence**

a. Am Wochenende wurde er im Garten von einer Hummel _gestochen_.  
   on.the weekend was he in.DEF garden by INDF bumblebee stung  
   ‘On the weekend he was stung by a bumblebee in the garden.’

b. Am Wochenende wurde er im Garten gestochen von einer Hummel.  
   on.the weekend was he in.DEF garden stung by INDF bumblebee  
   ‘On the weekend he was stung by a bumblebee in the garden.’

(17) a. **Context for informational status GIVEN**  
   Peter ist _allergisch gegen Insektenstiche aller Art_.  
   Peter is _allergic against insect.bites of.each kind_  
   ‘Peter is allergic to insect stings of any kind. Especially during the summer, this poses a problem when bumble bees are out in the garden.’

b. **Context for informational status NEW**  
   Manchmal _passieren Markus doch seltsame Sachen_.  
   sometimes _happen Markus PRT strange things_  
   ‘Sometimes strange things happen to Markus.’

c. **Context for informational status WH-FOCUS**  
   Sandra fragt Norman: ”Von _was wurde Stefan am Wochenende im Garten gestochen_?”  
   Sandra asks Norman by what was Stefan _on.DEF weekend in.DEF garden stung_  
   ‘Sandra asks Norman: “By what was Stefan stung on the weekend in the garden?”’

   Norman antwortet: [...]  
   Norman _answers: [...]_
d. Context for informational status WH-FOCUS

Anna sagt, dass Axel am Wochenende im Garten von einer Wespe verfolgt wurde. Aber Bernd berichtigt sie: „Nein, …”

‘Anna says that Axel was chased by a wasp on the weekend in the garden. But Bernd corrects her: “No, …”’

The results of the task are given in the following diagram. The blue bars represent the acoustic results, the red bars give the written results.

Diagram 1

The results show a clear difference in acceptability between PPs in the in situ and the ex situ position. This holds for the oral as well as for the written test. The result confirms the generally perceived (stylistic?) markedness of non-sentential phrases in postverbal position in German. It expresses an interesting disproportion between the production of postverbal non-sentential phrases, which in the corpus TüBa D/Z represent after all 18% of all postverbal constituents (Proske 2010), and the conscious perception of them, which is always conceived of as somehow degraded. The diagram also shows that the results from the written presentation of the target sentences are slightly degraded in comparison to those from the oral presentation, especially with the ex situ conditions. We take this to reflect the tendency to normative standards in written language. Very relevant for the discussion of the present subsection is the relative acceptability of the extraposed PPs across all information-structural conditions. This shows that apart from given and new postverbal PPs, which were attested by the data in (15), focused constituents (new in-
formation and contrastive focus) may also occur postverbally. Note that new information is considered slightly marked in the in situ position as well. In lack of an explanation for this observation, we attribute this result to independent factors.

The pragmatic facts of the German postverbal position do not suggest a specific syntactic analysis and we will therefore not pursue a syntactic approach for German here. As we will see in the next section, Turkish differs from German in that it exhibits clear information-structural restrictions with respect to the postverbal domain. Based on this insight, we develop a novel syntactic account of the Turkish clause, which is based on leftward movement exclusively.

4 IS restrictions in the Turkish postverbal domain

As opposed to what has been discussed in the previous section for German, discourse configurationality in Turkish is syntactically represented, see e.g. Erguvanlı (1984), Kornfilt (1997), Kural (1997), Şener (2010), not only by articulated functional projections for topic and focus above the propositional core V P, see Rizzi (1997), but also for low given, discourse anaphoric phrases (Bellettı 2004, Akan 2009). We assume that the postverbal domain is derived by raising the verb from its final base position above the respective category to the head of the focus phrase (FocP). All projections that the FocP dominate represent the postverbal domain.

4.1 Given and New information in the Turkish postverbal domain

4.1.1 Givenness in the postverbal position

With respect to information structure, Turkish postverbal constituents are considered to be backgrounded (Erguvanlı 1984, Kural 1997, Hoffman 1998, Kılıçaslan 2004, Kornfilt 2005, Doğruöz and Backus 2007, Özge and Bozsahin 2010), activated topics (Erkül 1986, Schroeder 1995) or tails (İşsever 2003). Postverbal constituents in Turkish may represent discourse given information. This is shown in (18) where a constituent mentioned in a wh-question, 'kitab' ('book') is repeated in the postverbal position in the answer.

\[(18)\] Q: Kitab-a ne ol-du?  
book-DAT what be-PST  
‘What happened to the book?’

A: Ahmet Berna-ya ver-miş o kitab-].  
Ahmet Berna-DAT give-EV.3SG that book-ACC  
‘Ahmet gave Berna that book.’

We assume that the givenness interpretation of the postverbal constituent results from movement of the given phrase to a discourse anaphoric phrase (DaP) above TP, which corresponds to the low topic phrase of Rizzi (1997), or to the Familiarity Phrase (FamP) of Frascarelli and Hinterhölzl (2007). As stated above, the verb raises to the head of the higher focus phrase. In (18), the indirect object is focused and moves to the specifier of FocP. The subject is a topic and appears in the high TopP. The full structure of (18A) is in (19).
Postverbal nominal elements must be case-marked. Case-marking is obligatorily correlated with definiteness in Turkish (Erkü 1986, Erguvanlı 1984, Nakipoğlu 2009) and we can easily account for this correlation: Given DPs always refer back to established discourse referents; hence they are naturally interpreted as definites. In Turkish this is expressed by case-marking. Since the postverbal domain hosts discourse anaphoric DPs, case-marking is required. In preverbal position, case-marking is optional as illustrated by the following example. We assume that the preverbal DP in (20A2) is a topic and as such exempt from the case-marking requirement.

As already discussed in section 2, the postverbal domain can host several discourse anaphoric phrases, as exemplified in (22). In order to account for multiple postverbal constituents, we assume that the DaP may be recursive. The structure of (21A) is given in (22).
(21) Q: Leyla-dan ne haber? Kaya-yla evlen-di mi?
Leyla-ABL what news Kaya-COM marry-PST Q
‘What about Leyla, did she and Kaya get married?’
A: Sen-in haber-in yok mu? ...
you-GEN news-POSS no Q
‘You don’t know?’
Oya evlen-di [iki yıl önce][Kaya-yl] / [Kaya-yl] [iki yıl önce]
Oya marry-PST two years ago Kaya-COM / Kaya-COM two years ago
‘Oya married Kaya two years ago.’

(22)

The two discourse anaphoric phrases are hierarchically ordered given that the first DaP c-commands the second. This can be shown with data involving binding. In (23), the reciprocal anaphor can only be bound by a DP from a preceding, i.e. structurally higher position. We will come back to hierarchy effect in the postverbal domain in section 5.

  love-PRS Ayşe and Ali each other-ACC
  ‘Ayşe and Ali love each other.’

b. *Sev-iyor birbirleri-ni Ayşe ve Ali
  love-PRS each other-ACC Ayşe and Ali

4.1.2 New information in the postverbal position

Discourse-anaphoric constituents are not the only possible information structural unit in the postverbal domain. New information, which we assume not to be focal, see section 2, can also be found in the postverbal domain in Turkish. According to Özge and Bozsahin (2010) and Kılıçaslan (2004), new information is common in the postverbal domain in Turkish. In (24B), two constituents appear in the postverbal domain. They are not given as they are not mentioned in the preceding
question. But they are also not focal since they do not introduce alternatives. Example (25) illustrates the same state of affairs with a simple postverbal constituent.

(24) A: Hayırdır, yüz-ün gül-üyor?
    what's up face-POS.2SG laugh-PROG
    'What's up, you are all smiles?'

B: Evet, çok sev-diğ-im bir arkadaş-im gei-di
    yes very love-NMLZ-POSS.1SG a friend-POS.1SG come-PST
de [bu sabah] [Adana-dan].
PRT this morning Adana-ABL
    'A good friend came this morning, from Adana.'
(Özge and Bozsahin 2010: 149, fn 16)

(25) Q: İstakoz-dan ne haber? On-u kim ye-di?
    lobster-ABL what news it-ACC who eat-PST
    'What about the lobster? Who ate it?'
    (Kılıçaslan, 2004, p. 731, (25))

A: On-u Ali ye-di [birkaç gün önce].
    it-ACC Ali eat-PST several day before
    'Ali ate it several days ago.'

For phrases encoding new information, which is neither focal, nor discourse anaphoric, we assume that they remain in their base position within vP. Focus constituents, on the other side move to the specifier of FocP. The tree in (26) shows the structure we propose for (25A).

(26) TopP
    onu FocP
    Ali Foc'
    Foc TP
    yedi T'
    vP T
    PP P
    birkaç gün

4.1.3 (Contrastive) Topics

Contrastive topics cannot appear postverbally in Turkish. The topic marker ise indicates a switch from one topic to another one and thus functions as a topic shifter. We assume that ise may realize the head of the TopP. This is shown in example (27). (27A) establishes the context, and
the last sentence in (27B1) shows the topic phrase preceding the subject in the left periphery, see the structure in (28). Phrases introduced by ise are excluded postverbally (27B2).

(27) A: Akşam misafir-ler için bir sürü yiyerek yap-muş-ti-m. evening guest-PL for a load food make-EV-PST-1SG
   Sen bil-iyor mu-sun, kim ye-di kek-i?
   you know-PRS QP-2SG who eat-PST cake-ACC
   'I made a lot of food for the guests last night. Do you know who ate the cake?'

B1: Bil-diğ-im kadar-ı-yla kek-i Ahmet ye-di,
    know-NMLZ-POSS.1SG much-POSS-COM, tart-ACC Ahmet eat-PST
    pasta-yı (ise) Hasan ye-di.
    cake-ACC TOP Hasan eat-PST
    'As far as I know, Ahmet ate the cake, Hasan ate the pancake.'

    Hasan eat-PST cake-ACC TOP
    'Hasan ate the cake.'

(28)

This concludes our survey on postverbal constituents in Turkish. In the following section, we consider focus realization in Turkish accounting for the fact that it is excluded from appearing postverbally.

4.2 No focus in the Turkish postverbal position

4.2.1 Presentational focus realization in Turkish

Not all scholars agree on the position of presentational focus in Turkish. Some argue that presentational focus is possible in any preverbal domain (Göksel and Özsoy 2000). Another school, however, argues that presentational focus is realized exclusively in the immediately preverbal domain (İşsever 2003, Akan 2009). We follow the second view. According to the present proposal, in presentational focus construction, the verb moves to the head of a FocP and the focal phrase moves to its specifier position. This is shown in (29) where the caseless DP in the immediate preverbal position represents the focus of the clause. The focus constituent may be preceded by another phrase, which is the topic of the sentence. Following Rizzi (1997), we assume that it is in the specifier position of the highest topic phrase. The structure of (29) is given in (30) below.
(29) Q: Gizem ne ye-di?
   Gizem what eat-PST
   ‘What did Gizem eat?’
A: Gizem dondurma ye-di.
   Gizem ice cream eat-PST
   ‘Gizem ate ice cream.’

(30)

```
TopP
  DP  FocP
  Gizem  ne/dondurma  Foc'
         yedi  TP
```

It is also argued that presentational focus may not be shifted further to the left (İşsever 2003:1034), showing that the verb necessarily moves to Foc. This is shown in (31) below. Again, the wh-pronoun and the focus constituent in the answer share the same position. If the focus phrase is moved across the topic, the sentence is ungrammatical, see (31A2).

(31) Q: Fatma’yı kim arı-yor?
   Fatma-ACC who call-PRES
   ‘Who is calling Fatma?’
   A2: *Ali Fatma’yı arıyor.  (İşsever 2003: 1034)

The focus position is available for all kinds of syntactic phrases. Apart from the focused subject, illustrated in (31), focused objects (32), focused adverbials (33), and focused PP-arguments (34) – all appear in the immediately preverbal position, which we claim to be SpecFocP.

(32) Q: Ali kim-i arı-yor?
   Ali who-ACC call-PRS
   ‘Who is Ali calling?’

(33) Q: Ali toplantı-da sınırlı mi-ydi?
   Ali meeting-LOC aggressive QP-PST
   ‘Was Ali aggressive in the meeting?’
   yes Ali room-ABL door-ACC slamming get.out-PST
   ‘Yes, he left the room slamming the door.’

(34) Q: Peki hırsız-ı gör-ünce ne yap-mış çocuk?
   well thief-ACC see-when what do-PST child
   ‘Then what did the child do when he saw the thief?’
   silently table-GEN under-POSS-DAT hide-PST
   ‘He hid under the table silently.’
Concerning predicate focus, we assume that the verb still moves to Foc, and the remant VP is raised to SpecFocP. This is illustrated for a distransitive VP in (35). No part of the focused VP may appear in postverbal position, cf. (35A2/A3).

(35) Q: Senem dün ne yap-tı?
Senem yesterday what do-ACC
‘What did Senem do yesterday?’

A1: Senem dün [VP kitab-ı Ahmet-e t₂] götür-dü tVP.
Senem yesterday book-ACC Ahmet-DAT take-PST
‘Senem took the book to Ahmet yesterday.’

A2: *Senem dün kitab-ı götür-dü Ahmet-e.
Senem yesterday book-ACC take-PST Ahmet-DAT

A3: *Senem dün Ahmet-e götür-dü kitab-ı.
Senem yesterday Ahmet-DAT take-PST book-ACC

If only the verb is focused, as in the following example (36), there is of course no restriction with respect to the other verbal constituents, which freely occur postverbally.

(36) Q: Senem tavuklar-ı ne yap-tı?
Senem chicken-PL-ACC what do-ACC
‘What did Senem do with the chicken?’

Senem chicken-PL-ACC cut-PAST
‘Senem slaughtered the chicken.’

A2: Senem kes-ti tavuk-lar-ı.
Senem cut-PST chicken-PL-ACC
‘Senem slaughtered the chicken.’

All new sentences are obligatorily verb final, see (37). The question triggers focus on the complete proposition. Hence the TP raises to SpecFocP after verb movement to the head of the FocP.

(37) a. Q: Bu gürültü ne?
this noise what?
‘What is this noise?’

A: Bahçivanlar çim-ler-ı biçiyor.
gardener-PL lawn-PL-ACC cut-PRS
‘the gardeners are mowing the lawn.’
4.2.2 Contrastive focus realization in Turkish

In the previous section, we have seen how presentational focus is realized in the immediate preverbal position. Now, we turn our attention to how contrastive focus is realized in Turkish. Contrary to presentational focus, contrastive focus is also possible in any preverbal domain in addition to the immediately preverbal area, which gives a clue that contrastive focus may be realized higher in the sentence and does not have to be verb adjacent. However, contrastive focus can never be realized in the postverbal domain. In this paper, we assume a high position for contrastive focus in addition to the low presentational focus position. Topics can be fronted and precede the contrastively focused constituent. In (38), contrastively focused constituents are marked with capital letters.

‘As far as I know, the teacher gave the students the books.’

B1: Bil-diğ-im kadar-i-yla, know-NMLZ-POSS.1SG much-POSS-COM
ÖDEV-LER-lı öğretmen öğrenci-ler-e ver-miş, kitap-lar-lı değil.
‘The teacher gave assignments to the students, not the books.’

B2: Bildiğim kadarya, öğretmen ÖDEVLERİ öğrenciilere vermiş, kitapları değil.
B3: Bildiğim kadarya, öğretmen öğrencilere ÖDEVLERİ vermiş, kitapları değil.

As can be seen in (38), the contrastively focused constituent does not have to be adjacent to the verb but may occupy higher structural positions. We assume that it moves through SpecFoc to receive a focus interpretation. It may then move up to a higher focus position below TopP. Following Rizzi (1997), we assume that topic phrases may be located above and below the highest FocP. (39) gives the tree for the example in (38B2). Note that the constituent preceding contrastive focus is interpreted as the topic of the sentence. The dative object öğrenci-ler-e (‘students-PL-DAT’) probably occupies a low TopP / a DapP, which is reserved for familiar / given phrases.

(39)

4.2.3 No focus in the postverbal domain

It was previously mentioned in the preceding sections that focus may not appear in the postverbal domain. The empirical data are given below. The constituent corresponding to the wh-word is the focus of the sentence and must appear preverbally. The postverbal position is excluded for wh-
words and foci (Göksel and Özsoy 2000, İşsever 2003, Öztürk 2005, Gürer 2015), hence, (40A) is not ungrammatical but inadequate as an answer to the preceding question.

(40) Q: Gizem ne ye-di?
   Gizem what eat-PST
   'What did Gizem eat?'
A: #Gizem ye-di [dondurma].
   Gizem eat-PST ice cream
   'Gizem ate ice cream.'

All focus types are excluded postverbally. The following examples illustrate for corrective focus (41) and selective focus (42), respectively.

(41) A: Derin dondurma ye-miş.
   Derin ice cream eat-EV.3SG
   'Derin has apparently eaten ice cream.'
B: #Derin değil, ye-di [Gizem].
   Derin NEG eat-PST Gizem
   'It was not Derin who ate it, it was Gizem.'

(42) Q: Ali-yi mi sev-iyor-sun Ozan-ı mı?
   Ali-ACC Q love-PROG-2SG Ozan-ACC Q
   'Do you love Ali or Ozan?'
A: #Sev-iyor-um [Ali-yi].
   love-PROG-1SG Ali-ACC
   'I love Ali.'

The restriction which bans focus constituents from postverbal position immediately follows from the proposed theory. Since focus constituents always move across the verb to the specifier of the FocP, they can logically never occur postverbally.

To sum up, information structure influences Turkish word order with respect to the postverbal position. Given and informationally new constituents may be realized postverbally, focal constituents, on the other side, are not licensed in this position. The assumption of a structured left periphery for Turkish is able to account for the observed word order. Postverbal constituents may either stay in their in situ position (new information focus), or move to a discourse anaphoric phrase (DaP) if they are given. We assume that every Turkish sentence exhibits an information-structural split. Thus, the assumed base-structure is hypothetical in that it is always modified by pragmatically driven movement operations. This even carries over to all-new focus as discussed in (37) where we assumed that the verb moves to Foc and the TP to the SpecFocP.

With respect to the postverbal domain, Turkish clearly differs from German. The German ‘Nachfeld’ is not restricted to constituents with a specific IS-status: it may host given, new and focused constituents. We conclude that Turkish exhibits discourse configurational properties also with respect to all clausal domains whereas German is discourse configurational mainly in the preverbal domain.
5 Further evidence for the proposal

In this final section, we corroborate our account of the Turkish postverbal field by adding evidence from quantifier scope. Based on the assumption that quantifier scope is determined overtly in Turkish, we elaborate an assumption made in the previous section where we discussed multiple occurrences of the DaP. We show that the available scope interpretations strongly suggest a hierarchy in the postverbal field just as predicted by our theory. In addition, we come back to the non-application of the Right Roof Constraint in Turkish and show how our theory is able to derive this fact.

5.1 Presentational focus realization in Turkish

Quantifier Scope is determined in the surface position in Turkish (Kural 1993). Thus, the surface position of quantifiers determines the scope relations at LF. The data in (43) prove a strictly hierarchical syntax of the pre- and postverbal domain, which could not be achieved by extraposition.

(43) a. Herkes bazı gün-ler şikayet ed-er.
   everyone some day-PL complaint make-AOR
   ‘Everyone complains on some days.’
   (∀ > Adv) (*Adv > ∀)

      some day-PL everyone complaint make-AOR
      ‘Everyone complaints on some days.’
      (Adv>∀) (*∀> Adv)

Example (43a) can be paraphrased that for everyone, there are some days where one complains. (43b) means that there are some days, like a complaint day, when everyone complains. That is, in (43a) the subject universal quantifier scopes over the adverb whereas in (43b) the adverb scopes over the subject universal quantifier. Surface scope also holds if one of the quantifiers appear postverbally. In this case the preverbal quantifier scopes over the postverbal one.

(44) a. Herkes şikayet ed-er bazı gün-ler.
    everyone complaint make-AOR some day-PL
    ‘Everyone complains on some days.’
    (∀ > Adv) (*Adv > ∀)

      some day-PL complaint make-AOR everyone
      ‘Everyone complains on some days.’
      (Adv>∀) (*∀> Adv)

Given the assumption that quantifier scope is established in surface position in Turkish, it can be concluded that the preverbal constituent is structurally higher than the postverbal one. The third case we consider are examples where both quantifiers occur postverbally. As indicated by the readings below, the first of the postverbal quantifiers takes scope over the second. We assume the structure in (46) for example (45B2) where the adverbial takes scope over the universal quantifier. The order of the recursive DaPs with the adverbial being higher than the universal quantifier reflects the surface scope.

(45) A: İçeri giri-rken gör-dü-m, çalışan-lar şikayet
      Inside enter-while see-PST-1SG employee-PL complaint
      ed-iyor-lar-dı dışarida.
      make-PROG-3PL-PST outside
      ‘When getting into the building, I saw that some employees were complaining outside.’
‘That’s right, everyone complains here on some days.’

‘That’s right, everyone complains here on some days.’

Surface scope also determines the relative interpretation of two adverbials, independently of their position with respect to the verb. This is shown in the following example, taken from Wilson and Saygin (2001), who argue in support of Cinque’s (1999) hypothesis that different classes of adverbs enter a rigidly ordered sequence, which is the same across languages. According to Cinque (1999), adverbials expressing the speaker’s attitude must be higher in the tree than circumstantial adverbials, see (47a). The reversed order, (47b), is ungrammatical.

(47) a. Açıkçası muhtemelen gel-me-yeceğ-im. (Wilson and Saygin 2001:3)
   ‘Frankly, I will probably not come.’

b. *Muhtemelen açıkçası gel-me-yeceğ-im.
   possibly frankly come-NEG-FUT-1SG

Interesting for the topic of this paper is the observation that the same restriction also holds if the two adverbials occur postverbally. Again, the circumstantial adverb muhtelemen cannot precede the attitude adverbial açıkçası. This shows that there is a hierarchy in the Turkish postverbal domain as well, which follows under the analysis we propose in (46).

(48) a. Gel-me-yeceğ-im açıkçası muhtemelen.
   ‘Frankly, I will probably not come.’

b. *Gel-me-yeceğ-im muhtemelen açıkçası.
   possibly frankly come-NEG-FUT-1SG
5.2 The Right Roof Constraint

As was previously exemplified in section 2, example (13), the Right Roof Constraint can be violated in Turkish. Under the assumption that postverbal constituents are not derived by rightward movement, as often assumed for German, see Büring and Hartmann (1997), but by movement to information-structural projections to the left, this is expected. This is illustrated in (49), where a nominalized propositional complement is embedded under the matrix verb ‘bil-iyor-um’ (‘I know’). The subject of the embedded complement moves leftwards to the specifier of the DaP remnant movement of the embedded complement to FocP. In an extraposition approach, a violation of the Right Roof Constraint would be expected since movement of the subject would cross the boundary of a propositional constituent.

\[I book-ACC read-NOMLZ-POSS-ACC know-PRS-1SG Ahmet-GEN \]
‘I know that Ahmet read the book.’

\[(50)\]

\[
\begin{array}{c}
\text{TopP} \\
\text{Ben} \\
\text{FocP} \\
\text{DP₁} \\
\text{D'} \\
\text{Foc} \\
\text{bilyorum₂} \\
\text{DaP} \\
\text{NP} \\
\text{D} \\
\text{DP} \\
\text{TP} \\
\text{DP₁ kitabi okuduğunu} \\
\end{array}
\]

The subject of the possessive construction may also be fronted to SpecTop of the matrix clause, in which case topic movement follows focus movement, see (51).

\[(51)\] Ahmet-in ben kitab-ı oku-duğ-u-nu \[bil-i-yor-um.\] Ahmet-GEN₁ book-ACC read-NOMLZ-POSS-ACC know-PRS-1SG 'I know that Ahmet read the book.'

6 Summary

This paper develops a novel syntactic analysis of the postverbal domain in Turkish with a comparison to another SOV language, German, where relevant. Turkish uses different strategies to implement postverbal constituents with respect to German. Although these two languages are both SOV languages, and they both have access to postverbal domain, the use of this domain takes place for very different purposes. Categorically, Turkish licenses DPs in the postverbal position as well as other syntactic categories, but German does not license DPs in the postverbal domain. From the perspective of information structure, Turkish is sensitive to the information
structural status of the postverbal constituents. We have seen that Turkish bans focused constituents from the postverbal position as opposed to German, which does not impose any information structural restrictions on postverbal constituents. Namely, constituents carrying different information-structural status can appear in this position in German. According to the proposal developed in this paper, in Turkish high and low functional projections host constituents with different IS-status. The verb targets pre-defined positions according the informational split of the clause. The constituents having certain information structure then move to the specifier positions of the relevant heads.

References

SOV-X: Syntactic and pragmatic constraints of the postverbal domain in Turkish


Kural, Murat (1997) Postverbal Constituents in Turkish and the Linear Correspondence Axiom. Linguistic Inquiry 28(3).


Thiersch, Craig (1978) Topics in German Syntax. PhD Dissertation M.I.T.


A predicate-final constraint for head-final languages

Jaklin Kornfilt
kornfilt@syr.edu

1 Introduction

This paper proposes a constraint for head-final languages, a “Predicate-Final Constraint” (PFC), such that, even in word-order free head-final languages, predicates must be clause-final: clauses must strictly represent the head-final property of the language. (See also Bayer, Schmid and Bader (2005) and Bayer (2009), i.e. work on head-final languages, where a very similar proposal is made, and where it is claimed, based mainly on German and Bengali, that in head-final CPs, the right edge of the CP must be “visible” in some fashion. For a very early instance of such an idea, see Kuno 1972.) This constraint is parameterized, such that it is absolute for some head-final languages (e.g. Japanese—but this needs further study, since for many speakers, Japanese seems to be more like Turkish in colloquial styles), but limited to embedded clauses in others (e.g. Turkish). A further weakening of the constraint in some languages, e.g. Turkish, is that the constraint holds only when the clause-final predicate is followed by material in the higher clause.

The Turkish facts form the core of this short paper. The relevant observations are summarized and illustrated in section 2, and the generalization concerning the importance of the unencumbered right edge of (embedded) clauses, along with the weakening of the PFC with respect to the importance of potential matrix material following the right edge is discussed. Section 3 compares the relevant Turkish facts with corresponding German ones (and also mentions Bangla), and addresses the issue of the nature of the constraint. Section 4 concludes.

* This paper is based on my presentation at the Workshop on Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages, held in May 2018, in Schloss Freudental. The observations concerning Turkish, showing the necessity of a (parameterized) final predicate constraint, are also to be found in Kornfilt (2019); however, section 2 of this paper, while addressing these observations, is not identical to the corresponding sections of that longer work. More importantly, sections 3 and 4 of the present paper are completely new. I would like to thank Yvonne Vissel, for her perfect organization of the workshop at Schloss Freudental, as well as the audience at the workshop for helpful questions and comments. Thanks are due in particular to Josef Bayer, for his help with finding relevant literature and for discussion. I would further like to thank Kai von Fintel for discussion of the German examples and for offering judgments on those as well as on additional examples which he sent me later. I am also grateful to R. Amritavalli for her detailed comments on the first draft of this paper, which would have benefitted even more from those comments, had I been able to follow them more closely. Antonia Braun deserves thanks for her help in the formatting of the paper. The usual disclaimers apply.
2 Observations with respect to scrambling, yes/no questions, and identical predicate ellipsis

2.1 Overview of main facts

The predictions made by this constraint are borne out in Turkish. The PFC is illustrated in this paper for scrambling, as in (1 - 4) and (7 – 9), for Yes/No questions, as in (10 -14), and for coordinate structures with identical predicates, as in (15 - 25), showing that the ellipsis of a predicate under identity obeys this parameterized constraint, i.e. it holds strictly for embedded clauses, but not for root clauses in Turkish. This has consequences for the directionality of such ellipsis: While both forward and backward ellipsis are possible in Turkish root clauses, only backward ellipsis is allowed in embedded clauses. Additional facts in coordinate structures with predicate ellipsis are shown to follow from this constraint, as well. I shall further argue that, at least for Turkish, the embedded clause which must be predicate-final when it precedes the matrix verb does not undergo any type of “clause union” with the matrix; from this point of view, Turkish is like Bangla, rather than like German (whose infinitival clauses are studied in this context in Bayer 2009 and in Bayer et al. 2005). However, it is different from both German and Bangla, in that the constraint in question holds even when the embedded and the matrix verb are separated by matrix material; in other words, the verbs do not have to be adjacent for the PFC to hold.

2.2 Scrambling in Turkish, and similarities to and differences from German

Turkish is a head-final language; this means that the basic word order is SOV in both root and embedded clauses, as in (1). The language is quite word-order free, with the possibility of leftward movement for topicalization, and rightward movement for backgrounding, the latter leading to non-verb-final word orders. However, it is important to note that the scrambled constituent can’t be post-verbal, if root material follows, as in (3); such a constituent can scramble “long-distance”, after the root predicate, as long as the embedded predicate is clause-final, as in (2); but such a constituent can be post-verbal locally, if the entire embedded clause is scrambled to follow the root predicate, as in (4a); when a matrix constituent follows the scrambled embedded constituent, as in (4b), on the other hand, the utterance deteriorates markedly:

1 This doesn’t mean that Turkish doesn’t have the phenomenon of clause union. Some reduced (infinitival or even tensed, but agreement-less) argument clauses, which do show the effects of the generalization about predicate-finalness of clauses, may, in fact, be open to a clause union analysis, in which case there would be an additional reason why the predicate-final constraint would hold for such clauses which, of course, would have to be verb-final, and would show up left-adjacent to the matrix verb, with no intervener possible. The statement in the text only asserts that clause union would not explain all occurrences of the PFC, given that in general, embedded clauses in Turkish, despite the fact that most of them are nominalized, do not show any evidence of clause union with the matrix.
Interestingly, German exhibits very similar facts, where the embedded clause is infinitival (the predicates are boldfaced):

(5) a. Ich habe ihn [sich dafür zu entscheiden] aufgefordert
   I have him there-for to decide asked
   ‘I asked him to decide on it’

b.*Ich habe ihn [sich zu entscheiden] dafür aufgefordert (Bayer 2009)

c. [Ich habe ihn [sich zu entscheiden] aufgefordert] dafür;

These examples correspond to the Turkish examples in (1), (3), and (2), respectively. Similarities are found in examples corresponding to (4a), as well, i.e. the constraint does not apply, when the embedded clause is post-verbal itself (and thus no root material follows it):

(5) Ich habe ihn aufgefordert [sich zu entscheiden] dafür
   I have him asked REF to decide there-for (Bayer 2009)\(^3\)

\(^2\) The basic pattern for embedded clauses is based on clausal nominalization; these clauses are quite similar to gerundive clauses in English. For details, see Kornfilt (1997), Borsley & Kornfilt (2000), Kornfilt & Whitman (2011), among others. The two main types of the predicates in such nominalized clauses bear the suffixes \(--\text{DIK} \) and \(--\text{mA} \), which I have glossed ‘indicative nominalization’ and ‘subjunctive nominalization’, respectively, based on arguments discussed in Kornfilt (2003) and (2007); these suffixes are positioned in the morphological slot of the verb where, in fully finite verbs, TAM (i.e. tense, aspect, and mood) morphemes would appear. Predolac (2018) is another source for this view of nominalized Turkish verbs with respect to this mood distinction.

\(^3\) I am indebted to R. Amritavalli for the information that this “paradigm” of examples is found in Kannada (a Dravidian language), as well.
However, things are different in the pre-verbal field, when the embedded predicate and the matrix one are not adjacent, and when the intervener is a matrix (rather than embedded) constituent. In German, this constellation enables an embedded constituent to scramble to a position after the embedded predicate:

(6) Ich habe ihn [[sich e zu entscheiden] dafür] schon mehrmals aufgefordert
I have him REF to decide there-for already several times asked
‘I already asked him more than once to decide on it.’ (Bayer 2009; material intervening between the embedded and the root predicates is underlined.)

However, in Turkish, material that intervenes between the two predicates does not invalidate the predicate-final constraint in question:

Hasan-GEN finish-INDICNOM -3.SG –ACC application-ACC nobody-DAT
tell -NEG-PST-1.SG
Intended reading: ‘I didn’t tell anybody that Hasan finished the application.’

The same observations can be made with other types of embedded nominalizations, too; in (8), we have an example with a nominalized subjunctive clause, rather than a nominalized indicative clause, as we did in (7):

(8) * Hasan-a [[pro e bitir-me -sin -i] başvuru-yuğ]
Hasan-DAT finish-SUBJNCTNOM-3.SG–ACC application-ACC
defalarca söyle-di-m.
many times tell -PST-1.SG
Intended reading: ‘I told Hasan many times that he should finish the application.’

The PFC remains in effect even when there is more than one constituent separating the two verbs:

(9) * [[Hasan-ın e bitir-me -sin -i] başvuru-yuğ] herkes-e defalarca söyle-di-m.
Hasan-GEN finish-SUBJNCTNOM-3.SG–ACC application-ACC everybody-DAT
many times tell -PST-1.SG
Intended reading: ‘I told everybody many times that Hasan should finish the application.’

I am indebted to R. Amritavalli for the information that from the point of view of matrix material that intervenes between the embedded clause and the matrix predicate not invalidating the PFC, Kannada appears to be like Turkish, and not like German. She also shared interesting examples of such topicalized
Now that we have seen the PFC at work with respect to word order, and in particular with respect to scrambling/backgrounding, let us turn to Yes/No questions, where the same constraint can be shown to pay a role, as well.

2.3 Yes/No Questions

In root clauses, the Yes/No Q-marker can attach to the predicate, as in (10), as well as to other constituents, as in (11). However, it can attach only to non-predicates in embedded clauses constituents, as can be seen in (12), which is well formed, in contrast with (13), which is ill-formed, because the particle is attached to the embedded predicate:

(10) Hasan başvuru-yu bitir-di mi?
Hasan application-ACC finish-pst Q
’Did Hasan finish the application?’

(11) Hasan başvuru-yu mu bitir-di?
Hasan application-ACC Q finish-pst
’Did Hasan finish the APPLICATION?’ (i.e. ‘Was it the application that Hasan finished?’)

Hasan-GEN application-ACC Q finish-INDICNOM-3.SG-ACC ask- PST-1.SG
’I asked whether Hasan finished the APPLICATION.’

Hasan-GEN application-ACC finish-DIK -3.SG-ACC Q ask- PST-1.SG
Intended reading: ‘I asked whether Hasan finished the application.’

(Note: (13) is OK under a wide-scope, i.e. root-level, Y/N question interpretation (‘Did I ask whether H. finished the application?’), whereby the embedded clause would be the questioned constituent at the level of the root clause, similar to başvuru-yu ‘application-ACC’ in (11); under that interpretation, the Q-particle would be a root-clause element, and the embedded clause would be predicate-final.5)

embedded clauses out of which sub-constituents are backgrounded to a position after the matrix verb (which is possible in Turkish, too). If a matrix constituent is also backgrounded after the matrix verb, it must precede the backgrounded embedded (post matrix-verb) constituent in Kannada—something which she ascribes to an obligation for “right-displaced material” to be “integrated hierarchically into the matrix and the embedded clauses” (p.c.). However, I would tentatively claim that this is due to the tendency, also seen in Turkish, for backgrounded sub-constituents of an embedded clause to “scramble” to the very end of the utterance. In corresponding Turkish backgrounded sub-constituents, the order with a following backgrounded matrix constituent is borderline, if not totally out (see example (4b)), probably due to the fact that in the Turkish matrix post-verbal field, (almost) “anything goes”; cf. Kornfilt 2005. For details concerning “local” and “long-distance” scrambling (also to post-verbal positions) in Dravidian, see Jayaseelan and Amritavalli (2005).

5 I am grateful to R. Amritavalli for the information that in Dravidian, Yes/No questions are somewhat similar, and that Yes/No questions can also be used to focalize a particular constituent, like in Turkish. In Dravidian, such constructions are clefts, or cleft-like. She implies that (13), under its wide-scope
To obtain the reading of a regular embedded Y/N-question, a coordinate predicate consisting of an affirmative and a negative part (similar to the “A-not-A” questions in Chinese—cf. Huang 1982, among others) has to be used, without the Y/N question particle:

(14) \[
\text{Hasan-in başvuru -yu bitir -ip bitir-me -diğ -in -i]}
\]
Hasan-GEN application-ACC finish-VBLCONJ finish-NEG-INDICNOM -3.SG -ACC
sor -du -m.
ask -PST-1.SG
'I asked whether Hasan finished the application (or not).'

Note that here, the embedded clause is predicate-final.

2.4 Identical predicate ellipsis in coordinate structures

Both forward (15) and backward (16) ellipsis of identical predicates is well-formed in root clauses, but only backward ellipsis is well-formed in embedded clauses (17 versus 18, 20 versus 22, 21 versus 23) in their canonical pre-verbal position, where the embedded clause is followed by root material (the gapping site is marked with an underlining, and by parentheses in the English translations):

(15) Hasan kitab -i oku -du, Fatma da gazete -yi ___.
Hasan book -ACC read -PST Fatma and newspaper-ACC
'Hasan read the book, and Fatma (read) the newspaper.'
(16) Hasan kitab -ı ___, Fatma te gazete -yi oku -du.
Hasan book -ACC Fatma and newspaper-ACC read -PST
‘Hasan (read) the book, and Fatma read the newspaper.’

(17) Zeynep [Hasan –ın kitab -ı ___, Fatma –nin da gazete -yi
Zeynep Hasan -GEN book -ACC Fatma -GEN and newspaper -ACC
read -INDICNOM -3.SG -ACC hear -PST
‘Zeynep heard that Hasan (read) the book, and Fatma read the newspaper.’

Zeynep Hasan -GEN book -ACC read -INDICNOM -3.SG -ACC Fatma -GEN and
gazete -yi ___] duy –du.
newspaper -ACC hear -PST

Intended: ‘Zeynep heard that Hasan read the book, and Fatma (read) the newspaper.’

This root/embedded asymmetry is easily explained by the PFC, which is violated in (18) and all
the other ill-formed examples just listed, but not in (17), nor in the other well-formed examples
listed above. Note that, once again, I am assuming the weakened form of the PFC here, where
the constraint holds only when the predicate which is supposed to mark the right edge of the
clause is followed by matrix material; in this weakened version, the PFC is not violated in
instances of forward gapping in root sentences, such as in (15): even though the entire
coordination is not predicate-final in such instances, the PFC is not violated, given that no
constituents belonging to a higher clause follow the site of the gapping, and the utterance is
well-formed.

Just like with post-verbal scrambling, the PFC can be violated, when the embedded clause is
post-verbal itself; forward predicate ellipsis in the embedded coordination becomes well-formed:

(19) Zeynep e, duy-du [Hasan –ın kitab -ı oku-duğ –un –u, Fatma –nin de
Fatma –nin de gazete -yi ___].
Fatma -GEN and newspaper -ACC
‘Zeynep heard that Hasan read the book, and Fatma (read) the newspaper.’

Just like with (15), the gapping site is not followed by any constituent belonging to the higher
clause; thus, even though here, the embedded clause is not predicate-final (given that it has no
final predicate, due to gapping), the weakened version of the PFC is not violated, and the
utterance is well-formed, in contrast with the ill-formed (18), its pre-verbal counterpart.

These contrasts are independent from the nominalized character of the typical embedded
clauses in Turkish; non-nominalized embedded clauses exhibit the identical contrasts:

(20) Ahmet [Hasan kitab -ı ___, Fatma da gazete -yi oku -du]
Ahmet Hasan book -ACC Fatma and newspaper -ACC read -PST
san -iyor.
believe -PRPROG
‘Ahmet believes Hasan (read) the book and Fatma read the newspaper.’
Non-nominalized embedded clauses cannot be scrambled to post-verbal positions\(^6\); however, they can be base-generated as head-initial CPs (but with the lower phrasal architecture head-final) in such positions, a pattern copied from Persian, along with a complementizer; in that pattern, once again, just like in (19) with a nominalized coordination, forward gapping is possible, due to the post-verbal position of the embedded clause, given that it isn't followed by root material:

(24) Ahmet san -iyor [ki [Hasan kitab-ı -ı oku -du, Fatma da gazete -yi ___]]

Ahmet believe-PRPROG that Hasan book-ACC read-PST Fatma and newspaper-ACC

‘Ahmet believes Hasan read the book and Fatma (read) the newspaper.’

(25) Ahmet isti -yor [ki [Hasan kitab-ı oku -sun, Fatma da gazete -yi ___]]

Ahmet want -PRPROG that Hasan book-ACC read-SUBJUNCT Fatma and newspaper-ACC

‘Ahmet wants (that) Hasan should read the book and Fatma (should read) the newspaper.’

Clearly, the (parameterized) PFC can easily and successfully deal with all the contrasts illustrated so far.

---

\(^6\) This prohibition is getting weakened in recent colloquial usage, and in such utterances, the result of forward gapping is well-formed, as predicted in the (weakened) PFC-based approach I am proposing here on gapping:

(i) Ahmet çok isti-yor [Hasan kitab-ı oku -sun, Mehmet te gazete -yi ___]

Ahmet very want-PRPROG Hasan book-ACC read-SUBJUNCT Mehmet and newspaper-ACC

‘Ahmet wants very much Hasan should read the book and Mehmet (should read) the newspaper.’
3 The nature of the constraint

3.1 Some like-minded prior proposals

In looking for an explanation of the facts in German and Bangla, Bayer (2009) refers to Escribano’s (2004) Head-Final Filter (HFF): Base-generated pre-head modifiers must be head-final. But clearly, the constraint which is needed to explain the facts we have seen so far must hold not just for modifiers, i.e. for adjuncts, but also for complements, as in the examples addressed here; note that all the embedded clauses in the Turkish examples are arguments of the root predicates, i.e. they are complements. For those, Bayer (2009) refers to the Final-over-Final Constraint (FOFC) of Biberauer, Holmberg & Roberts (2008), while noting that it is not sufficient to explain all relevant examples.

One issue with the FOFC is that its definition doesn’t make it clear that it is relevant to the type of examples addressed here—at least not to the Turkish data; e.g. in Biberauer and Roberts (2013), the relevant part of the FOFC is defined as follows:

(26) “... If α is a head-final phrase, and β is a phrase immediately dominating α within the same Extended Projection, then β can be head-initial or head-final.” (cf. Biberauer & Roberts 2013:16.)

Biberauer & Sheehan’s (2013) definition of the FOFC is essentially the same. In the examples we have seen from Turkish, the issue was not whether β is head-initial or head-final, but rather whether α, while being head-final, is allowed to have an overt (i.e. filled with phonological material) head or not.

However, the reference to the constraint in Biberauer, Newton & Sheehan (2009) points to an effect of the FOFC which makes it more relevant to what this paper is interested in:

(27) “As an absolute principle, FOFC rules out the possibility of a head-final phrase dominating a categorically non-distinct head-initial phrase.” (Biberauer, Newton & Sheehan 2009: 701.)

If we change this statement’s end to “… a categorically non-distinct non-head-final phrase” (i.e. expanding the statement, by replacing “head-initial phrase” with the more general “non-head-final phrase”), we have a statement which is essentially equivalent to the PFC proposed here (at least from an empirical point of view), and to its precursor, a filter proposed in George & Kornfilt (1980):

(28) *... X\S ...V (where V c-commands S), unless X = V. (Please note that, in the spirit of the 1970’ies and early 1980’ies, “S” here stands for a clause=sentence, and should thus be updated to CP or TP.)

All of these principles or constraints are supposed to hold for modifiers as well as for complements. We cannot stop here, however, without asking what the motivation for any one of those principles or constraints (which are ultimately very similar to each other) might be.
An explanation for why head-final languages should have a filter or constraint of this sort is proposed in Bayer (2009):

(29) The FOFC-violating constellation \([\beta \bar{P} [\alpha \bar{P} \gamma P] \beta]\) is ruled out because \(\alpha \bar{P}\) and \(\beta\) fail to be in an agreement (/feature valuation) relation.

Notice that not only the constellation in (29), but also the following one should block the agreement (/feature valuation) relation in question:

(30) \(\beta \bar{P} [\alpha \bar{P} \gamma P] \delta P \beta\)

It seems that in Turkish, just as in German, the \(\gamma P\) in (29) would indeed block agreement/feature valuation, if this is the correct explanation of the relevant constraint. But while the same phrase, i.e. \(\gamma P\), would not block valuation in (30) for German, as we saw in the examples with the complement clause “scrambled” or topicalized in the pre-verbal field, by moving leftward (this lack of blockage would be due to the \(\delta P\) which somehow renders the condition inoperative), \(\gamma P\) would still block valuation for Turkish, even when a matrix \(\delta P\) is present, as we also saw: the leftward scrambling of the embedded clause does not free up that clause from the necessity of being predicate-final in Turkish, in contrast to German (and Bangla).

To phrase this differently, in Turkish, for the constraint in (29) to hold, the embedded verb and the matrix verb do not have to be adjacent. I shall return to this issue later on, especially with respect to German.

The relevant feature which must undergo valuation under Bayer’s approach is referred to in Bayer (2009) and in Bayer, Schmid & Bader (2005) as “status” (following Bech 1955/57), whereby different types of matrix verbs access different types of “edges”:

(31) “In the pertinent structures of head-final languages HP has only a right edge which consists of the head H…” (Bayer 2009:8). This is important with respect to Chomsky’s (2001) Phase Impenetrability Condition (PIC).  

For the complement to undergo valuation/agreement, its edge must be “visible” at spell-out. This is where the various versions of the filter/condition we have seen come in.

It appears that this visibility must be guaranteed not only at spell-out, but also at PF, as also acknowledged in Bayer (2009):

“However, at spell-out only the edge of XP is accessible. Unlike core-syntactic movement, PF-movement has a linear effect. It alters the right edge of, say, AP.” Bayer 2009:7.)

As a consequence, the boldfaced (sub-)structures in (32) (corresponding to example (26) in Bayer 2009) are inaccessible for the agreement process.

---

7 The Phase Impenetrability Condition (PIC): The domain of head H is not accessible to operations outside HP, but only to H and its edge (the edge being SpecHP or positions adjoined to HP). (Cf. Chomsky, 2001:13)
Bayer states that (32a) (his (26a)) can be used for agreement / feature valuation (with respect to the head adjective) while (32b) (his (26b)) cannot be used in this way. This is because the rightward movement ("extraposition" in Bayer 2009) of the adverb in (32a), resulting in (32b) is characterized as a PF-movement, which alters the linearity of the right edge of the phrase it adjoins to (here, an Adjective Phrase).

Can we apply this proposal to the Turkish data, and the generalizations we drew from them? I shall argue at the end of the next subsection that an agreement-based explanation is not available for Turkish, given that it doesn’t have concord, which is the relevant type of agreement for the kind of German data illustrated in (32).

3.2 How important is adjacency? Is predicate ellipsis a phenomenon? It is in Special Predicate Ellipsis, but not in regular Predicate Ellipsis

In Turkish, too, linearity at PF plays a role. In what follows, I shall address a particular kind of predicate ellipsis process in Turkish coordinate structures which, I believe, shows quite clearly that at least in such instances, (special) predicate ellipsis is best analyzed as a PF-process. This special type of ellipsis, which I call Special Coordinate Ellipsis (SCE), involves predicates of clauses which are not directly coordinated. The following pair of examples illustrates this phenomenon:


‘Ali (heard) that Oya (cooked) the shrimp and Zeynep heard that Fatma cooked the lobster.’

(The elided verbs are signalled by underlinings in the examples, and by parentheses in the English translation, here and in the next example.)

8 Although, as we see in (i) versus (ii), matrix "interveners" (as in (ii)) do not "count" for the PFC proposed (as mentioned in the text earlier), while embedded "interveners" (as in (i)) do, because only the latter obliterate the edge of the embedded clause as the head of that clause:


Intended reading: ‘I didn’t tell anybody that Hasan finished the application.’


‘I didn’t tell anybody that Hasan finished the application.’
In these examples, the embedded predicates are not in an embedded coordination, but rather are part of clauses which are embedded within matrix clauses which themselves form a coordinate structure. In other words, the coordination is at the matrix level. Therefore, the elided embedded predicates can’t have undergone coordinate ellipsis by themselves, at least not in the usual sense of coordinate ellipsis. Note also that the English translation of (34) is well-formed only if the only elided predicate is the matrix predicate. The utterance becomes ill-formed if the embedded predicate is elided, as well. In Turkish, however, the resulting utterance with both the matrix and the embedded predicates elided is perfectly fine.

The SCE observed here in both directions must be a PF-operation rather than a strictly syntactic one: For the elision of predicates that are not in a direct coordinate structure containing the identical predicate that licenses the ellipsis, string adjacency of the embedded predicate with the elided matrix predicate is crucial.

Note that in our two examples, the surviving embedded predicate is string-adjacent to the matrix predicate. It therefore makes sense to assume that the identical, elided embedded predicate is (or was, before the elision) also string-adjacent to the elided matrix predicate. (Note that the elided matrix predicate is in a regular coordinate structure with the identical matrix verb that licenses the ellipsis.)

We can see the importance of string adjacency for these observations by altering the word order via scrambling:

\begin{align*}
\text{(35)} & \quad \boxed{\text{`ALI (heard) that Oya (cooked) the shrimp and ZEYNEP heard that Fatma (cooked) the lobster.'}} \\
& \quad \boxed{\text{`ALI (heard) that Oya (cooked) the shrimp and ZEYNEP (heard) that Fatma (cooked) the lobster.'}} \\
\end{align*}

These examples show clearly that without the string adjacency with the elided matrix predicate, ellipsis of the embedded predicate is ill-formed in these examples. Since string adjacency is a typical PF-phenomenon but is not a condition on syntax in the strict sense (cf. Chomsky 1995), I conclude that this type of SCE is a PF phenomenon.
Does this mean that other types of coordinate ellipsis are PF-phenomena, as well? I leave the answer to this question to future research; note, however, that in regular coordinate ellipsis, adjacency plays no role.

3.3 Back to regular coordination and the PFC in Turkish; additional facts in German

This brings us back to the issue of the string adjacency of the embedded and the matrix verbs with respect to the visibility of the right edge of the embedded clause: In the German (and Bangla) examples, that type of adjacency was important: once the embedded clause in those examples was not adjacent to the matrix verb (and thus the two verbs were not adjacent any longer), due to either the clause’s being extraposed to post-verbal position, or its being scrambled leftward, the edge of that clause could be “obliterated”, by the placement of a clause-internal constituent after the embedded predicate. In Turkish, on the other hand, we saw that this is not possible, and that the PFC, however formalized, holds independently from the adjacency of the two verbs. To be more exact, this contrast between German (and Bangla) on the one hand, and Turkish on the other, is true only for the instances of scrambling/topicalization; in the post-verbal field, in contrast with the pre-verbal field, all three languages behave the same way: the embedded clause does not have to be predicate-final, when it is placed after the matrix verb.

Why should there be a difference among these languages in the pre-verbal field, and not in the post-verbal area? And, linked to this question, are the Turkish facts with respect to the pre-verbal field more surprising, or the German (and Bangla) facts with respect to that same structural area?

With respect to the first question, the informal, “functionalistic” weakening of the PFC which I have referred to earlier has an answer: Suppose that not only Turkish, but also German and Bangla are subject to this weakened version of the PFC. In other words, we say that in (certain) head-final languages, the clause must be predicate-final, but only if clause-external material follows it. Functionally speaking, the clause must clearly show that it is head-final, i.e. predicate-final. However, this is obligatory only when the clause-final head, which is, at the same time, the right edge, i.e. the boundary of the clause, must resist any potential confusion with strings that follow it—strings which are clause-external, but might be perceived as being clause-internal. When the clause is not followed by any clause-external material, it is less important that the clause’s head fulfill its function as the rightmost boundary of the clause, and thus it can be followed by clause-internal material (as in rightward scrambling), as well as being absent with respect to other phenomena obliterating the linear status of the predicate as the rightmost element of the clause, such as predicate ellipsis.

This would take care of those instances in all three languages where the embedded clause has been extraposed to a post-verbal position, and where the embedded clause is not strictly predicate-final.

Now we can turn to those instances where Turkish differs from German and Bangla, i.e. where, in Turkish, the PFC holds even when the embedded clause has been scrambled/topicalized, but where in German and Bangla, it apparently does not hold.

From the point of view of the weakened version of the PFC, the Turkish facts are not surprising: Under leftward movement of the embedded clause, we obviously find matrix material that follows that clause: This is the constellation in (30). Thus, we would expect that the PFC should hold, and it does. Seen from this perspective, German and Bangla are the surprising languages.
In these two languages, as we saw, the embedded verb and the matrix verb have to be adjacent for the PFC to hold. But it is not clear why this should be so; as Bayer (2009) and Bayer, Schmid and Bader (2005) have shown, clause union is irrelevant (and would not hold for any of the relevant Bangla examples, nor for many of the relevant German examples).

In addition, it is not crystal-clear why leftward-scrambled/topicalized embedded clauses should not exhibit any freezing effects with respect to (dis-)allowing subconstituents that are extracted from them. As a matter of fact, they do, at least in German, and at least with respect to wh-movement; thus, Müller (1997) offers the following examples, whereby (37a) is ill-formed, due to the attempt at extracting a subconstituent from the scrambled object, while the same extraction is fine in (37b), due to the fact that the object has not moved from its base position:

(37) a. *[PP Worüber ] hat [NP ein Buch t1] keiner t1 gelesen?
   about-what has a book no-one read

   ‘About what has nobody read a book?’ (Müller 1997:11)

Bayer (1999) has similar examples. Leaving Bangla aside, this makes the well-formedness of German examples such as (6) surprising—although, according to Bayer (p.c.), only topicalization (i.e. scrambling to the highest possible topic position) induces freezing effects with respect to subextraction, and scrambling to intermediate positions (i.e. to positions in the “middle field”) does not. Thus, it would be interesting to see if examples such as the following one is considerably worse than (6):

(6)’ [[Sich dafür zu entscheiden] dafür], habe ich ihm schon mehrmals
   REF to decide there-for have I him already several-times
   t1 aufgefordert
   asked

   ‘I already asked him more than once to decide on it.’

I leave this example without a grammaticality judgment (but see below), given that I was unable to ask native speakers about the contrast between it and (6); I don’t find it worse than (6), but then, I don’t like (6), either.

With respect to the status of (6), I am not the only speaker of German who does not like such examples. I am indebted to Kai von Fintel for the information that he, too, does not like them—as a matter of fact, he stars both (6) and (6)’; and finds them ill-formed to the same extent. If it is true that intermediate scrambling positions do not exhibit freezing effects in German with respect to subextraction, the reason for the ill-formedness of (6) for speakers such as von Fintel may, indeed, be something like the PFC being operative in German in a similar way that it is in

---

9 The Turkish examples in this paper would not be amenable to a clause union analysis, either. As mentioned earlier, such an analysis would potentially be an option for infinitival embedded clauses, which never have predicate-subject agreement morphology, or tensed clauses without such agreement (such as those found in ECM (=Exceptional Case Marking)/SOR (=Subject to Object Raising) contexts), while none of the examples used in this paper have infinitival embedded clauses or otherwise agreement-less clauses and are thus not candidates for clause union.
Turkish. However, before deciding that this is so (and I will tentatively propose that it is), I should report that von Fintel also rejects examples such as (5), which I repeat here for the convenience of the reader:

(5) Ich habe ihn **aufgefordert** [[sich dafür **zu entscheiden**] dafür]
    I have him asked       REF to decide    there-for
    ‘I have asked him to decide for it (i.e. to decide in its favor).’

At first glance, this judgment poses a problem to the weakened form of the PFC as proposed here: the embedded clause is post-verbal and is not followed by any root material; thus, the PFC shouldn’t have to be applied in a strict fashion, and a constituent of such a clause should be able to scramble out.

However, the ill-formedness for von Fintel of this example is not problematic, as it turns out. Von Fintel reports that his judgment of ill-formedness has to do with the “lightness” of the backgrounded, rightward-moving word dafür ‘for it’, and its pronoun-like nature. When such backgrounded constituents are made heavy, von Fintel accepts them:

(38) Ich habe ihn **aufgefordert** [[t, aufzuhören] mit dem Zappeln]],
    I have him asked to-stop with the fidgeting
    ‘I have asked him to stop with the fidgeting’

This is completely in line with the facts we saw for Turkish and Bangla, as well as with Bayer’s judgment of (5) as well-formed: When the embedded clause is extraposed to a post-verbal position in the matrix, and is not followed by any matrix material, a subconstituent of the embedded clause can be placed after that clause’s predicate; the clause does not have to be verb-final in this position.

Also, von Fintel’s evaluations of the following two examples\(^\text{10}\) are crucial:

(39) ???/* Ich habe ihn [[t, aufzuhören] mit dem Zappeln]], **schon mehrmals** t, aufgefordert.
    I have him to-stop with the fidgeting already repeatedly asked

Intended: ‘I have repeatedly asked him already to stop with the fidgeting.’

    I have him to-stop with the fidgeting asked

Intended: ‘I have asked him to stop with the fidgeting.’

While (40) is ill-formed, just as it would be in Turkish and Bangla, and just as it would also be evaluated as ill-formed by German speakers such as Bayer, interestingly, (39) is also ill-formed

---

\(^{10}\) All the examples having to do with fidgeting were volunteered, along with their judgments, by Kai von Fintel, to whom I am very grateful for his help.
for German speakers such as von Fintel (and myself), just as it would in Turkish. The weakened version of the PFC would predict these evaluations for both examples. This version of the PFC does obviously not predict Bayer’s evaluation of (5) as well-formed, nor does it predict the well-formedness of the Bangla equivalent of such an example (and, I would surmise, the Bangla equivalent of (39)). Hence, an approach such as Bayer’s would need to be used, where any version of the PFC would have to be limited to instances where the embedded and the matrix predicates are adjacent, instead of the weakened version I have proposed here (which, however, is independent from adjacency of the verbs), based on Turkish data. It is interesting that von Fintel’s judgments align with mine, given that his judgments cannot possibly be influenced by Turkish, as mine might be.

What about the FOFC? As mentioned earlier, in the text following the statement of the FOFC in (27), it would have to be expanded somewhat, so as to rule out non-head-final projections under head-final ones, rather than prohibiting head-initial projections under head-final ones, so as to be relevant for many of the data addressed here. Once this is done, the FOFC essentially expresses the same idea as the PFC proposed here.11

A question to be posed at this point is about the reason for the existence of either one of these constraints. This is what Bayer (2009) and Bayer, Schmid and Bader (2005) have attempted to do. As stated in Bayer in (2009), an account of agreement would be needed, according to which an XP must have a right edge that enables it to check agreement in the specifier of the selecting head. In other words, one would need an account of agreement as concord. The data, mainly from German and Dutch, which are used to support this idea, exhibit morphological concord in right edge positions. This is quite convincing; however, for a language such as Turkish, where there is no (overt) morphological concord with respect to any conceivable feature (i.e. person, number, gender or case), it is not clear whether this agreement-based explanation for any principle or constraint that imposes right-edge “visibility” in head-final languages is sufficiently motivated.

4 Conclusions

This paper has proposed a constraint for head-final languages, imposing predicate-final word order on clauses. This predicate-final constraint (PFC) is parameterized across head-final languages, so that it may hold in absolute terms (e.g. possibly in Japanese) or in a weakened form (e.g. in Turkish), the latter only when the embedded clause is followed by matrix material. The weakened form of the constraint has the consequence that only embedded clauses will obey it, and that they would do so only when they are not followed by a matrix constituent, which is the case when the embedded clause is in a post-verbal position (i.e. follows the matrix verb). This consequence is found for Turkish, as well as for German and Bangla. However, German (at least for certain speakers) and Bangla allow for instances where the embedded clause is scrambled in the pre-verbal domain of the matrix, and where the constraint does not hold, thus allowing for non-predicate-final orders in the embedded clause in pre-verbal

---

11 As we saw in (27), the FOFC is formulated in such a way as to require identical categories for the higher and the lower projection in a constellation where the higher projection is head-final and the lower one cannot be non-head final. In the relevant German and Bangla examples, these projections were CPs in the examples we have seen or have at least alluded to. However, in most (but not all) Turkish examples, the clauses embedded under a CP were nominalized; it would make sense to analyze those as DPs dominating CPs (or as DPs dominating at least Asp(ect)Ps; see Borsley & Kornfilt 2000 and Kornfilt & Whitman 2011 for discussion). Thus, it is less than clear that with respect to such gerund-like clauses, there is categorial identity between the embedded and matrix clauses—a point which would challenge the relevance of even the expanded version of the FOFC for the data addressed in this paper.
positions, as well. We saw that for some other German speakers, on the other hand, this latter constellation is not possible, just as it is not for Turkish, and that the embedded clause must be predicate-final in scrambled positions, too. Whether the PFC can be derived from the FOFC is an open question, although some doubts were raised, given that the categories of the higher, head-final clause and of the lower clause whose head-final status is addressed by the PFC do not necessarily match in Turkish. Furthermore, at least for Turkish, it is not clear whether a concord-based explanation for the existence of either one of these constraint is convincing, given that the language does not show any evidence of (overt) concord. Instead, an informal, cognition- or processing-based explanation was proposed, such that in (some) head-final languages, the end, i.e. the right edge of an embedded clause must be cognitively prominent when it is not utterance-final, i.e. when the boundary between clause-internal and clause-external constituents needs to be made clear.

Abbreviations

1.SG  First person singular  
3.SG  Third person singular  
ACC  Accusative  
DAT  Dative  
GEN  Genitive  
INDICNOM  Indicative nominalization  
MASC  Masculine  
NEG  Negative  
PRPROG  Present progressive  
PST  Past  
Q  (Yes/No) Question marker  
REF  Reflexive  
SUBJNCTNOM  Subjunctive nominalization  
SUBJUNCT  Subjunctive  
VBLCONJ  Verbal conjunction

References

Bayer, J. 2009. “The Final-over-Final Constraint meets the Head-Final Filter”; handout of a talk presented at the conference on theoretical approaches to disharmonic word orders, Newcastle University.
A predicate-final constraint for head-final languages


Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive

Katalin É. Kiss
e.kiss.katalin@nytud.mta.hu

1 Introduction

This paper argues that the Ob-Ugric sentence displays a partial fusion of grammatical functions and discourse roles, and this has consequences for the licensing conditions of cases, agreement, and passive. In Ob-Ugric, the subject is also primary topic. The conflation of the subject and topic roles is attained by means of a liberal version of passive, involving the demotion of the [-topic] highest argument (whether it is an agent or a patient), and the promotion of a [+topic] internal argument (if there is one). Hence internal argument promotion, i.e., NP movement, is not case-driven; it can also target arguments bearing a semantic, adverbial case. The object is either information focus, in which case it remains in the VP, eliciting no verbal agreement and bearing no morphological case, or it is a VP-external secondary topic, in which case it elicits object–verb agreement and, in some of the dialects, it bears accusative case. In sentences with three arguments, the argument with the least prominent thematic role can only be secondary topic if it is promoted to object, and it can be primary topic if it is further promoted to subject. Another peculiarity of Ob-Ugric grammar is the possibility of oblique case for subjects functioning as shifted topics in active sentences. It is argued that these facts require the modification of case theory. So as to account for Ob-Ugric differential object marking, we either have to divorce NP licensing and morphological case marking, or we have to give up the Case Filter. The distribution of Ob-Ugric locative subjects could best be explained in a Fillmore (1968) style framework where structural case neutralizes an underlying semantic case.

The paper is structured as follows: Section 2 describes the relevant facts of Ob-Ugric. Section 3 provides a structural analysis of the facts surveyed. Section 4 discusses the theoretical implications of the proposed analysis. Section 5 is a summary.

2 The Ob-Ugric sentence

2.1 The subject-topic

The Ob-Ugric languages of the Uralic family, Khanty and Mansi, are structurally closely related SOV languages (at present subjected to strong Russian influence). The Ob-Ugric subject bears a morphologically unmarked nominative case, and elicits verbal agreement. It is external to vP – as shown by the fact that it precedes not only the VP-internal material but also vP-adjuncts. The focal object is caseless, triggering no agreement. It is left-adjacent to the verb.
Passivization is also triggered in presentative sentences containing a single non-specific complement. The verb assumes passive morphology, and the theme bears locative or lative case.

---

1. Khanty and Mansi have no generally accepted spelling conventions, therefore, I preserve the spellings of the examples cited from various sources.

2. This constraint is becoming less strict in the „Russified” language variety of the younger generations.

3. In (4a), the verb agrees with the definite object. sg<3sg in the glosses means that the verb bears an object agreement marker cross-referencing a singular object, and a subject agreement marker cross-referencing a 3rd person singular subject.
The subject position appears to be empty, but it may be filled by an implicit personal pronoun or spatiotemporal pradverb. The verb bears default 3SG agreement.4

(6) a. puwlapsi-na e:t-s-a. (Khanty; Nikolaeva 1999: 32)
   tumor-LOC enter-PAST-PASS.3SG
   ‘A tumor appeared.’ Lit.: ‘(pro) was entered by a tumor.’

b. āk-māst komne tāwl-aw-s (Mansi; Kulonen 1989: 184)
   suddenly man-LAT appear-PASS-PAST.3SG
   ‘Suddenly a man appeared.’ Lit.: ‘Suddenly, (there) was appeared by a man.’

The promotion of the oblique arguments to subject in (5) is not motivated by their lack of case; on the contrary, they lose their lexically selected oblique cases in the course of movement to subject position. The trigger of passivization in all cases is the [-topic] feature of the underlying subject, which leads to its demotion to adjunct status. The promotion of the internal argument to subject must be due to the EPP.

2.2 Focus objects versus topic objects

In active transitive sentences, the object is either in situ, bearing no morphological case and eliciting no verbal agreement, interpreted as information focus – see (1) above, or it is in a second topic position, following the subject-topic and preceding vP-adjuncts – see (7). (The morpheme preceding the symbol < in the glosses is an object agreement suffix, and the morpheme following the symbol < is a subject agreement suffix, occasionally fused with the object agreement suffix.) The topicalized object gets accusative case and elicits number (singular, dual, or plural) agreement on the verb (see Nikolaeva (2001), and Sosa (2017) on Khanty, and Skribnik (2001), Virtanen (2014; 2015), Sipőcz (2015), and Bíró and Sipőcz (2017) on Mansi).5 Accusative case is morphologically salient on lexical noun phrases only in the eastern dialect of Mansi. In other dialects, it is only visible on personal pronouns – but tentatively I assume a phonologically null accusative morpheme on topicalized lexical noun phrases, as well:

(7) Petra mo:jpar-Ø/luw-e:l u:r-na wa:nt-sa-lii (Khanty; Nikolaeva 2001: 28)
    Peter bear-ACC itl-ACC forest-LOC see-PAST-SG<3SG
    ‘Peter saw the bear in the forest.’

The claim that the Ob-Ugric object elicits verbal agreement if and only if it is a topic has been supported by various kinds of evidence. Nikolaeva (2001) analyzed the discourse role of the object

4 Kiparsky (2013) claims that Khanty marginally also allows transitive impersonal passives, on the basis of example (i) cited from Kulonen (1989):

(i) nōgat ile onalta-s-i (Khanty; Kulonen 1989: 267)
   you-ACC PRT teach-PAST-PASS.3SG
   you were taught

This construction is known from Northern Russian; the isolated example in (i) must be a loan translation – especially because it occurs in a fairy tale with a well-known Russian original (The fisherman an the gold fish).

I also disagree with Kiparsky’s analysis of (ii). He claims (ii) to be a subjectless impersonal sentence:

(ii) ima-ne χāt  χara pōwta t’ōylat-aj (Khanty, Kulonen 1989: 269)
    woman-LOC house floor-LAT blow-INF start-PASS.3SG
    ‘the woman started to blow onto the floor’

In my view, (ii) is a regular passive construction with the infinitival phrase ‘to blow onto the floor’ functioning as the subject of the passive verb.

5 Accusative marking has been found to be restricted to specific or definite objects in several other languages, as well, among them Turkish (Enc 1991), and Hindi (Bhatt and Anagnostopoulou 1996).
in nearly 1100 Khanty transitive clauses recorded by Pápay (1906–8). The object triggers agreement in 677 of the clauses. 87% of the agreeing objects are contextually given and an additional 7% have a contextually given possessor, i.e., altogether, 94% of the agreeing objects are familiarity topics. Among the 412 non-agreeing, i.e., focused, objects, the proportion of previously activated objects is 11%. The role of givenness in licensing object–verb agreement has also been proven by elicited examples. Observe the following minimal pair of Nikolaeva (2001):

(8) a. What did you do? (Khanty; Nikolaeva 2001: 16)
    *Ma tam kalaŋ ʷeːl-s-əm /weːl-eːm
    I this reindeer kill-PAST-1SG.kill-PAST-3SG<1SG
    'I killed this reindeer.'

b. What did you do with this reindeer?
    *Ma tam kalaŋ ʷeːl-s-əm /weːl-eːm
    I this reindeer kill-PAST-1SG.kill-PAST-3SG<1SG
    'I killed this reindeer.'

In (8a), the object conveys new information, hence it is VP-internal, eliciting no agreement. In (8b), on the contrary, the object is given, hence it is in topic position eliciting agreement.

The Eastern Mansi data in (9a,b) are cited by Virtanen (2014) from coherent stories, where the discourse status of the objects is made clear by the context. In (9a), the focal object bears no case ending and elicits no verbal agreement. (9b) contains a given object, which is marked by overt accusative case and triggers agreement on the verb.

(9) a. kom jowt-ryõõl wø-s (E Mansi; Virtanen 2014: (17),(11))
    man bow-arrow take-PAST.3SG
    'The man took a bow and an arrow'

b. õõw-ëm øät kont-iiɭ-øm
    door-ACC NEG find-SG<1SG
    'I don’t find the door.'

The correlation between agreement and topicality is not surprising. Givón (1975) argues on the basis of evidence from child language and Creol languages that subject agreement and object agreement suffixes were originally topic doubling pronouns cliticized to the verb. Among the present-day languages, the phenomenon is not unique to Ob-Ugric; Dalrymple and Nikolaeva (2011) discuss scores of languages from various language families where object–verb agreement encodes the topic status of the object.

The claim that agreeing Ob-Ugric objects occupy a VP-external topic position behind the primary subject-topic, whereas non-agreeing focal objects are in situ inside the VP can be supported by the following data of Nikolaeva (1999).

(i) Topical objects precede VP-adverbs, whereas focal objects follow them (10a,b). The focal object is left adjacent to the verb; only particles (e.g., the negative particle) can intervene between them:

(10) a. pro Šiti woj wel-l-øt (Khanty; Nikolaeva 1999: 61)
    so animal kill-PRES-3PL
    'So they kill an animal.'

b. *pro Wøj šiti wel-l-øt
    animal so kill-PRES-3PL

languages, including Hungarian (cf. É. Kiss of possessive suffixes, and their appearance on personal pronouns is a common feature of Uralic

This “Person Case Constraint” (the relics of which are also present in Hungarian, a sister language) is attributed by É. Kiss (2013, 2017) to an Inverse Topicality Constraint, which rules out

In some Ob-Ugric dialects, 1st and 2nd person objects cannot get accusative case and/or cannot elicit verbal agreement – as illustrated by the elicited Eastern Khanty examples in (13). The 1st and 2nd person objects in (13a,b) trigger no agreement, unlike the 3rd person object in (13c).

In Eastern Mansi, it is the accusative morpheme that is blocked in the case of 1st and 2nd person pronominal objects:

This „Person Case Constraint” (the relics of which are also present in Hungarian, a sister language) is attributed by É. Kiss (2013, 2017) to an Inverse Topicality Constraint, which rules out

In (14a), the 1st person object bears a 1st person singular possessive suffix. The non-possessive use of possessive suffixes, and their appearance on personal pronouns is a common feature of Uralic languages, including Hungarian (cf. É. Kiss and Tánczos 2018).
constructions where the structural hierarchy of topics is contrary to their ranking in the following topicality hierarchy: speaker participant > participant > non-participant of the discourse.

2.3 Topicalization of oblique internal arguments via promotion to object

The secondary topic position of the Ob-Ugric sentence can only be taken by an internal argument if it is promoted to object role. In ditransitive constructions with a beneficiary or goal promoted to object, the theme argument demoted from object bears instrumental (in a different terminology, instructive-final) case. Observe a Northern Mansi minimal pair analyzed by Skribnik (2001). In (15a), the so-called directive construction, the theme is object-topic, marked by accusative case and verbal agreement, whereas the goal bears a lative case ending. In (15b), the so-called secundative construction, the goal is promoted to object-topic, bearing accusative case and eliciting agreement; the theme has an instrumental case suffix.

(15) a. Who do you relate the tale to?
   Am mōjt tawen mōjt-i-lum. (N Mansi; Skribnik 2001: 228)
   I tale he.LAT tell-PRES-SG<1SG
   'I relate the tale to him.'

   b. What do you tell him?
   Am tawe mōjt-al mōjt-i-lum.
   I he.ACC tale-INST tell-PRES-SG<1SG
   'I tell him a tale.'

In the (a) example of (16), a Northern Khanty minimal pair, both the caseless theme object and the lative-marked goal are part of the verb phrase. In (16b), the goal is promoted to object-topic; it bears no visible case (in this dialect, only pronouns have overt accusative marking); but it elicits verbal agreement. Northern Khanty has no instrumental/instructive-final case; the demoted theme bears locative case.

(16) a. ānte-l ńāwrem-al-a ńāh ma-l. (Khart; F. Gulyás: 2015: (15), (16))
   mother-3SG child-3SG-LAT bread give-PRES.3SG
   'The mother gives bread to her child.'

   b. ānte-l ńāwrem-al ńāh-an ma-l-li.
   mother-3SG child-3SG bread-LOC give-PRES-SG<3SG
   'The mother gives her child some bread.'

In the Eastern Khanty sentence in (17a), the VP contains a dative-marked goal and a caseless theme. In (17b), the goal has become a secondary topic, bearing accusative case. It is a 1st person pronoun, hence the Inverse Topicality Constraint blocks verbal agreement with it.

(17) a. lūw mantem kat quλə-γən maj. (Khart; Sosa 2017: 118)
   (s)he I.DAT two fish-DU give,PAST.3SG
   'She gave two fish to me.'

   b. lūw mant kat quλə-γən-at maj.
   (s)he I.ACC two fish-DU-INS give,PAST.3SG
   'She provided me with 2 fish.'

A beneficiary or goal argument promoted to object tends to function as a secondary topic; but we also attest examples where it remains in the VP, eliciting no verbal agreement, e.g.:

---

8 The possessive suffix on *mother* can agree with a pro possessor coreferent with *her child*, but – as is common in the Uralic languages – it can also function as a (non-possessive) definiteness marker.
Though such examples are sporadic, they indicate that the secundative construction is not necessarily the result of a discourse-motivated movement rule. In fact, native speakers’ explanations suggest that the theta roles in the directive and secundative construals of a verb phrase are not completely identical. In the directive variant, the target of the action is the theme, whereas in the secundative variant, the target of the action is the beneficiary, with thetheme functioning as a means of providing for him/her. Therefore, I assume that the secundative construction is not a transformational derivative of the directive construction; the two constructions are the mappings of two different theta grids associated with ditransitive verbs.

2.4 Oblique subjects in active sentences

In Khanty, not only the demoted agent of passive sentences bears lative or locative case; the subject can bear the same oblique case in active sentences, as well. Oblique subjects are sporadic in Northern Khanty, occurring mainly in archaic folklore texts (19); they are rare in most parts of the Eastern Khanty area (20); but they are quite common in its easternmost Vasyugan subdialect (21).

(19) só:mì-na pos-ij-ǝl (N Khanty; Nikolaeva 1999: 43)
gold-LOC float-IMPF-PRES.3SG
‘The gold is floating down.’

(20) pan tom aj poy-ǝli-n wu-λ-tay. (E Khanty; Sosa 2017:189)
and this one boy-DEM.LOC know-PRES-SG>PRES.3SG
‘And this little boy knows it’

Second, locative subjects co-occur with an accusative object, as in (23):

(23) Lūy-an lūy-ǝt wijnat il-wel-te (Khany; Honti 1971: 432)
he-LOC he-ACC deliberately PRT-kill-PAST.3SG
‘He killed him deliberately.’

The Ob-Ugric locative subject is not an Icelandic-type quirky subject, either – because it is not lexically selected; it can occur with any verb. The function of the locative case of active subjects has been clarified recently by Sosa (2017); it can be used to mark a subject functioning as an unexpected shifted topic.

3 Analysis

The facts surveyed in Section 2 suggest that the highest noun phrase in the Ob-Ugric sentence functions as both grammatical subject and primary topic. That is, the AgrSP phrase, whose specifier
it occupies, is the projection of a head specified as both [+nominative] and [+topic]. AgrSP subsumes an AgrOP, the projection of a [+accusative, +topic] head. The specifier of AgrOP is the landing site of objects functioning as secondary topics. The AgrOP projection need not be activated in every transitive sentence, i.e., the object need not be topicalized. The projection and the filling of AgrSP, however, is obligatory – which is a manifestation of the EPP. The AgrS and AgrO heads are realized as subject agreement and object agreement suffixes. If the order of the suffixes on the Ob-Ugric verb mirrors the order of the functional projections headed by them, then in Mansi, AgrOP subsumes TenseP, VoiceP, vP and VP, as shown in (24). In Khanty, the order of VoiceP and TP is reversed; VoiceP dominates TP.

\[
(24) \quad \begin{array}{c}
\text{AgrSP} \\
\text{NP1} \quad \text{AgrS}' \\
\text{AgrOP} \quad \text{AgrS} \\
\text{NP2} \quad \text{AgrO}' \\
\text{TP} \\
\text{AgrO} \\
\text{T} \\
\text{VoiceP} \\
\text{T} \\
\text{Voice} \\
\text{vP} \\
\text{Voice} \\
\text{v} \\
\text{VP} \\
\text{v} \\
\text{V} \\
\text{V'} \\
\end{array}
\]

The noun phrase moved to Spec,AgrSP checks the phi-features of AgrS, which assigns nominative case to it, as usual; in addition, however, the noun phrase in Spec,AgrSP must also check the [+topic] feature of the AgrS head. (The [+topic] feature is essentially identical with the specificity feature proposed by Enç (1991), encoding that its carrier is identical with, or is a subset of, a previously established referent.) The constituent moved to Spec,AgrSP passes through Spec,VoiceP. The Voice head is either active or passive. An active Voice head selects a thematically complete verb phrase, whereas a passive Voice head selects a verbal projection (vP or VP) whose thematically highest argument is existentially bound but syntactically unrealized.\(^9\) The movement rule filling Spec,AgrSP targets the closest syntactically active constituent. If this argument has inherent case, its case is replaced by the phonetically null nominative assigned by AgrS.

AgrOP can be projected (or activated) in active transitive sentences; in other words, the AgrO head selects an active VoiceP subsuming a vP. If the directive and secundative constructions of ditransitive verbs discussed in Section 2.3 are indeed mappings of different theta grids, as suggested in Section 2.3, then movement to Spec,AgrO always targets the internal argument closest

\(^9\) A TP intervening between the Voice head and the vP or VP in Khanty is assumed to be transparent for selection.
to the verb. The noun phrase in Spec,AgrOP checks the number feature of AgrO, and – at least in Mansi – AgrO assigns accusative case to it. Crucially, the filler of Spec,AgrOP must also check the [+topic] feature of AgrO.

The tense, voice, object agreement, and subject agreement suffixes realizing the T, Voice, AgrO and AgrS heads are bound morphemes spelled out suffixed to the verb, which apparently remains in situ in the verb phrase.

4  **Theoretical implications**

The Ob-Ugric facts surveyed above necessitate the rethinking of certain theoretical assumptions of generative linguistic theory. In standard generative theory, passive NP movement is case-driven; it is triggered by the NP’s need of case. In Ob-Ugric, NP-movement can also target complements bearing an oblique case – which will be replaced by nominative case assigned by AgrS. Ob-Ugric NP movement is motivated by discourse requirements: Spec,AgrSP and Spec,AgrOP are to be filled by [+topic] constituents to check the [+topic] features AgrS and AgrO. AgrS also has an EPP feature, which makes movement to Spec,AgrSP obligatory (unless the verb has no complement other than a demoted subject).

The type of differential object marking attested most clearly in Eastern Mansi also raises questions concerning case theory. In Eastern Mansi, objects with no topic feature (i.e., non-referential objects and objects with unfamiliar referents), remaining in situ in the verb phrase, bear no visible case; objects extracted into Spec,AgrOP, however, are assigned an accusative case suffix – as was illustrated by (9a,b).

In the Khanty dialects, the accusative case ending has only been preserved on pronouns. In most Khanty dialects, 3rd person pronominal objects, practically always referring to previously introduced individuals, are always case-marked, whereas the case-marking of 1st and 2nd person pronominal objects is blocked by the Inverse Topicality Constraint. This is not the case in the Vasyugan dialect though; some data of Filchenko (2003) indicate that the accusative-marking of a 3rd person pronoun is not obligatory but depends on the [+topic] feature of the pronoun. In (25a,b), the object is non-topical (it is in the VP, following a VP-adjunct, eliciting no object-verb agreement); accordingly, it bears no accusative case in the Vasyugan example in (25a). (25b) illustrates the Surgut dialect, where accusative marking has been extended to [-topic] pronominal objects, as well.

(25) a. pay-ǝl-ǝn qoy juy waγa-γǝn.  (Vasyugan Khanty; Filchenko 2003: 113)
   son-3,SG-LOC long she call-PST.3SG
   ‘Her son called (for) her for a long time.’

   b. ma λəw-ǝt tɔŋmɛ-λ-ɛm.  (Surgut Khanty; Sosa 2017: 180)
   I he-ACC understand-PRES-1SG
   ‘I understand him.’

The Vasyugan example suggests that the discourse-motivated differential accusative marking preserved in Eastern Mansi may have been more general in Ob-Ugric.

The pattern attested in Eastern Mansi is problematic for case theory because it shows a dissociation of object licensing and accusative assignment. The default position of the object in the Mansi sentence (and in the Ob-Ugric sentence, in general) is the immediately preverbal position. In this position, however, the object is unmarked morphologically; it assumes accusative case if and only if it is raised to Spec,AgrOP, where it establishes an agreement relation with AgrO.

---

10 Northern Khanty does not have the Inverse Topicality Constraint.
The easiest explanation of this situation would be to assume that the VP-internal object has no case because it is incorporated into the verb. However, a VP-internal object is often clearly referential; e.g., it can be a complex possessive construction as in (26):

(26) ton k’än-pōāl-nø seemōl-nyoxøs, sājøng-nyoxøs jālp-øn
toāgl-äät nok-posyg-øs. (E Mansi; Virtanen 2014: (25))
cloth-3SG up-pull-PAST.3SG
‘Upon that he put on his sacred costume of black sable, white sable.’

Another possibility would be to give up the Case Filter, requiring that every overt noun phrase have case, and to accept that an object in situ is caseless in Ob-Ugric. A third option is to divorce syntactic Case (the feature responsible for object licensing) and morphological case as proposed by Bobaljik (2008), and to assume that the object complement is licensed by the V under government, and its accusative case is assigned by AgrO in a specifier–head configuration.

The locative case appearing on the subjects of active sentences, too, presents a problem for standard case theory. As discussed in Section 2.4, a subject-topic in Spec,AgrSP can optionally bear the oblique case of subject candidates demoted to adjuncts – if the subject functions as a shifted, recurring topic. That is, the semantic case that marks agents and other demoted subject candidates can replace the phonetically null structural case of an active subject in order to make it more salient. This possibility does not seem to fit into the framework of generative case theory; it could, at best, be analyzed as a PF phenomenon, a kind of stylistic variation. It could most easily be accounted for by a Case for Case (Fillmore 1968) style approach, where thematic roles are encoded by different case endings and adpositions. The thematically determined cases and adpositions are neutralized in subject and object positions, but remain accessible, and can be put to use when e.g. a recurring subject-topic needs to be emphasized. This approach would also account for the disappearance, or neutralization, of the oblique case of locative, goal, and beneficiary arguments promoted to subject in passive constructions.

5 Summary

In the Ob-Ugric sentence, subject agreement and object agreement also encode the topic function of the subject and the object, respectively. The fusion of the subject and topic roles has far reaching consequences in Ob-Ugric grammar. The constituent in the specifier of AgrSP must have the feature [+topic], therefore, a [-topic] subject candidate has to be demoted, and – owing to the EPP feature of AgrS – a [+topic] internal argument must be promoted to subject. Subject demotion and internal argument promotion are realized via passivization. Subject demotion can involve subjects of transitive, unergative and unaccusative verbs alike, and internal argument promotion, i.e., NP movement, can target objects, oblique internal arguments, and even adjuncts. Consequently, NP-movement is not a case-driven operation; it is triggered by the interplay of the [+topic] feature of AgrS in need of checking, and the EPP.

In transitive clauses with a [+topic] subject and a [+topic] object, the object undergoes NP-movement to Spec,AgrOP, a secondary object position. In Eastern Mansi, the object moved to Spec,AgrO and entering an agreement relation with AgrO is also assigned accusative case by it. Focal objects remain in the VP caseless. These facts indicate that object liensing by the verb under government, and case assignment by AgrO via specifier–head agreement are separate processes – as suggested by such recent versions of case theory as Bobaljik (2008).

A further discourse-motivated property of Ob-Ugric is the optional replacement of the nominative case of subjects of active verbs functioning as recurring topics with an oblique case. This phenomenon – as well as the replacement of the oblique cases of internal arguments promoted to subject with nominative case – argues for a Fillmore (1968) style approach to case. Namely, every
noun phrase has a thematically motivated case ending or adposition, which is neutralized in subject or object position, but is accessible if needed, e.g., for the encoding of some discourse feature.

References


What microvariation can show us: An analysis of the Basque discourse particle ote

Sergio Monforte
sergio.monforte@ehu.eus

1 Introduction\

There have been two main topics on the syntax of modal or discourse particles lately:1 On the one hand, the debate on the syntactic status of these particles between those positioning in favour of considering discourse particles as heads and those claiming that particles are phrasal occupying specifier positions. On the other hand, the position of discourse particles in the syntactic structure also has received attention in the literature with some proposals locating them in the TP-domain (Elordieta, G. 1997; Coniglio 2007; Bayer 2009), some others in the CP-domain (Kuong 2008; Coniglio and Zegrean 2012; Kuwabara 2013) and (Elordieta 1997; Coniglio 2007; Bayer 2009) finally, other studies which claim that they occupy a position in the Speech Act layer above CP (Munaro and Poletto 2002; Li 2006; Haddican 2014).

This paper deals with the Basque discourse particle ote and its microvariation in eastern dialects; the data found in those varieties brings up the question whether the behaviour of discourse particles in Basque agrees with that described for heads, as has been claimed in the literature (Albizu 1991; Elordieta 1997; Elordieta, A. 2001; Haddican 2008; Arregi and Nevins 2012; Lizardi-Ituarte 2017), or they should be also considered maximal projections (Etxepare 2010; Etxepare and Uria 2016). In addition to this, two novel behaviours of discourse particles in Basque grammar will be presented: the combination of the particle and wh-words creating a single constituent and the use of ote at the end of the utterance, similar to sentence final particles (Izutsu and Izutsu 2013). The paper is organised as follows: I briefly introduce the literature concerning the syntactic status of discourse particles before I move on to the analysis of the discourse particle ote in Basque; then in section 2, after giving basic notions on the grammar and syntax of Basque, I present data of standard Basque and Eastern Basque concluding that Basque discourse particles behave not only as heads but also as weak adverb; in section 3 I described the construction formed by a wh-word and the discourse particle ote and I argue that they form a single constituent, unlike the analysis claimed by Munaro and Poletto (2002) for similar structures in North Italian dialects; finally, in section 4 I provide another piece of dialectal data related to ote occurring at the end of the utterance and conveying an intersubjective interpretation.

---

1 I am grateful to the audiences of the Clause Typing and the Syntax-to-Discourse Relation in Head-Final Languages workshop and Norio Nasu, Ricardo Etxepare and Xabier Artiagolitxu for their comments. This study has been made possible thanks to the research project PGC2018-100686-B-I00 from the Spanish Ministry of Science, Innovation and Universities. All errors are mine.

1 I will use the term ‘Discourse particle’ to refer to those which have been traditionally referred to as ‘Modal particles’.
2 About the syntactic status of discourse particles

The term particle, as pointed out by Paul (2015), has been used for those words which have not successfully been assigned a category. Besides, since discourse particles in languages such as German or Italian are diachronically related to other items with a clear category, both discourse particles and their historically related counterparts have been claimed to belong to the same category, no matter what their function is in the proposition. For instance, some of the discourse particles in German derive from adverbs (durch, einfach, auch, eingentlisch, wohl, bloß and schon according to Meibauer, apud Struckmeier 2014: 18); therefore, some authors (Jacobs 1991; Zimmermann 2004; Cardinaletti 2011) consider them a subclass of adverbs (Struckmeier 2014).²

The vagueness of their character complicates not only the categorisation of those words which may belong to the group of discourse particles, but also the identification of their status within the generative framework. If we consider two well-studied languages such as German and Italian, the two hypotheses mentioned earlier have been claimed for those languages. Analyses in favour of the headness status take into account the following evidence: a) they have a fixed order, unlike the majority of adverbs;³ b) they cannot occur in first position in V2 languages such as German; c) they behave mostly as clitics; d) they are not stressed;⁴ e) they are diachronically related to elements of other categories; f) they cannot be the only element as a reply to a question; g) they cannot be coordinated or modified; h) they cannot be topics or focus.⁵ Apart from those, other general properties are also usually listed when describing modal particles, for instance, that they cannot all occur in every kind of clause, i.e. they are clause-type dependent, or that they can only be used in clauses containing full illocutionary force, hence, they can appear in root clauses and in some embedded clauses.

Some authors (Bayer and Obenauer 2011; Struckmeier 2014) claim that German discourse particles behave just as described above; therefore, they consider German particles to be heads, namely the head of the ParticleP located in the Middle field. In a similar way Coniglio (2008) concludes that Italian discourse particles are heads. Interestingly, he also examines German discourse particles and arrives at the conclusion that they are Maximal projections occupying a specifier position.

Coniglio is not the only one denying the head status to German particles; indeed, Cardinaletti (2011) supports this idea. Based on a previous work on pronouns (Cardinaletti and Starke 1999), she distinguishes three levels for adverbs: full adverbs, weak adverbs and clitic adverbs. She claims that discourse particles in German are weak adverbs considering their syntactic and phonological behaviour: Let us briefly illustrate the main arguments they use in favour of the non-head status (Munaro and Poletto 2002; Coniglio 2007; Cardinaletti 2011): a) discourse particles are closer to full words than clitic functional words concerning their phonological properties and prosody; b) if they were heads, they should block V raising, for instance, in German where they

² Lindner (1991: 163) asserts that discourse particles have been traditionally called adverbs.
³ As pointed out by a reviewer, not all adverbs enjoy free distribution, in fact, some of them such as just, well, often, right, even have limited distribution.
⁴ This property is an opened issue in the description of modal particles, since authors do not agree whether modal particles receive stress or not (Thurmair 1989; Coniglio 2007; Cardinaletti 2011; Egg and Zimmermann 2011; Struckmeier 2014).
⁵ Following Gutzmann (2015) and Scherf (2016), I do not consider that properties (g) and (h) present evidence of the syntactic status since those restrictions are not related to the syntax of modal particles, but to their semantics.
are supposed to occupy a position between TP and vP, or, otherwise, particles should move along with V; c) scrambled DPs and PPs can appear between two discourse particles, not an expected distribution if they were syntactic heads.

Therefore, Cardinaletti (2011) and Coniglio (2007; 2008) conclude that discourse particles in German occupy a specifier position where they remain since they are weak adverbs, i.e. because of the lack of certain phonological properties.⁶

So far, I have reviewed briefly the hypotheses on the syntactic status of discourse particles. In the following, I will present data of the Basque language to conclude that discourse particles in that language behave as clitic-heads. Nevertheless, microvariation related to the discourse particle ote found in Eastern Basque shows that this particle can also behave as a weak adverb. This will be also examined below.

3 Discourse particles in Basque

Along these lines I will prove that discourse particles in Basque are syntactically heads. Nevertheless, first some properties of the Basque language will be explained so that the reader can fully understand the data to be dealt with below.

3.1 Some notions on Basque grammar

Basque is mostly classified as a non-rigid SOV language (Villasante 1980; Hualde and Ortiz de Urbina 2003; Rijk 2008); that means that, although SOV order has been identified as the neutral one, other phrase combinations are possible conveying different pragmatics at the level of information structure, for instance:

1 Xabier etxera etorri da.
   Xabier house.ADL come AUX.PRES.IND.3SG.ABS
   ‘Xabier came home.’

2 XABIER etorri da etxera.
   Xabier come AUX.PRES.IND.3SG.ABS house.ADL
   ‘XABIER came home.’

3 Xabier, etxera etorri da.
   Xabier house. ADL come AUX.PRES.IND.3SG.ABS
   ‘Xabier, he came home.’

As can be observed in these examples, finite verbs are for the most part analytic, i.e. composed of a morphologically independent lexical verb carrying aspectual information and an auxiliary form bearing tense, mood, and agreement with the arguments and, in some cases, also the addressee (Oyharçabal 1993; Miyagawa 2012). Additionally, there are about 12 verbs which can be synthetic but only when the aspect is punctual (Laka 1996), for instance:

---

⁶ According to Cardinaletti (2011) elements have three levels: phonological, syntactic and semantic. If, at least, one level lacks, they cannot be considered full words and they must be either weak elements or clitic ones.
Leirek hori dakar.
Leire.ERG that.ABS bring.PRES.IND.3SG.ERG.3SG.ABS
'Leire is bringing that.'

These synthetic forms present the same information as periphrastic ones but realised in a single constituent. Also, they show similar syntactic behaviour as auxiliary verbs.

Lexical and inflected verbs are usually adjacent; indeed, nothing can occur between them, such as adverbs:

5 * Xabier etxera etorri lehen da.
   Xabier house.ADL come before AUX
   'Intended: Xabier came home before.'

Nevertheless, the adjacency between lexical and auxiliary verbs can be broken in some contexts such as in negative (main) clauses:

6 Ez naiz Lindaura joan ez naiz
   not AUX Lindau. ADL go
   'I didn't go to Lindau.'

Furthermore, Eastern Basque offers another context where this adjacency does not arise: while the standard procedure to form focal structures and wh-questions is fronting both lexical and auxiliary verbs, in eastern dialects the inflected verbal form can be the only constituent to trigger movement to be next to the focal element or wh-word, leaving the lexical verb in-situ:

Standard Basque:

7 MAITENAK [erran du] MAITENAK hori [erran du]
   Maitena.ERG say AUX that.ABS
   'MAITENA said that.'

Eastern dialects:

8 MAITENAK [du] MAITENAK hori erran [du].
   Maitena.ERG AUX that.ABS say
   'It was Maitena who said that.'

Nevertheless, the non-adjacency of these elements is not limited to contexts where the finite verb triggers movement; indeed, discourse particles also constitute another context where this adjacency is broken; in fact, the canonical position of discourse particles in declarative sentences is between the lexical and the auxiliary verbs:

9 Ikasle guztiek liburua irakurri omen dute.
   student all.ERG book.ABS read PRT AUX
   'All students apparently read the book.'

Concerning the formation of questions, Basque is an SOV language with wh-movement and question particles (cf. Cheng 1997; Bruening 2007); therefore, in wh-questions not only the wh-word moves to the CP-domain in questions, but also the verb. In the case of polarity questions, there is also fronting of the constituent composed by the lexical and inflected verbs; let us briefly exemplify:
Finally, complementizers in Basque function as clitics attached to the finite verb; accordingly, in a neutral embedded context this appears at the end of the sentence. However, if the verb moves because there is, for instance, a focal constituent, the complementizer moves along with the finite verb.\footnote{In terms of the syntactic structure, I follow Ortiz de Urbina’s (1999) proposal that complementizers occupy the head of Finite Phrase following the cartographic approach; this phrase is head-final, unlike those above it which are head-initial.}

```
10 [\text{\texttt{[CP Zer \{[^{C}_{C^{0}} \text{erosi du]} [\text{TP Asierrek \text{zer \text{erosi du}}] \} ] ?}}
\text{What buy AUX Asier.ERG}
\text{'}What did Asier buy?’\]

11 [\text{\texttt{[CP \{[^{C}_{C^{0}} \text{Erosi du]} [\text{TP Mikelek etxea \text{erosi du}}] \} ] ?}}
\text{buy AUX Mikelek.ERG house.ABS}
\text{’Did Mikel buy the house?’}\]
```

Now that we have established some basic properties of Basque grammar, I will proceed to examine the syntax of discourse particles in Basque in the next section, with special focus on the particle \textit{ote}.

### 3.2 Head status of Basque discourse particles

Although the list of discourse particles in Basque is not as long as in German, they can be classified into different groups: those conveying evidentiality (\textit{ei, omen}), those expressing epistemicity (\textit{ahal, bide, ote}) and the interrogative particle (\textit{al}); for instance:

```
13 \text{\texttt{\text{Txiki lurretan lanean hasi omen da.}}}
\text{\texttt{\text{ Txiki lurreta.IN work.IN begin PRT AUX}}}
\text{\texttt{’Apparently, Txiki began working in lurreta.’}}

14 \text{\texttt{\text{Elurra botako ahal du!}}}
\text{\texttt{\text{ snow.ABS throw.FUT PRT AUX}}}
\text{\texttt{’I wish it snows!’}}
```

Another shared characteristic is that all discourse particle function as proclitics to the inflected form. Evidence of this relationship is found in 1) negative contexts (as shown in 15), 2) focus contexts in eastern dialects (as shown in 16), and 3) non-inflected clauses where discourse particles cannot occur (as shown in 17 and 18):

```
15 \text{\texttt{\text{Ez al duzu egunkaria \text{erosi ez al duzu?}}}}
\text{\texttt{\text{ not PRT AUX newspaper.ABS buy}}}
\text{\texttt{’Didn't you buy the newspaper?’}}

16 \text{\texttt{\text{Jonek ote dia Jonek hori erran ote du?}}}
\text{\texttt{\text{ Jon.ERG PRT AUX.PRT that.ABS say}}}
\text{\texttt{’Was it Jon who said that? (I'm wondering)’}}
```
17 Horierosi (*al) eta bestea bota al duen galdetu dut.
this.ABS buy PRT and other.ABS throw PRT AUX.C ask AUX
'I asked whether s/he bought this one and threw away the other one.'

18 Ez dakit nora joan (*ote)
not know where.ADL go PRT
'I don’t know where to go.'

Additionally, being a proclitic to the inflected verb implies that nothing can intervene between the particle and the finite verb; this prediction is borne out:

19 Txiki lurretan lanean hasi omen (*berriro) da.
Txiki lurreta.IN work.IN begin PRT again AUX
‘Apparently, Txiki began working in lurreta.’

Indeed, if the particle stays in-situ after the lexical verb, the sentence would be grammatically wrong:

20 * Jonek dia Jonek hori erran ote du?
Jon.ERG AUX.PRT that.ABS say PRT

Furthermore, evidence that discourse particles behave as clitics comes from the fact that they do not count as a constituent when the verb is in clause initial position. As described above, finite verbal forms are banned from the very first position and, therefore, an expletive morpheme ba is required before the verb as showed in (21). Clauses containing a discourse particle still have this requirement although particles occur to the left of the finite verb in first sentential position:

21 * (Ba) omen daki.
CL PRT know
‘Apparently, s/he knows it.’

So far, this pattern is attested in all particles; indeed, the only difference they present concerns their interpretation.

Based on this evidence, Basque grammarians have considered discourse particles as clitic-heads which occupy a position in the TP-domain (Albizu 1991; Elordieta 1997; Elordieta 2001; Monforte 2018b), since they always move along attached to the head of TP.

In addition to function as clitics, particles in Basque also display a fixed order, since they always occur precedent to the inflected verb, as observed in (15-16). Also, discourse particles in Basque cannot be topic or focus since particles cannot occur in first position as shown in (21); in fact, particles can occur when there is a focalized constituent, however, the particle will never be the focus.8 Let us recall that nothing can intervene between the focus and the verb and that it is the position right before the verb the one focal elements occupy; therefore, if the discourse particle in (22) were in [spec, FocP], XABIER could not be the focus of the proposition:

8 Gutzmann (2015: 222) claims that discourse particles are scopeless because they cannot be the topic of a question and, in the same way, they cannot give rise to alternatives, i.e. they cannot be considered semantically focused.
Concerning whether discourse particles can receive stress or not, although the prosodic characteristics of particles in Basque should be studied deeply, at first glance they seem to form a prosodic unit with finite verbs.\(^9\) Interestingly, these particles may suffer apheresis in some contexts as the result of the position of the accent:

23  Loak artu mertzun.  [omen > men]
    sleep.ERG take PRT.AUX
    ‘Apparently, s/he fall asleep.’ (Barandiaran 1972)

24  Nun utzi te (d)et periodikua?  [ote > te]
    where leave PRT AUX newspaper.ABS
    ‘Where did I leave the newspaper? (I'm wondering)’

They cannot occur by themselves; this property can be applied to particles in Basque too since they need to occur always attached to the inflected verbal form:

    arrive AUX David Boston.ADL PRT
    ‘- Did David arrive in Boston? - Reportedly.’

Finally, it is an acknowledged fact that discourse particles in German or Italian (Thurmair 1989; Coniglio 2008; Cardinaletti 2011; Bayer 2012; Bayer and Struckmeier 2017) are historically related to other categories such as adverbs or conjunctions. In Basque we do not see a clear relation between discourse particles and other categories; the only exceptions may be bide (DiscP) / bide ‘way’ (N) and ahal (DiscP) / ahal ‘to be able to’ (Modal V).\(^{10}\)

In this section I have demonstrated that those properties cross-linguistically claimed for discourse particles are also found in Basque ones. However, what it is relevant is that those properties give rise to the idea that particles are heads; therefore, the hypothesis that Basque discourse particles are heads is reinforced here.

### 3.3 Microvariation in eastern dialects: the discourse particle ote

Discourse particles have been shown above to function as heads based not only on their syntactic behaviour, but also on the data collected by testing the cross-linguistic properties of discourse

\(^9\) In the Basque of Lekeitio (Hualde et al. 1994: 57) the presence of the MPs ei or ete has an impact on the location of the stress provoking that the particle and the auxiliary constitute a new prosodic compound which differs from the one made up of the auxiliary and the main verb:

1  Gâur etorri- diras?
    today come- AUX
    ‘Did they come?’

1  Gâur etorri ete diras?
    today come PRT AUX
    ‘Did they come? (I'm wondering)’

\(^{10}\) See Monforte (2018a) for an analysis of the question particle al based on the grammaticalization of the epistemic discourse particle ahal. Similar processes have been also claimed for the German discourse particle denn (Bayer 2012) and the Ladin particle po (Hack 2014).
particles. Now I will provide evidence that *ote* can behave either as a head or as a weak adverb in Eastern Basque.

As observed before, discourse particles have a fixed position in the clause, i.e. precedent to the inflected form. Nevertheless, data from eastern dialects go against this statement, since *ote* occurs also in a post verbal position:

26 Ez dūa ote eginahala egin?
not AUX.PRT PRT effort.ABS do
‘Did s/he do everything possible OTE11?’ (Coyos 2013)

27 Bena ezpitakit nik lamina horiek zer zien othe!
but not.c.know I.ERG mermaid those.ABS what were.C PRT
‘But I don’t really know what those mermaids were OTE.’ (Camino 2017)

28 Eta gero Jainkoa zertako dugun ote samur, estonatuko gira gu!
and then God.ABS why have.c PRT soft astonish.FUT
AUX we
‘And then we’ll be astonished why God is hard on us OTE.’ (Hiriart-Urruty 1972)

The fact that *ote* always occurs after the verb is not surprising since we are dealing with questions, and those, as described above, always show the verb fronted in matrix questions and embedded contexts, if they are *wh*-questions, or polar questions containing a focal constituent. Nevertheless, it does not appear right after the finite verb, because complementizers, the interrogative mark -a or the discourse particle *bada* occur in between:

29 Nor deitzen du bada ote Peiok egun guziz hain goizik?
who call.IPFV AUX PRT PRT Peter.ERG day all.INS so soon
‘Who does Peter phone every morning so early? OTE’

However, this kind of *ote* also have a specific position in the syntactic structure, after the finite verb and DP *bada* and before the subject; in fact, appearing in other positions brings ungrammaticality:

30 Nor deitzen du (ote) Peiok (*ote) egun guziz (*ote)?
who call.IPFV AUX PRT Peter.ERG PRT day all.INS PRT
‘Who may Peter phone every morning?’

This agrees with the properties proposed by Munaro and Poletto (2002) and Cardinaletti (2011), i.e. that weak adverbs, unlike full adverbs, may only appear in a fixed position.

Following these authors, proof that the particle does not function as a clitic to the finite verb comes from the fact that they do not form a prosodic unit; in fact, *ote* forms a single prosodic unit by itself. Consequently, it cannot be phonetically reduced in contrast to what can happen when *ote* behaves as a head (see 24):

---

11 Although I have translated the particle *ote* as ‘I wonder’ in previous examples, following the traditional interpretation (Rijk 2008), this particle may have a different interpretation in the following contexts; therefore, I will not provide paraphrases for the next examples and mark their contribution by using ‘OTE’.
This is also attested in Austrian German (Coniglio apud Bayer and Obenauer 2011); in this variety the adverb *vielleicht* ‘maybe’ can be reduced to *leicht* but only when it functions as a discourse particle, i.e. as a head:

32 * [vielleicht > leicht] ist DIE schlau!
   this.FEM is PRT PRT smart
   ‘My God, how smart this one is!’

As stated above, German discourse particles cannot occur in first sentential position; *vielleicht* can occur in that position but only if it had an adverbial reading and not a particle reading.

In addition to this, Coniglio (2007) claims that modal particles should block the rise of the finite verb to CP or they should move along together, if they were syntactic heads. This behaviour is found in Basque modal particles as observed in examples (15 & 16); nevertheless, the kind of *ote* under examination now does not show those properties, since it does not block the movement to the left of the finite verb and it stays in-situ, as can be observed in example (29).

As for the fact that scrambled PPs or DPs can intervene between two modal particles, this does not seem to be an option in Basque, since the only two modal particles which do not behave as clitic heads, i.e. *omen* and *ote*, are not semantically compatible, hence, they cannot occur in the same clause.

Based on this evidence, I conclude that *ote* in eastern dialects can behave not only as a head, but also as weak adverb12 following Cardinaletti’s (2011) terminology. We find this last syntactic status in (26-28) examples since it does not have to occur adjacent to the finite verb, i.e. it does not behave as a clitic, and, indeed, it does not have the properties clitic elements usually present us with.13

3.4 Interim conclusions

At the beginning of this section, I brought up the question whether discourse particles are syntactically heads or phrasal. In order to answer this question, I presented data from Basque. This language has several discourse particles which clearly behave as heads. However, as it is acknowledged, Basque is rich in its microvariation; in fact, the analysis of the data offered by the particle *ote* in eastern dialects provides evidence of a distinguished behaviour: *ote* can behave not only as a head but also as a weak adverb (cf. Etxepare 2010; Etxepare and Uria 2016).

Therefore, based on the data found in the Basque language, both approaches seem not to be mutually exclusive even in the same language; in fact, future analyses on the syntactic status of

---

12 More evidence in favour of the adverbial status comes from the fact that postpositions can adjoin *ote*, for instance: *otean* (*ote* + inessive) or *oterik* (*ote* + partitive case). Further research is needed.

13 Etxepare (2010) and Etxepare and Uria (2016) examine the evidential particle *omen* which also shows microvariation in the same dialects and claim a similar analysis in terms that *omen* can function as a head and as a phrase.
discourse particles considering the properties displayed by Basque particles (as heads) should improve and clarify the vague distinction between the two statuses since both are found in the same language and are undoubtedly differentiated.

4 On an (un)usual context of discourse particles

Discourse particles mostly show behaviours as those described above, i.e. they are free morphemes, or they attach to a concrete word in the clause, usually the closest head in its c-commanding domain. Beside these patterns, there is another one which has recently received more attention in different languages, i.e. ‘wh-word particle’ configuration, although it does not seem as common as the use of particles described above. This construction consists of a discourse particle attached to a wh-word, following the order wh-word - particle:

34 wh-word P / *P wh-word

Configurations formed by wh-words and discourse particles are found crosslinguistically and have received distinct analyses: some posit that they form a single constituent (Bayer and Obenauer 2011; Bayer and Trotzke 2015; Endo 2018), whereas others (Munaro and Poletto 2002) claim that the particle always occupies the same position in all types of constructions. Consider languages such as German (Abraham 1991; Bayer and Obenauer 2011; Bayer and Trotzke 2015), Italian (Munaro and Poletto 2002; Coniglio 2008), Dolomitic Ladin (Hack 2014) and Japanese (Endo 2018): 14

35 Von wem schon kann man das sagen?
of who PRT can one say
‘Who can one say that about? About nobody!/ Hardly about nobody!’

36 Cosa mai avrebbe Gianni potuto fare in quel frangente?
what PRT would.have Gianni could do in that occasion
‘What could Gianni do on that occasion?’

37 Nani-yo John-tara kidotteru wa
what-PRT John-TOP vain mood
‘John is so vain/ John acts cocky’

Similarly, Basque also has this configuration, although it can only be found in north-eastern dialects (Trotzke and Monforte 2019):

38 Non ote utzi dut kazeta?
where PRT leave aux newspaper.ABS
‘Where did I leave the newspaper OTE?’

39 Zer ote ari da haur hori?
what PRT PROG AUX child that
‘What is that child doing OTE?’

14 See Hagstrom (1998), Kishimoto (2005) and Cable (2008) for an interesting analysis of the interaction between particles and wh-words in Sinhala, Tlingit and Japanese considering that the particle can appear not only attached to a wh-word but also at the edge of the clause.
40 Ez dakit nik nola ote egiten ahal zükean.
not know l.ERG how PRT do.IPFV can AUX.C
'I don’t know how it could be done OTE?'

41 Nork ote jan züan?
who.ERG PRT eat AUX
Who ate it OTE?

This construction arises not only compound by a wh-word and the particle ote, but also by the discourse marker ba(da):

42 Nondik bada zetozen eskatu zien.
where.ABL PRT come c ask AUX
'S/he asked them where they were then coming from.' (Borda 2005)

43 Zergatik bada erraten dizkiodan burutik pasatzen zaizkanak oro.
why. PRT say.IPFV AUX head.ABL pass.IPFV AUX.C.ABS all
(…) why I didn’t tell him all that was going through my head. (Landart 1999)

Returning to the analysis of ote in these configurations, although I accept the idea that wh-words and the discourse particle ote form a single constituent, two other hypotheses could be claimed: on the one hand, ote could be attached to the verbal complex, i.e. the merge between lexical and finite verbs, instead of being adjacent to the finite verb as in the standard behaviour. In other words, [ ote [ V Aux ] ] would be the analysis, in contrast to the general pattern [ V [ ote Aux ] ]. On the other hand, analyses within the Italian languages claim that the particle occupies the same position in this kind of configuration as when it functions as a weak adverb not attached to another constituent (Munaro and Poletto 2002; Coniglio 2008); the difference between them is related to the movement of the verb or the whole CP:

44 Quando, po, eli rivadi?
when PRT have.they arrived

45 Quando eli rivadi, po?
when have.they arrived PRT

46 a. [FP wh [fo particle] [CP t [IP …t….]] ] [Derivation of (44)]
   b. [FP CP [fo particle] [CP wh [IP …t….]] ] [Derivation of (45)]

Nevertheless, as the following data shows, these approaches must be rejected, at least, regarding the ‘wh-word ote’ combination:

47 Ez dakit non ote kazeta utzi dudan.
not know where PRT newspaper.ABS leave AUX.C
'I don’t know where I left the newspaper OTE.’

48 *Ez dakit non kazeta ote utzi dudan.
not know where newspaper.ABS PRT leave AUX.C

49 Ez dakit non ote utzi dudan kazeta.
not know where PRT leave AUX.C newspaper.ABS
'I don’t know where I left the newspaper OTE.’
On the one hand, these examples prove that *ote* is attached to the *wh*-word since 1) in North-Eastern Basque the adjacency between the *wh*-word and the verb is not obligatory, at least, in embedded contexts, as can be observed in (47); in those cases the *wh*-word and the particle move along together to be fronted and they appear contiguous, no matter whether the verb, too, moves to be adjacent to the *wh*-word, (49), or not (47); 2) *ote* cannot occur before an in-situ verbal complex as in (48).

Also, it is an acknowledged fact that nothing can intervene between the *wh*-word and the verb when the latter has been fronted to the CP-domain:

50 * Zer atzo erosi zuen Asierrek zet erosi du?
   What yesterday buy AUX Asier.ERG
   Intended: 'What did Asier buy yesterday?'

On the other hand, I do not agree with hypotheses which claim that the particle in this construction and the one examined in (3.3) occupies the same position and the surface distribution between the particle and the *wh*-word varies on the movement of the verb (Munaro and Poletto 2002), i.e. the analysis that if the verb triggered movement to the CP-domain, we would get the order of configurations described in section (3.3) and, if it did not, the construction dealt with in this section would arise:

51 a. [XP whi [VP *ote [ZP .... t..... ] ] ]
   b. [XP whi [Vj ] [VP *ote [ZP .... t, tj... ] ] ]

However, I have shown that movement of the verb does not determine the occurrence of one structure or the other one, since the configuration between *wh*-word and particle occurs in (47) where the verb stays in situ and in (49) where the verb moves to FocP.

Another analysis could also be suggested: the *wh*-word, in its movement to the CP-domain, merges with the weak adverb *ote* which occupies the specifier position of a Phrase located between FocP and TP and then they, *wh*-word and *ote*, move to FocP.

Even though this is an elegant analysis for this construction, there is data to discard it. If we accept the idea that the *wh*-word ‘why’ does not occur below TP and does not trigger movement to CP because it already occurs in CP (Cecchetto and Donati 2012), then the *wh*-word could not adjoin to the particle by head movement to the left periphery; however, there is data showing that ‘why’ merges with *ote*:

52 Zergatik ote galdegin dit Peiok hori?
   why PRT ask AUX Peter.ERG that.ABS
   ‘Why did Peter ask me that OTE?’

53 Zergatik ote Peiok hori galdegin dit?
   why PRT Peter.ERG that.ABS ask AUX

54 * Zergatik Peiok hori ote galdegin dit?
   why Peter.ERG that.ABS PRT ask AUX

Based on this evidence I conclude that the hypothesis of a structure ‘*wh*-word *ote*’, or generally stated ‘*wh*-word Particle’, is appropriate for the analysis of configurations such as those described here. Indeed, prosodic data reinforces this approach, since the particle is not only pronounced
forming a prosodic unit with the wh-word, but also the wh-word, and as a result the [wh-word ote], receives an extra strong accent.15

5 Ote as a sentence final particle

As remarked above, Basque is rich in microvariation not only among dialects, but also among subdialects. The particle ote has been a good example of this richness, since it shows differentiated behaviour in eastern dialects. Ote can function as a head, a weak adverb and it can also be attached to wh-words. Furthermore, there is a fourth distinguished syntactic use of ote in these varieties: the sentence final particle ote.

We have seen in section (3.3) that the particle ote can appear after the verb as in:

55 Egiazko apeza dea ote hori? real priest:ABS is:PRT PRT that
   ‘Is that person a real priest OTE?’ (Salaberry 1978)

However, some examples are not as clear regarding the position of ote since it could be considered to occupy a position in the right periphery:

56 Jeina, ni naza ote?
   Lord I be:PRT PRT
   ‘Lord, is it me OTE?’ (Estornès 1982)

57 Orain gazte batekin zirea ote?
   now young one:COM be:PRT PRT
   ‘Are you with a young one now OTE?’ (Thikoipe 2009)

Both examples show ote at the end of the clause, but I will argue that they do not occupy the same position. First, if we change the position of ote and place it preceding the finite verb, as in the general pattern, differences arise: the former is acceptable also for speakers of western dialects, whereas the latter is not. This leads me to the idea that these uses of ote may not convey the same interpretation; however, further research should be done on this topic in order to clarify this point.

Syntax also provides evidence that we are dealing with a different kind of ote. If we compare the following examples with those in (3.3), we observe that in (58) the particle appears sandwiched between the finite and lexical verb, as expected for constituents occupying a position between FocP and TP; nevertheless, in (59-60) the particle occurs after all the components of the sentence even the lexical verb in negative sentences:

58 Ez düa ote eginahala egin?
   not AUX:PRT PRT effort:ABS do
   ‘Did s/he do everything possible OTE?’

59 Ez girea bikurarat gomitatuak ote?
   not AUX:PRT meeting:ADL invited PRT
   ‘Weren’t we invited to the meeting OTE?’

15 See Bayer (2009), Bayer and Trotzke (2015), Trotzke and Turco (2015) for an interesting analysis of this phenomenon relating to emphasis and mirativity.
What microvariation can show us: An analysis of the Basque discourse particle ote

60 Ez da joan ote?
   not AUX go PRT
   ‘Didn’t s/he go OTE?’

In fact, if the particle ote is moved to be next to the finite verb, the judgments of native speakers on examples such as (59) clearly reflect that this use of ote has not the same interpretation as the previous ones, since they reject the use of the particle:

61 * Ez ote girea bikurarat gomitatuak?
   not PRT AUX.PRT meeting.ADL invited
   ‘Weren’t we invited to the meeting OTE?’

In conclusion, this is a separate use of ote which seems to appear in what has been referred to as the Right Periphery and occupy a position above CP. Further analysis of the syntax of this use should be conducted in the future. Also, the interpretation of this kind of ote is beyond the scope of the aim of this paper; however, based on the context of these examples, I could arrive at a first conclusion: unlike the interpretation of ote in the other cases, this has an intersubjective meaning which apparently matches the fact of being in the utterance-final position (Izutsu and Izutsu 2013) since the speaker seems to expect a specific response from the addressee by using the particle ote at the end of the utterance.

6 Conclusion

The discourse particle ote has been examined here considering its microvariation. After proving that discourse particles in Basque behave as heads, I have turned to data gathered from eastern dialects where the particle ote can function not only as a head but also as a weak adverb based on the syntax (the former is a clitic always adjacent to the finite verb, the latter is fixed in a specific position but not attached to another constituent) and on phonological data (the former can suffer apheresis and forms a prosodic unit with the finite verb, whereas the latter cannot be phonetically reduced and it shows a distinct intonation contour, i.e. it is prosodically independent). On the other hand, I have presented original data from eastern dialects which gives evidence of two distinguished uses: the ‘wh-word Particle’ configuration, also found in German, Italian, Dolomitic Ladin or Japanese, and the utterance final ote presumably conveying intersubjectivity.

References


Addresses of the Contributors

Tamer Akan
Ankara University
Department of Linguistics
06100 Sihhiye/Ankara
Turkey

R. Amritavalli
63, "Srinivasa Kripa", 1st Main, 1st
Cross
Thyagarajnagar
Bengaluru 560 028
India

Rahul Balusu
Department of Linguistics and Phonetics
#99, Main Building
English and Foreign Languages University
Hyderabad 500 007
India

Probal Dasgupta
Kaberi, First Floor
123A, Purbachal Kalitala Road
Kolkata 700 078
India
Katalin É. Kiss
Research Institute for Linguistics
Benczúr u. 33.
1068 Budapest
Hungary

Yoshio Endo
Graduate School of Language Sciences
Kanda University of International Studies
1-4-1, Wakaba, Mihama-ku, Chiba, 261-0014
Japan

Katharina Hartmann
Institut für Linguistik
Goethe Universität Frankfurt
Norbert-Wollheim-Platz 1
60629 Frankfurt am Main
Germany

K. A. Jayaseelan
63, "Srinivasa Kripa", 1st Main, 1st Cross
Thyagarajanaagar
Bengaluru 560 028
India
Jaklin Kornfilt
Syracuse University
Department of Languages, Literatures, and Linguistics
340 HBC
Syracuse, NY 13244-1160
U.S.A.

Annu Kurian Mathew
Department of Linguistics and Phonetics
School of Language Sciences
English and Foreign Languages
University
Hyderabad 500 007
India

Sergio Monforte
Department of Linguistics and Basque Studies
Faculty of Arts
University of the Basque Country UPV/EHU
Unibertsitateko Ibilbidea, 5.
01006 Vitoria-Gasteiz
Spain

Norio Nasu
Kobe City University of Foreign Studies
9-1 Gakuen-higashi-machi
Nishi-ku, Kobe-shi
651-2187
Japan

Die elektronisch verfügbaren Arbeitspapiere können auch über das Konstanzer Online-Publikations-System (KOPS) der Bibliothek der Universität Konstanz unter folgender Adresse bezogen werden:

http://kops.uni-konstanz.de/handle/123456789/45/browse?authority=uniknseries1&type=konstanzseries

75. Susanne Günthner: From Subordination to Coordination? Verb-second position in German causal and concessive constructions. Februar 1996. DM 5,–.
77. Zvi Penner: From Empty to Doubly-Filled Complementizers. A Case Study in the Acquisition of Subordination in Bernese Swiss German. Oktober 1996. DM 10,–.
79.* Klaus von Heusinger und Urs Egli (Hrsg.): Proceedings of the Konstanz Workshop „Reference and Anaphorical Relations“. November 1996. DM 13,–.
83. Björn Wiemer: Displaced speech - systematic account and acquisitional background (illustrated by Polish and German). Oktober 1996 / Februar 1997. DM 7,–.
88.* Christoph Schwarze & Aditi Lahiri: Einführung in die französische Phonologie. April 1998. DM 10,–.