The maximal size of infinitives: a truncation theory of finiteness*

Deniz Satık, Harvard University
deniz@g.harvard.edu

This paper argues for the following finiteness universal: an infinitive cannot co-occur with a high complementizer (such as *that* in English). Although such an observation may seem trivial, assuming Rizzi (1997)'s articulated CP allows one to redefine *that*. In a vein similar to Wurmbrand and Lohninger (2019), I propose that infinitives can come in different sizes. This paper combines Pesetsky (2021)'s arguments that finiteness is a matter of clause size together with truncation theories of infinitives such as Shlonsky and Soare (2011)'s to argue for a novel understanding of finiteness, proposing precise and falsifiable definitions for finite and nonfinite clauses. Beyond arguing for this finiteness universal, this paper also discusses the cartographic predictions that result from maximal size of infinitives in a given language, and concludes that infinitives can come it at least eight different sizes crosslinguistically.

Keywords: finiteness, complementizer, infinitive, clause size, left periphery

1 Introduction

One of the most poorly understood notions in generative grammar is the notion of finiteness. For descriptive grammarians, this is relatively simple: finiteness is seen as a property of the verb. As Nikolaeva (2007) points out, in Latin, the finite/nonfinite distinction was originally just the presence or absence of agreement of the verb, though other properties were later considered to be relevant for finiteness as well–the most important of which is tense.

This works straightforwardly to analyze finiteness within a European context, but such a definition of finiteness cannot be extended crosslinguistically. Landau (2013) lists a number of languages with inflected infinitives, such as Turkish, Brazilian Portuguese, Basque, Hungarian and Welsh which have nonfinite complements that are inflected for agreement. An example from European Portuguese is provided in (1) below from Raposo (1987):

(1) Será dificil [eles aprovar-em á proposta].
It will.be.difficult they to.approve-3PL the proposal
'It will be difficult [for them to approve the proposal].' European Portuguese

One, as Raposo (1987) does, may claim that agreement is not the relevant property for the finite/nonfinite distinction: instead, the distinguishing property is tense. But this does not work either. In Tamil, as McFadden and Sundaresan (2014) points out, we see the opposite scenario with

*Thanks first and foremost to Susi Wurmbrand and David Pesetsky for extensive help. I would also like to thank Jonathan Bobaljik, Jim Huang, Jay Jasanoff, Stefan Keine, Eva Klingvall, Idan Landau, Ur Shlonsky and Höskuldur Thráinsson for helpful comments. All errors are mine. This paper is dedicated to Erdoğdu Satık.

a gerundival participle in (2) below, in which a clause is embedded with tense, but lacks agreement, yet appears to be nonfinite given its inability to stand alone:

(2) Raman$_i$ [EC$_i$ Seetha-vae naaleeki paar-pp-adaagae] so-nn-aan.
Raman EC Seetha-ACC tomorrow see-FUT-GER-ACC say-PST-3MSG
'Raman$_i$ spoke of [EC$_i$ seeing Seetha tomorrow].' Tamil

Another property that has been commonly assumed to distinguish finite clauses from nonfinite clauses is whether the clause licenses overt subjects, such as by Chomsky (1977a). For example, A-movement out of a finite clause is not possible, as in (3a), but it is from a nonfinite clause, as in (3b). Although in the past such a distinction was tied to Case and agreement, in more recent proposals such as by Pesetsky (2021) it is tied to clause size: (3a) involves a clause as large as CP, which precludes the possibility of subject extraction, whereas (3b) involves a clause that is smaller than CP, which allows the possibility of subject extraction:

(3) a. * David$_i$ seems [that t$_i$ likes exfoliation].
b. David$_i$ seems [t$_i$ to like exfoliation].

Relating finiteness to clause size predates Pesetsky's work. Bouchard (1984), Koster (1984) and Hornstein and Lightfoot (1987) all argue that object extraction correlates with the size of the embedded clause; more recently, Müller (2020) has proposed a similar theory to Pesetsky's. This paper defends this theory, but with more fine-grained distinctions than just CP and TP.

This observation helps us with languages like Mandarin which have no inflectional morphology whatsoever, and hence, no tense and agreement. As has been noted by many in the literature on Mandarin, clausal complements of verbs such as *like* cannot have an overt subject or a null pronoun that does not refer to the matrix subject. In other words, we seem to observe a controlled PRO in the complements of such sentences, as in (4) from Ussery et al. (2016) below. This indicates there might be a finite/nonfinite distinction in Mandarin after all:

(4) Xiaoming xihaun (*ta) chi shousi.
Xiaoming like he eat sushi
'Xiaoming likes to eat sushi.' Mandarin

McFadden and Sundaresan (2014) raises further challenges for this line of reasoning, however, based on evidence from languages such as Tamil, Sinhala, Modern Irish and Middle English which have clauses that are clearly nonfinite–that lack tense and agreement–yet allow subjects to be licensed, as in the Modern Irish example in (5) below.

(5) Ghoillfeadh se orm [tu me a ionsai].
would.bother it on.me you.ACC me INF attack
'It would bother me for you to attack me.' Irish

As Raposo (1987) points out, even inflected infinitives in European Portuguese allow overt pronominal subjects–which Raposo ties to agreement. Regardless, McFadden and Sundaresan undermine the correlation between subject licensing and finiteness, not just for simpler models of subject licensing via Agreement in the GB and Minimalist framework like Raposo (1987)'s, but also for Landau (2004) and Szabolcsi (2009), who assume a more complex relationship between tense, agreement and subject licensing in clauses.

Another potential distinguishing property, briefly alluded to above in our discussion of Tamil, is the ability of a clause to stand alone. This seems difficult to reconcile with the existence of

imperatives like *Catch her!* which, even in languages with very rich inflectional morphology, have little inflection, and yet can stand alone. Therefore, although I have simplified the empirical terrain somewhat, many works, such as Nikolaeva (2007), have concluded that there is no single morphosyntactic definition or single semantic function associated with finiteness.[1]

Although I agree with this conclusion, I will argue that there is a single syntactic property that nonfinite clauses crosslinguistically have in common, providing further evidence for Pesetsky (2021)'s presupposition that finiteness is a matter of clause size. In this paper, I will propose that there is in fact at least one specific clausal projection which all nonfinite clauses lack. In particular, I would like to bring the attention of the reader to a seemingly trivial fact: an infinitival clause can never co-occur with *that*, which is often referred to as a finite complementizer:

(6) Caitlin seems (*that) to be pretty.

I will argue that (6) is true of all nonfinite clauses. This is, so far, a trivial observation: a finite complementizer cannot head a nonfinite clause. But this merely means that we ought to sharpen what we mean by a *finite complementizer*. Why are finite complementizers only associated with finite embedded clauses? We can answer these questions if we adopt works which split up the CP domain following Rizzi (1997), to change our conception of what *that* actually is. And this will allow us to bypass this circularity and make a non-trivial crosslinguistic generalization.

Following Rizzi (1997), I split up the C domain in a manner which is schematized below. Further details will be provided in section 3 of this paper, but I will first note that I have eliminated Rizzi's labels of ForceP and FinP, and replaced them simply with CP2 and CP1. As we will see, this splitting-up is justified by the possibility of double complementizer constructions crosslinguistically, and the existence of complementizers which seem higher and lower in the C domain:

(7) **CP2 (high)** > IntP > FocP > TopP > WhP > **CP1 (low)** > PropP > TP

I define a *high complementizer* as a complementizer that heads CP2. It precedes topics and focus-marked elements. I argue that high complementizers never appear with nonfinite clauses. A complementizer that heads CP1, on the other hand, is a *low complementizer*. It often appears with nonfinite clauses, but it need not. It follows topics and focus-marked elements, but only if the TopP and FocP layers have not already been truncated.

It is in fact possible to distinguish between these complementizers even in English: I will uncontroversially claim that *that* is a high complementizer. *For* may be a low complementizer. Although many such tests will be presented throughout this paper, I will provide a simple illustrative example. For example, notice that, as Haegeman (2012) points out, topicalization is possible in the embedded clause complements of non-factives, and in this case *that* precedes the topic:

(8) I said that Manufacturing Consent$_i$, Chomsky wrote t$_i$.

That is a high complementizer in Rizzi's system. On the other hand, infinitives in English never allow topicalization or focalization, indicating that infinitives are truncated in the C domain.

[1]Many morphosyntactic categories have been suggested to be responsible for finiteness in the literature: mood, tense, aspect, person marking, illocutionary force, nominal morphology on the verb, and markings that mark dependent clauses in certain languages. Given that a full discussion of these properties would take us out of the scope of the paper, the reader is referred to Nikolaeva (2007) for further discussion. Works like Wurmbrand et al. (2020) claim that different morphosyntactic categories are responsible for finiteness in different languages–such as agreement in the South Slavic languages.

(9) * Chomsky claimed Manufacturing Consent$_i$, to have written.

Languages differ in this regard. For example, Hungarian and Hebrew infinitives seem to display almost the entire range of the properties of the C domain, allowing *why*-embedding, topicalization, focalization and more, according to Shlonsky (2014):

(10) ani roce [et ugat ha pereg]$_i$ lenasot t$_i$.
I want DOM cake the poppyseed to.try
'I want to try the poppyseed cake.' Hebrew

And yet, Hebrew infinitives crucially cannot be headed by the high complementizer *še*:

(11) ani roce (*še) lenasot et ugat ha tapuxim.
I want (*that) to.try DOM cake the apples
'I want to try the apple cake.' Hebrew

More revealingly, there are languages–at least Icelandic, Mandarin, Serbian and Spanish–which have elements that behave as high complementizers in finite clauses, but cannot behave as a high complementizer in nonfinite complements. For example, what has been called the infinitival marker in Icelandic, *að*, appears only with control complements. But another element, *að*, behaves like *that*, appearing with finite embedded clauses. They have different properties: infinitival *að* does not allow topicalization, while finite *að* does.

(12) a. * Risarnir lofa [að [á morgun]$_i$ éta ríkisstjórnina t$_i$].
the-giants promise to to-morrow eat the-government
'The giants promised to eat the government tomorrow.'

b. Risarnir segja [að [á morgun]$_i$ éti þeir ríkisstjórnina t$_i$].
the-giants say that tomorrow eat they the-government
'The giants said that they will eat the government tomorrow.' Thraínsson (1993)

Therefore, the main crosslinguistic generalization that I will argue in this paper, is in (13) below.[2]

(13) **Infinitive Size Generalization (ISG)**: No infinitive projects CP2.
No infinitive can co-occur with a high complementizer.

This allows for a definition of finiteness in terms of the truncation of the C domain, and hence clause size. My goal in this paper is to investigate the clause size of infinitives more generally, and see whether generalizations beyond (13) above can be made. I also discuss the consequences of defining finiteness in terms of clause size.

This paper is structured as follows. Section 2 introduces the reader to Rizzi (1997)'s structure for the C domain. Section 3 provides the empirical background for this paper, and I discuss the generalizations that result from it. Section 4 discusses the theoretical consequences of my theory of finiteness: consequences for Pesetsky's Exfoliation framework, and potential counterexamples to my definitions of finiteness. Section 5 concludes.

[2]The scope of this paper is to cover only the clausal size of infinitives; as such, I will only focus only on infinitives. For the most part, I will not discuss gerunds or other kinds of nonfinite clauses in this paper, leaving it to future research. However, it is likely that conclusions that I make concerning infinitives can also be made concerning gerunds as well. I will also leave subjunctives open to future research.

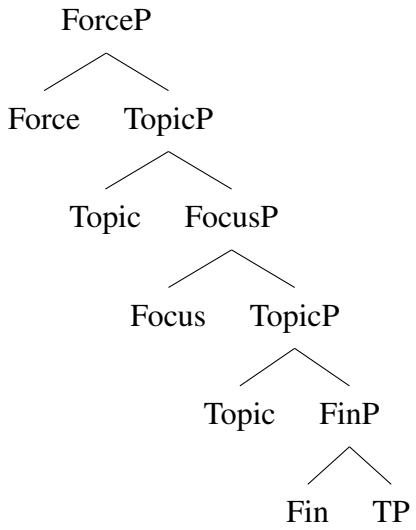
2 Splitting up the C domain

This section will lay the foundation for the theory of finiteness that I propose in this paper: namely that finiteness itself is a property of the C domain. I present Rizzi (1997)'s arguments in favor of splitting up the C domain into many (and potentially ordered, crosslinguistically) different functional projections. I provide evidence for there being high and low complementizers–two separate complementizers–in the C domain. I discuss existing accounts of the truncation of infinitives. At the end, I also provide my update to Rizzi's structure, changing the labels of Rizzi's ForceP and FinP. This provides the background for getting rid of FinP: I argue instead that finiteness itself can be derived via truncation.

2.1 Rizzi (1997)'s split-CP structure

Rizzi (1997) provides arguments for splitting up the C domain as follows in (14). If we had just one C projection–CP, as is commonly assumed–it would be impossible for a single projection to be responsible for all of these properties that I will discuss in this section.[3]

(14)



Rizzi (1997) argues that two complementizers in Italian, *che* and *di*, are realized by Force and Fin respectively. ForceP is the locus of the semantic force of the clause (such as an assertion, a question or an imperative). FinP, on the other hand, simply encodes whether the clause is finite or not. Under Rizzi's account but not mine, finiteness is to be understood as a very rudimentary specification of mood, tense and agreement in the IP domain. Fin itself does not have a semantics but it is endowed with certain features that allow this aforementioned specification to take place.

[3]This raises the interesting question of what exactly is a phase in this structure. This is at odds with Chomsky (2001) to some degree given that there are many potential phase candidates but it is not obvious which one is the phase head. At the very least, I assume that ForceP–the highest projection of the C domain–is a phase head. Given that wh-movement takes place to a position right above FinP, as I will argue later in this next section, and successive cyclic wh-movement, it might be assumed that FinP is a phase as well. But this is at odds with Carstens and Diercks (2009)'s observations of FinP never being phasal in Lubukusu. Regardless, apart from the phasehood of ForceP, it is out of the scope of this paper to determine what potential phase heads in this structure are.

Topic and Focus, on the other hand, are projections with an independent semantics of their own, and their specifier position is for topicalized and focalized DPs respectively. There is a difference between focalization and topicalization: they can be teased apart by using different contexts. Rizzi contrasts between these two in Italian: while (15a) involves Clitic Left Dislocation (CLLD), (15b) involves focus fronting in a context with contrastive focus:

(15) a. Il tuo libro, lo ho letto.
the your book, it I have.read
'Your book, I have read it.'

b. Il tuo libro ho letto.
the your book I have.read
'Your book I have read.' (but not his)

Furthermore, TopicP in (14) is *recursive*, in that it can appear both before or after FocusP–or before or after other projections between ForceP and FinP; it is commonly assumed that there are. Rizzi provides evidence from this in Italian, which we need not go into; in this paper, I will assume for simplicity that FocusP is always ordered above TopicP.

2.2 What are high and low complementizers?

This sets the stage to allow us to distinguish between *high* and *low* complementizers, which are complementizers realized at Force (my CP2) and Fin (my CP1) respectively. Rizzi was the first to note this contrast, which will be essential for the theory of finiteness in this paper. We see in (16) below that it is impossible to topicalize to a position to the left of the high complementizer *che* (which Rizzi calls a finite complementizer), but it is possible to topicalize to its right.

(16) a. Credo che, il tuo libro, loro lo apprezzerebbero molto.
I.think that[+fin] the your book them it will.appreciate much
'I think that they will appreciate your book very much.'

b. * Credo, il tuo libro, che loro lo apprezzerebbero molto. Italian

This contrasts with the behavior of the low complementizer *di* (which Rizzi calls a nonfinite complementizer), which only allows topicalization to its right in (17):

(17) a. Credo, il tuo libro, di apprezzar-lo molto.
I.think the your book that[-fin] appreciate-it much
'I think that they will appreciate your book very much.'

b. * Credo di, il tuo libro, apprezzar-lo molto. Italian

This indicates that *di* in Italian cannot be in the same position as *che*: but if *di* is a low complementizer in FinP whereas *che* is a high complementizer in *ForceP*, these facts would immediately be explained. Some languages like Spanish even allow double complementizer constructions:

(18) Susi dice **que** *a los alumnos* (**que**) les van a dar regalos
Susi says that DAT the students that cl. go to give presents
'Susi says that they are going to give the students presents.' Spanish

There is a great deal of evidence of high and low complementizers, and even double complementizer constructions even outside of Romance. Even in English, Haegeman (2012) notes two such examples below. Because *that* never behaves as a low complementizer alone, It appears that *that* in FinP can only be licensed if *that* is also realized in ForceP:

(19) a. She maintained **that** when they arrived **that** they would be welcomed.

b. He reminds me **that** in the days of Lloyd George **that** business leaders were frequently buying their way in.

Larsson (2017) provides a survey of double complementizer constructions across the Scandinavian languages, providing an example from Icelandic, from Thráinsson (2007) below. *Sem* is a relative complementizer, and *að* can follow it. It appears that *að* is in CP1 in this case:[4]

(20) þetta er bokin sem (að) eg keypti
This is book.DEF that that I bought
'This is the book that I bought.' Icelandic

I conclude this subsection with evidence that some Bantu languages distinguish between a high, phasal complementizer and a low, non-phasal complementizer. Carstens and Diercks (2009) shows that in Lubukusu, some clauses are transparent for hyperraising, which is raising out of a finite clause, while others are not transparent for it. Here are some examples from Lubukusu, where what they call hyperraising is possible with the complementizer *mbo*:

(21) Mikaeli a-lolekhana **mbo** a-si-kona.
Michael 1SA-seem that 1SA-PRES-sleep
'Michael seems to still be sleeping.' Lubukusu

But this raising is not possible with the complementizer *-li* which agrees with the matrix subject:

(22) * Mikaeli a-lolekhana **a-li** a-si-kona.
Michael 1SA-seem 1CA-that 1SA-PRES-sleep
'Michael seems to still be sleeping.' Lubukusu

Under this analysis, *mbo* is the low, non-phasal complementizer, and *-li* is the high, phasal complementizer. With this difference established, we now move onto infinitives.

2.3 Infinitives are truncated in the C domain

Adger (2007) notes a contrast between English and Italian that we will build further upon in section 3.1: topicalization is not allowed at all in English infinitives (Hooper and Thompson (1973)):

(23) * I decided, [your book]$_i$, to read t_i.

Adger also notes that the complementizer *for* in English rejects topics. As Adger suggests, I agree with him that this indicates that *for* is a low complementizer in Fin:

(24) * I propose, [these books]$_i$, for John to read t_i

Following Adger among others such as Haegeman (2006), Barrie (2007) and Shlonsky and Soare (2011), I also take this to be evidence that infinitives are truncated: as we will see, this truncation can differ between languages like English and Italian.

[4] Icelandic allows infinitival relatives but they cannot contain *sem*; instead they have the preposition *til*:

(i) Þetta er bón [til að bóna bíla með _].
this is wax for to polish cars with
'This is wax to polish cars with.'

Höskuldur Thráinsson (p.c.) has pointed out to me that *til* behaves as a preposition in such constructions rather than a complementizer, based on the fact that the genitive form of *það* 'it,' *þess*, can be inserted between *til* and *að*.

There is reason to believe that there are many more projections than what Rizzi (1997) has initially claimed, and the number of functional projections has indeed increased in works since then such as Haegeman (2012). For our purposes, I will present only the additional projections which are relevant to infinitives–IntP and WhP in particular.

The layer IntP is short for InterrogativeP, which according to Rizzi (2001) is higher than FocusP: Spec,IntP houses *why*. Shlonsky and Soare (2011) provides a convincing argument that *why* is base-generated in position lower than Spec,IntP and moves up to it, in the form of infinitives. Note that the infinitive form is very marginal at best, but the finite form is fine:[5]

(25) a. ?? I asked Bill why to serve aubergines.
b. I asked Bill why I should serve aubergines.

Given that we have already seen that TopicP is truncated in English infinitives, it is unsurprising that a functional projection ordered even higher is truncated as well.

Let us move to WhP. The fact that focalization is impossible with English infinitives whereas wh-infinitives in English do exist, ex. *I know what to eat*, is not expected under Rizzi's original account, where all wh-words move to Spec,FocP. As such, Barrie (2007) and Shlonsky and Soare (2011) have assumed the addition of a further functional projection on top of FinP, WhP, which wh-elements first move into prior to moving to Spec,FocP.[6] Even in a language where fronted focus is possible such as Italian, which also has wh-infinitives, Haegeman (2006) and Bocci (2007) note that focalization is very marginal:

(26) ?? Gli sembra le sedie di aver venduto (, non il tappeto)!
To him-seems the chairs to have sold (, not the carpet)
'It seems to him that the chairs have sold! (not the carpet).' Italian

Infinitives, as I will argue in section 3, are *always* truncated in the C domain. If this is true, this indicates there is a relationship between clause size and finiteness. As such, there is redundancy between the notion that infinitives are truncated, and the theoretical tool of FinP to begin with. What seems more reasonable is that these infinitives are nonfinite *because* they are truncated, and this makes sense if finiteness is a matter of clause size.

There are more general problems with Rizzi's definition of finiteness, as well. It is circular, in that whether a clause is finite iff its finiteness feature is encoded as + at FinP, following Adger

[5]Although it is not relevant for our purposes, Shlonsky and Soare (2011)'s argument that it is base-generated lower is as follows. The following question can be construed in two ways: one in which *why* is construed within the matrix clause, and one in the embedded infinitival clause:

(i) Why did you ask her to resign?
a. What is the reason x, such that for x, you asked her to resign?
b. What is the reason x, such that you asked her to resign for that particular reason x?

[6]It seems that there is a WhP on top of ForceP as well. Henry (1995) notes that Belfast English permits indirect questions introduced by a wh-element that isn't a subject, to the left of the high complementizer *that*:

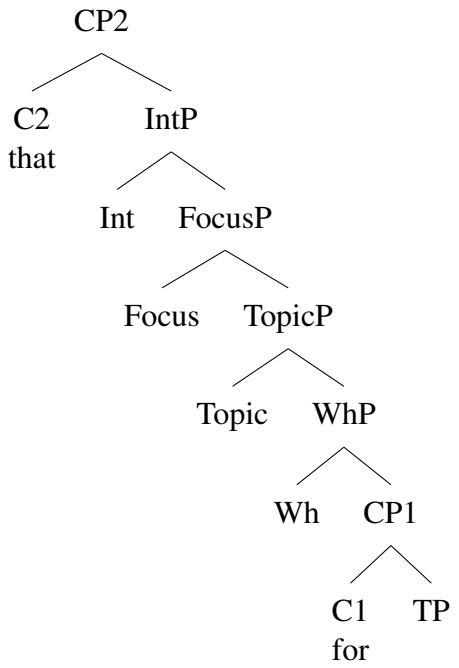
(i) I wonder which dish that they picked.

This seems to be very common crosslinguistically; Larsson (2017) notes that several Scandinavian languages allow such constructions. At this point, an obvious question to be asking is why there isn't yet another FocusP, TopicP, IntP etc. on top of CP2 as well. But it simply seems to be the case that this is not empirically attested. So this does not put my definition of a high complementizer in jeopardy.

(2007). It may be possible to define finiteness in terms of other features, such as past tense, agreement and indicative mood on FinP. But even then, it is not a fully explanatory theory of finiteness: truncation theories of finiteness in terms of clause size do make testable predictions.

Thus, I propose getting rid of the labels of ForceP and FinP and replacing them simply with CP2 and CP1 respectively:

(27)



This is what I hope to be the novel idea of the paper. While I am far from the first to assume that infinitives are truncated, I am synthesizing the approach to finiteness as a matter of clause size together with Rizzi's work on the split C domain. This allows us to bypass any circularity or redundancy in defining finiteness, and thus make novel empirical observations in the next section.

3 The size of infinitives and empirical generalizations

I present a crosslinguistic survey of infinitive sizes in 3.1. In 3.2, I give evidence from four languages that distinguish between high and low complementizers (or lower clausal heads), very similarly to what Rizzi (1997) noted in Italian above, but these are with elements with the exact same phonetic form instead. This evidence, I believe, shows a fundamental inability for nonfinite clauses to co-occur with high complementizers. Section 3.3 provides further cartographic generalizations concerning the order in (27). Section 3.4 discusses the findings of the section.

3.1 Infinitives can differ in size, but are always truncated

In this subsection, I have two goals. First, I provide a survey of 22 languages in which the maximal infinitive size is not CP2–that is, they cannot co-occur with high complementizers. Second, in 19 out of these 22 languages, I also determine the maximal size for infinitives. The final list that we will be left with is as follows in (28). It appears that TP is the *minimal maximal size* for

infinitives crosslinguistically, whereas IntP, and crucially not CP2, is the *maximal maximal size* attested. I will argue that this observation is the key to understanding finiteness:[7]

(28) a. **Hierarchy**: CP2 > IntP > FocP > TopP > WhP > CP1 > PropP > TP
b. **Maximally TP Infinitives**: Turkish, Serbian, Hindi, Bangla, *Jordanian Arabic
c. **Maximally PropP Infinitives**: German
d. **Maximally CP1 Infinitives**: Icelandic, Swedish, Norwegian
e. **Maximally WhP Infinitives**: English, Spanish, French, European Portuguese, Dutch, *Mandarin
f. **Maximally TopP Infinitives**: Italian, Catalan
g. **Maximally IntP Infinitives**: *Serbian$_{\mathrm{DA}}$, Hungarian, Hebrew
h. **Maximally CP2 Infinitives**: $\emptyset$
i. **Unspecified but not CP2**: Middle English, Old Norse, Old Swedish

Let us see how the empirical tests in this section will work. I now provide a quick summary of the properties of the C domain of English infinitives:[8]

(29) a. *Infinitival complementizers*: I am eager for Caitlin to please.
b. *Wh-infinitives*: I know what to eat.
c. *No topicalization within infinitives*: *I wanted this book, to read.
d. *No focalization within infinitives*: *I wanted THIS BOOK to read (not that one).
e. *No why-infinitives*: ??I asked Caitlin why to eat salad.
f. *No high complementizer*: I seem (*that) to be happy.

This indicates that English infinitives are *maximally* as large as WhP. The maximal size of an infinitive is the most crucial notion of this paper: Languages appear to vary as to the maximal size of their infinitive, and there are at least six different maximal sizes which are attested. I will start from the languages that can have infinitives as large as IntP.

[7]Languages or constructions preceded by a * are those whose status of finiteness is controversial; ex. whether Mandarin has a finiteness contrast is unclear. Unmarked languages are not controversial.

[8]One puzzle is the difference between *whether* and *if* in infinitives. These words are often interchangeable, ex. *I asked my mom whether/if I should take out the trash.* But only *whether* is permitted in infinitives:

(i) a. Caitlin asked whether to take out the trash.
b. * Caitlin asked if to take out the trash.

Following Shlonsky and Soare (2011), one explanation is to suppose that *whether* can be Merged in either Spec,WhP or Spec,IntP, whereas *if* must be Merged in Spec,IntP. It then follows that only *whether* can be licensed in infinitives. But it does not appear to be possible to assume that *whether* is always Merged in Spec,WhP, as Jonathan Bobaljik (p.c.) has pointed out to me with the following contrast. Adjuncts which precede the embedded subject must follow *whether* in finite embedded clauses. Adjuncts in the C domain will be discussed further in 5.1.

(ii) Caitlin asked whether under any circumstances she should leave.

(iii) * Caitlin asked whether under any circumstances to leave.

3.1.1 IntP Infinitives

Shlonsky (2014) notes that Hebrew infinitives appear to be almost untruncated in the C-domain, allowing focalization and even why-infinitives, as shown in (30a)-(30b) below:

(30) a. ani roce [et ugat ha pereg]$_i$ lenasot t_i (lo et ugat ha tapuxim).
I want DOM cake the poppyseed to.try (not DOM cake the apples)
'I want to try the poppyseed cake (not the apple cake).'

b. ani lo mevin lama la'avor dira.
I not understand why to move apartment
'I don't understand why to move apartments.' Hebrew

But there seems to be at least one property which its C domain lacks: the ability to co-occur with the high complementizer *še*. I conclude that Hebrew infinitives may be as large as IntP.

(31) ani roce (*še) lenasot et ugat ha tapuxim.
I want (*that) to.try DOM cake the apples
'I want to try the apple cake.' Hebrew

In Hebrew, given the presence of *why*-infinitives, we would predict all of the properties below IntP to be present as well, if there is indeed a cartographic hierarchy. Shlonsky shows almost all of these in Hebrew, with the exception of the infinitival complementizer. According to Landau (2013), Hebrew has the dedicated complementizer *me-*, appearing with control infinitives.

Hebrew is not alone in having infinitives with a nearly complete left periphery. Szécsényi (2009) reports that this is also the case in Hungarian, allowing infinitives with topicalization and focalization. I have verified independently that Hungarian also allows *why*-infinitives below, like Hebrew. Hungarian infinitives do not allow the presence of finite, or high, complementizer *hogy*.

(32) John meg kérdezte, minek en-ni.
John VM asked why eat-INF
'John asked why to eat.' Hungarian

Moving onto Serbian, whether its *da*-constructions are nonfinite is controversial. But in section 3.2 I will discuss the differences between Serbian infinitives and *da*-constructions in greater detail, arguing that the latter are in fact nonfinite. I have verified that Serbian *da*-constructions also allow an embedded *why* preceding *da*. If the clause is finite, *why* must follow *da*. On the other hand, I have verified Serbian infinitives appear to lack the C domain entirely, not allowing wh-elements inside them, or topicalized elements, or *why*.

(33) a. Pitao sam zašto da jedem. b. * Ne znam šta jesti.
asked.SG.M AUX.1SG why DA eat.1SG NEG eat.1SG what eat.INF
'I asked why to eat.' 'I don't know what to eat.' Serbian

I will hence classify Serbian infinitives as maximally TP.[9]

[9]Languages with focalized elements inside infinitives but not *why* does not seem to be attested crosslinguistically. I am unable to answer why this is the case at this time: perhaps the sample in the paper is not large enough, or alternatively, there should not be a split between IntP and FocP.

3.1.2 TopP Infinitives

Some languages, such as Italian and Catalan, have infinitives that are larger than WhP but smaller than ones of Hebrew and Hungarian. Although fronted focus in Italian infinitives is borderline unacceptable, repeated in (34a), CLLD is in fact completely acceptable in (34b):

(34) a. ?? Gli sembra le sedie di aver venduto (, non il tappeto)!
To him-seems the chairs to have sold (, not the carpet)
'It seems to him that the chairs have sold! (not the carpet).'

b. Gli sembra, il tappeto, di averlo venduto.
To him-seems, the carpet, to have-it sold
'It seems to him that the carpet has sold.' Italian

We have already seen in (17) that Italian has a low complementizer *di*. Given the ordering TopP > WhP, wh-infinitives should exist in Italian. According to Kayne (1981), they do, as seen below. I conclude that Italian infinitives can be slightly larger than English ones, or as large as TopP:

(35) Gli ho detto [dove andare].
Him I told [where go.INF]
'I told him where to go.' Italian

Catalan also allows CLLD and wh-elements inside infinitives, according to Villalba (2009), but Spanish, French and European Portuguese actually don't allow CLLD inside infinitives.

Further, although it appears to no longer be attested in Germanic today, Faarlund (2015) and Kalm (2016) claim that arguments can precede both the infinitival marker and the verb in Old Norse and Old Swedish respectively. I provide an illustrative example from Faarlund (2015) below. As in languages like Icelandic, *at* appears to be the phonetic form for both the infinitival marker and finite complementizer in Old Norse and Old Swedish.

(36) ek hafða nú ætlat [sex skip ór landi]$_i$ at hafa t$_i$
I had now intended six ships from country to have
'I had now intended to take six ships out of the country.' Old Norse

This leads these authors to reject that *at* in these contexts is a complementizer. But if we treat *at* as a low complementizer inside infinitives rather than a high one, as I will argue in section 3.2, this conclusion will not be necessary. However, given the lack of solid data–such as the possibility of wh-infinitives–and impossibility of further investigation, I have not classified the languages together with Italian and Catalan. But given that the infinitives show the inability for *at* to be a high complementizer–that is, obligatorily precede topicalized elements–I have classified Old Norse and Old Swedish as languages without CP2 infinitives.

3.1.3 WhP Infinitives

WhP is the case in English, as seen in (29) prior, lacking *why*-infinitives, topicalization or focalization in infinitives. Dutch is the only Germanic companion to English among my sample, also allowing wh-infinitives (from Wheelock (2015)):

(37) Ik weet niet [wie te bezoeken].
I know not [who to visit.INF]
'I do not know who to visit.' Dutch

This is also the case in other Romance languages with CLLD, in fact: Rizzi (1997) reports that the Italian facts in (34a)-(34b) are not acceptable in French. Villalba (2009) and Barbosa (2001) also report that Spanish and European Portuguese respectively pattern with French, rather than Italian or Catalan, in not allowing CLLD within infinitives.

Although I will save further discussion of finiteness in Mandarin–which is controversial–to section 3.2, Ussery et al. (2016) note that both control and raising complements in Mandarin allow wh-elements, while only control complements in English allow wh-elements. This leads me to classify Mandarin as a WhP language as well, even though it does not have infinitives, because it appears to have a finite-nonfinite distinction, as mentioned in section 1 above.[10] In (38a)-(38b) below, *how* inside a raising complement is not acceptable in English, but it is in Mandarin:[11]

(38) a. * Alex seems how to have gotten fat.

b. Xixi kanqilai zenme zhangpang le.
Xixi seem how grow fat ASP
(lit.) 'Xixi seems how to have gotten fat.' Mandarin

3.1.4 TP Infinitives

With the size of English infinitives established, let us move onto TP languages prior to considering CP1 languages. TP, as far as I am aware of, represents the minimal maximal size for infinitives based on my crosslinguistic survey.

Keine (2020), based on tests from Wurmbrand (2001) among others, provides convincing arguments that Hindi nonfinite complements are smaller than English infinitives. For example, the wh-element *kyaa* 'what' can take scope within the finite embedded clause, as in (39a). But it cannot take embedded scope inside the infinitive, as shown in (39b). Keine reports that the sentence is acceptable as long as the wh-element takes matrix scope ("what do you know to do"):

(39) a. tum jaan-te ho [(ki) us-ne kyaa ki-yaa]
you know-IPFV.M.PL be.PRES.2PL that he-ERG what do-PFV.M.SG
'You know what he did.'

b. * tumhe [kyaa kar-naa] aa-taa hai
you.DAT what do-INF.M.SG come-IPFV.M.SG be.PRES.3SG
'(Intended) You know what to do.' Hindi

But it is not only WhP that is missing. Keine also makes a stronger claim: that the C domain is entirely missing in Hindi infinitives. (40) below involves illicit A'-movement from a doubly embedded finite clause into the embedded infinitive. A'-movement is allowed from finite clauses, so the reason why (40) is ruled out, for Keine, is because the infinitive lacks a C domain:

[10] We will see in section 3.2 that Mandarin does not allow external topicalization to the left periphery in infinitives, like languages like Italian or Hebrew, but merely topicalization to a verb-medial topic or focus position. This is why I have chosen to classify Mandarin as a WhP language rather than, say, a TopP or IntP language.

[11] It appears that, following a derivational approach like Pesetsky (2021), this indicates that raising and control complements have slightly different operations in English, given that the former can only be as large as TP while the latter can be as large as WhP. Though I leave open the details of this solution.

(40) * [mai caah-taa huu [kitaab-ko kah-naa [ki mai-ne parh-aa
I want-IPFV.MSG be.PRES.1SG book-ACC say-INF.M.SG that I-ERG read-PFV.M.SG
hai]]]
be.PRES.3SG
'(Intended) I want the book, to say that I read.' Hindi

Like English, Hindi does not allow a high complementizer to co-occur with the infinitive:

(41) siitaa [(*ki) prataap-ko dekh-naa] caah-tii thii
Sita that Pratap-ACC see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
'Sita wanted to see Pratap.' Hindi

Based on these data, Keine concludes, as I do, that Hindi infinitives are only as large as TP. Dasgupta (1982) reports that Bangla, another Indo-Aryan language, patterns with Hindi in that it lacks wh-infinitives and the complementizer *je* cannot co-occur with infinitives. This is despite the fact that according to Hsu (2015), *je* in Bangla can occur in either ForceP (my CP2) or FinP (my CP1), indicating further that it is genuinely a TP language.

Kornfilt (1996) also reports that although Turkish has infinitives, it lacks infinitival wh-questions. My own judgment is that infinitival complementizers are not present in the language, either. An illustrative example is given below:

(42) * Ahmet Ayşe-ye [PRO ne oku-mak] söyle-di.
Ahmet Ayşe-DAT what read-INF say-PST
'Ahmet told Ayşe what to read.' Turkish

A particularly interesting case of a language that appears to have a finiteness contrast despite not having traditional infinitives is in Jordanian Arabic. Al-Aqarbeh (2011) argues that finite complements are those which project a C domain, and nonfinite complements are those which do not project a C domain. Two illustrative examples are given, in which the complement clause cannot have a complementizer or an embedded topicalized element:

(43) a. 9ali bid-uh (*innu) il-banaat yi-l9ab-an.
Ali want-3SG.M (*that) the-girls 3-play-PLF
'Ali wants the girls to play.'

b. * 9ali bid-uh il-ghurfah il-banaat yi-naththif-an-ha.
Ali want-3SG.M the-room the-girls 3-clean-PLF-it
'Ali wants the girls to clean the room.' Jordanian Arabic

On the other hand, a complementizer and embedded topicalized elements may appear with propositional complement clauses. Although it would be out of the scope of this paper to discuss in detail the semantics of the complements are nonfinite and those which are finite, nonfinite complements cannot have propositional semantics, at least in Arabic. Hence, Al-Aqarbeh (2011) relates the presence of the C domain to finiteness in Jordanian Arabic.

3.1.5 PropP languages

The languages that I have classified as TP languages are those which appear to lack propositional semantics in their infinitives entirely, along with all properties of the C domain, as in Jordanian

Arabic. For example, according to Bhatt (2006), Hindi lacks raising predicates like that of English entirely. German appears to be the only language in my sample which appears to lack an infinitival complementizer–following Sabel (2006) among others–but has raising constructions with propositional semantics, as in (44a). Wh-infinitives are impossible in German, as shown in Wheelock (2015)'s example (44b).

(44) a. Er scheint [intelligent zu sein].
he seems intelligent to be
'He seems to be intelligent.'

b. * Ich weiß nicht [was zu kaufen].
I know not [what to buy.INF]
'I do not know what to buy.' German

For Keine, German infinitives are the same size as Hindi's, but he does not discuss the lack of propositional infinitives in Hindi. I believe that this indicates the need to distinguish between German on one hand, and languages like Hindi on the other. I follow Wurmbrand and Lohninger (2019) in assuming that the propositional semantics is a part of the C domain, but belonging to a head lower than CP1, which I call PropP. Although I am unable to improve on this stipulation at this time, this allows us to straightforwardly capture the difference between German and Hindi.

3.1.6 CP1 Infinitives

There appears to be a final size in between WhP and PropP/TP, which I have saved for last given that it is likely to be the most controversial. The most difficult maximal size to determine is that of languages like Icelandic, Swedish and Norwegian, because their finite complementizers share the same phonetic form as the so-called infinitival marker. This is unlike that of German: *zu* is not the phonetic form of the finite complementizer *dass*. But I would like to provide an analysis of languages like Icelandic in which this phonetic similarity is not a mere coincidence.

As mentioned in section 1, *að* seems to come in two different varieties: finite clauses allow embedded topicalization to the right of *að*, whereas the *að* found in control infinitives does not allow topicalization. It appears, then, that TopP in Icelandic infinitives is always truncated. Data from Thraínsson (1993) is repeated in (45a)-(45b) below.

(45) a. * Risarnir lofa [að [á morgun]$_i$ éta ríkisstjórnina t_i].
the-giants promise to to-morrow eat the-government
'The giants promised to eat the government tomorrow.'

b. Risarnir segja [að [á morgun]$_i$ éti þeir ríkisstjórnina t_i].
the-giants say that tomorrow eat they the-government
'The giants said that they will eat the government tomorrow.' Icelandic

Faarlund (2015) points out that argument preposing of this kind is not possible in Norwegian embedded infinitives, either, while it is in Old Norse. Furthermore, Icelandic lacks wh-infinitives, as Sabel (2006) points out, indicating the absence of WhP.

Thraínsson (1993) took (45a)-(45b) as evidence that *að* in AgrSP–right above TP but below CP. Further evidence of this is the fact that Icelandic has V-to-T (V-to-I in older frameworks) movement in infinitives, unlike English, and the verb still occurs after *að*. I believe the first to note this was Sigurðsson (1989). Notice that in (46a), the movement of the auxiliary to T precludes the movement of the embedded verb to T, but this is not the case in (46b), and it does move to T. In control infinitives, V to T is still possible and it is to the right of *að*, as in (46c).

(46) a. Risarnir segja [að þeir hafi stundum [$_{VP}$ étið ríkisstjórnir].
the-giants say that they have sometimes eaten governments
'The giants say that they have sometimes eaten governments.'

b. Risarnir segja [að þeir éti$_i$ stundum [$_{VP}$ t$_i$ ríkisstjórnir]].
the-giants say that they eat sometimes governments
'The giants say that they sometimes eat governments.'

c. Risarnir lofa [að éta$_i$ oft [$_{VP}$ t$_i$ ríkisstjórnir]].
the-giants promise to eat frequently governments
'The giants promised to eat governments frequently.' Icelandic

The evidence that *að* is above TP seems to be strong. Assuming that it is in AgrSP would not contradict the goal of this paper. However, with Rizzi's split-CP structure, we do not need to give up the idea that *að* in Icelandic is *always* a complementizer–it could simply be a low complementizer realized in CP1 if it is not first realized in CP2. This has an advantage over Thraínsson (1993)'s account of infinitival *að* in AgrSP, given that it would be mysterious as to why the two *að*–even with different properties–have the same phonetic form.

Another CP1 language appears to be Swedish. Platzack (1986) notes that the Swedish complementizer *att* is similar to that of Icelandic. We've already seen that Swedish lacks wh-infinitives; Engdahl (1986) also provides evidence that Swedish infinitives are not full CPs either, based on the inability of pied-piped material to appear in infinitival relative clauses:

(47) a. ett rum att arbata i _
a room to work in
'a room to work in.'

b. * ett rum i vilket att arbata [$_{PP}$ _]
a room in which to work
'a room in which to work.' Swedish

Finally, Wheelock (2015) notes that both Swedish and Norwegian lack wh-infinitives. I believe that the data given above imply the need to stipulate a different projection, WhP, above CP.[12]

3.1.7 Conclusion

I understand that some of the classifications may be controversial–for example, assuming a separate projection WhP for wh-movement and PropP purely for propositional semantics, without a complementizer. But ultimately, none of this contradicts the primary goal of this paper, which is to show that infinitives cannot co-occur with high complementizers. This seems to be true.

Regardless, our survey is almost complete. But it has been claimed that Middle English infinitives project ForceP. According to van Gelderen (1998), it is possible for *ai* in (48) below to be a focus marker; in which case, *til* would be in ForceP (my CP2), flatly falsifying my upcoming generalization: no infinitive projects CP2. My attempt at glossing her ideas is below:

(48) Til [all oure bale] ai for to bete
COMP all our sorrow FOC COMP to heal
'For all our sorrow to heal...' Middle English

However, according to Jay Jasanoff (p.c.), it appears that this is not a double complementizer construction. *Til* plays the role of complementizer *for* in this construction, making it as large as

[12]We might try to avoid not splitting up the C domain into WhP and CP by assuming that C has a wh-feature in English but not in, for example, Icelandic. But I believe that this would miss the upcoming cartographic generalizations in section 3.3. Further evidence for splitting up WhP and CP1 will be provided.

CP1. *Ai* is not a focus marker but rather a word that means *forever*, whereas "for to" in Middle English is itself the infinitive marker, (cf. *to* in English). When this sentence is translated with modern lexical substitutions into its syntactic structure, we obtain *for all our sorrow forever to amend*, which is not so exotic after all.

Infinitives appear to never project the full C domain; in particular, Rizzi's ForceP, or my CP2. I have shown that even in Hebrew and Hungarian, with the largest attested infinitives, infinitives cannot co-occur with the so-called high complementizer *še* in Hebrew or *hogy* in Hungarian.

Of course, one might be allege that this might simply be because finite complementizers don't select nonfinite clauses. But I believe this simply begs the question of *why* finite complementizers (in our terminology, high) do not select nonfinite clauses, and does not lead to a greater understanding of this fact. To explain this, I present a potential finiteness generalization in (49):

(49) **Infinitive Size Generalization (ISG)**: No infinitive projects CP2.
No infinitive can co-occur with a high complementizer.

But we do not yet have enough evidence to conclude that the ISG is true, of course. Absence of evidence is not evidence of absence: the fact that there does not seem to be a language reported in the literature with a high or double complementizer construction–with the exception of Middle English, which we have rejected–does not mean that we have a universal. In other words, the pattern seen in (28) is not enough to conclude that infinitives are *always* truncated, and that finiteness can be defined in terms of the presence or lack of the CP2 layer. But in the next subsection, I attempt to present evidence of absence in favor of the ISG, in which I argue that nonfinite clauses are fundamentally unable to co-occur with a high complementizer.

3.2 Languages with the same phonetic form for high complementizers and other clausal heads

This subsection presents further evidence for the generalization in (49) above. We will be investigating a specific pattern in several languages. In particular, all of these languages have an element which is uncontroversially high complementizer, corresponding to *that* in English. An element bearing the same phonetic form as the high complementizer may appear in nonfinite clauses, which might seem as a genuine counterexample to the ISG. But in a nonfinite clause, it turns out that this element has very different properties: it cannot behave as a high complementizer.

Italian is not one of these languages. But recall the data from (16)-(17) above, repeated in (50) below. It is possible to topicalize to the right of the high complementizer *che* in Italian but not to its left; it is also possible to topicalize to the left of the low complementizer *di* but not to its right:

(50) a. Credo che, il tuo libro, loro lo apprezzerebbero molto.
I.think that[+fin] the your book them it will.appreciate much
'I think that they will appreciate your book very much.'

b. * Credo, il tuo libro, che loro lo apprezzerebbero molto.

c. Credo, il tuo libro, di apprezzar-lo molto.
I.think the your book that[-fin] appreciate-it much
'I think that they will appreciate your book very much.'

d. * Credo di, il tuo libro, apprezzar-lo molto. Italian

It turns out that similar contrasts are seen crosslinguistically, even with elements that share the same phonetic form. The first of which is, of course, Icelandic, which we have already discussed in subsection 3.1 prior: it allows topicalization to its right in finite contexts as in (45a) above, but not at all in control infinitives, as in (45b). This, in my view, is because *að* cannot behave as a high complementizer in control infinitives, because CP2 is truncated.

Similar facts are seen in Norwegian and Swedish, according to Faarlund (2015) and Kalm (2016) respectively. In Old Norse, we saw in (36) prior that unlike Icelandic, Swedish and Norwegian, argument preposing inside infinitives is allowed. Although it is not clear whether this involves topicalization, it does at the very least show that *at* does not behave as a high complementizer in this context. Here is a similar example from Kalm (2016) illustrating this in Old Swedish:

(51) þa ær han skyldugher han at ola
then is he obliged him to oil
'Then he is obliged to oil him.' Old Swedish

A language similar to Icelandic in some respects is Spanish, according to Villa-Garcia (2012), for which (52) is repeated below. Villa-Garcia (2012) refers to the first bolded *que* as a high complementizer, just like *that*, whereas the lower *que* he refers to as a "jussive/optative" complementizer, which is characteristic of subjunctives. (52) shows that topicalization occurs to the right of the high variety of *que*.

(52) Susi dice **que** *a los alumnos* (**que**) les van a dar regalos
Susi says that DAT the students that cl. go to give presents
'Susi says that they are going to give the students presents.' Spanish

It seems that the complementizer *que* cannot occur in Spanish infinitives; according to Lujan (1980) a separate complementizer *de* is used instead, so the facts would not be very different from Italian. But there are other nonfinite contexts outside of infinitives in which low *que* can be used, such as imperatives. In (53) below, Demonte and Fernández-Soriano (2009) point out that the topic *a ese alumno* 'to that student' moves to the left of *que*. They analyze the two *que* precisely as I and Villa-Garcia do: *que* comes as both a high and a low complementizer, and the low variety is only present in nonfinite contexts like (53).

(53) A ese alumno, que los profesores no lo dejen salir hasta las 6.
to that student that the teachers not CL.3SG allow leave until the 6
'Let the teachers not allow that student to leave before 6.' Spanish

We are now moving onto Serbian, for which Wurmbrand et al. (2020) has already provided us with a well-developed analysis of complementation that will lay the foundation for the arguments in this subsection–although I will disagree with their conclusion on what finiteness in Serbian is. Wurmbrand et al. (2020) notes that Serbian allows both "finite" and nonfinite complements of verbs like *try*. We see two forms that can be the complement of *try* in (54a): the bare infinitive form without *da*, and *da* together with agreement on the embedded verb. But the infinitive is impossible with the propositional complement of *claim*, as in (54b) below:

(54) a. Pokušala sam {da čitam / čitati} ovu knjigu.
tried.SG.F AUX.1SG DA read.1SG / read.INF.IPFV this book
'I tried to read this book.'

b. Tvrdim {da čitam / *čitati} ovu knjigu.
claim.1SG DA read.1SG / *read.INF.IPFV this book
'I claimed to be reading this book.' Serbian

Our objection of investigation is this *da*. For Wurmbrand et al. (2020), assuming the framework of Wurmbrand and Lohninger (2019), the complement of *try* is an event complement, is associated with vP. On the other hand, the complement of *claim* is a full CP. Under this account, *da* itself is not a complementizer, but rather a lower clausal head that can mark vPs, TPs or CPs. I will adopt this analysis for Serbian and Mandarin, but not Icelandic or Spanish. Given the presence of verb-medial focus and topic positions in Serbian and Mandarin, but not in Icelandic or Spanish, it is difficult to determine whether focalization or topicalization would take place within the V or C domain, as Jim Huang (p.c.) points out. Even so, *da* can never precede topics in vP complements.

What Wurmbrand et al. and I diverge on is the nature of finiteness. For them, finiteness is a language specific property, and it is agreement in Serbian. So, the complement of *try* may be finite. By contrast, I will claim that finiteness is in fact not a language specific property, and it is merely the presence of an untruncated C domain. As such, under my account, the complement of *try* is never finite, as it is only as large as vP, but rather something akin to an inflected infinitive.

Todorović and Wurmbrand (2016) notes that tenseless complements of predicates such as *try* and propositional complements of predicates like *claim* allow topicalization and focalization, but with different word order. This is possible given that Serbian has verb-medial topic and focus positions. Topicalization in the embedded complement of *try* must precede *da*, but follow *da* with the complement of *claim*. I present my own illustrative examples below:[13]

(55) a. Pokušala sam [ovu knjigu]$_i$ da čitam t_i.
tried.SG.F AUX.1SG this book DA read.1SG
'I tried to read this book.'

b. * Pokušala sam da [ovu knjigu]$_i$ čitam t_i.

c. Tvrdim da [ovu knjigu]$_i$ čitam t_i.
claim.1SG DA this book read.1SG
'I claimed to be reading this book.'

d. * Tvrdim [ovu knjigu]$_i$ da čitam t_i. Serbian

This looks like Italian. On one hand, we see *da* behave as a high complementizer in the complement of *claim*, as evidenced by (55d). On the other, *da* must behave as a lower clausal head, as shown in (55b), in which *this book* moves to a verb-medial focus or topic position. Once again, I believe that this is evidence of a fundamental inability of nonfinite clauses to co-occur with high complementizers, which language specific accounts of finiteness do not predict.

One area of concern at this stage is that I have rejected an independent diagnostic for finiteness in Serbian–agreement. But I have used the lack of agreement to diagnose finiteness uncontroversially in the subsection prior. Therefore, a potential objection could be a charge of circularity: I do not have independent evidence that the *da*-complements of *try* are nonfinite. But as

[13] I have verified that with a control predicate like *decide* which takes situation complements, the complement allows topicalization both before and after *da*, as predicted by Wurmbrand and Lohninger (2019)'s ICH, which Wurmbrand et al. (2020) assumes and is based on. That predicates like *decide* can take both finite and nonfinite complements is true in English, as well.

expected, the complement of *try* requires OC PRO, which is usually associated with nonfinite clauses, but that of *claim* can license overt subjects.[14]

Mandarin has a similar pattern to Serbian. Huang (2018) makes precisely the same argument that I made for Serbian, but in Mandarin instead–his analysis can be straightforwardly translated to mine. As Huang (2018) convincingly shows, *shuo* behaves as a finite complementizer (in our terminology *high*) when it heads a finite embedded clause. In (56), topicalization is only allowed within the embedded clause, because the complement of *believe* must be finite.

(56) a. Wo xiangxin [shuo Lisi [zhe-pian baogao]$_i$ xie-wan-le t_i].
I believe SHUO Lisi this-CL report write-finish-PFV
'I believe that Lisi has written this report.'

b. * Wo [zhe-pian baogao]$_i$ xiangxin [shuo Lisi xie-wan-le t_i].

But *shuo* behaves as a lower clausal head when it heads a nonfinite embedded clause, such as the complement of *try*, with which the pattern in (56b) is possible. The complement of *try* in (57), which appears to be nonfinite–as evidenced by the requirement of a controlled PRO–involves restructuring, as it allows the embedded object to move up and precede the verb:

(57) Wo [zhe-pian baogao]$_i$ hui shefa [shuo jinkuai xie-wan t_i].
I this-CL report will try SHUO as-soon-as-possible write-finish
'I will try to finish this report as soon as possible.' Mandarin

Once again, we see the fundamental inability of a high complementizer to co-occur with nonfinite contexts. The untruncated CP2 layer blocks topicalization to a matrix verb-medial topic or focus position, as in (56b). But restructuring, and removal of the CP2 layer, allows for this movement to take place, as in (57). Concerning subject licensing, the complement of *like*–a predicate that takes vP complements similar to *try*–requires an OC PRO but that of *hope* does not, which according to Grano (2017) takes a CP, as predicted:

(58) a. Xiaoming$_i$ xihuan (*ta$_{i/j}$) chi shousi.
Xiaoming like he eat sushi
'Xiaoming likes to eat sushi.'

b. Xiaoming$_i$ xiwang (ta$_j$) chi shousi.
Xiaoming hope he eat sushi
'Xiaoming hopes to eat sushi.'

This section, in my view, shows that complementizers, when put into nonfinite clauses, cease to behave as high complementizers: depending on the language they must either behave as low complementizers or as lower clausal heads. This is strong evidence for the ISG.

3.3 Cartographic predictions

If the ordering IntP > FocP > TopP > WhP > CP1 > PropP is correct, we would be able to make further cartographic predictions on the nature of infinitives crosslinguistically, providing further evidence for Rizzi (1997)'s cartographic approach and hence, my account. For example, we would expect WhP languages to have infinitival complementizers, TopP languages to have wh-infinitives and infinitival complementizers, and IntP languages to have all of these properties.

Sabel (2006) was the first to do such a survey of infinitives, concluding that if a language has wh-infinitives, then it also has infinitival complementizers. This is evidence for the ordering WhP

[14]Landau (2004) discusses the phenomenon of finite control, in which OC PRO may appear in certain finite clauses–but only those with a subjunctive meaning. Finite control in the complement of *try* is not expected.

> CP1 under a Rizzi framework, though Sabel does not assume it. For Sabel, wh-movement simply takes place to Spec,CP, so the presence of wh-movement necessitates the presence of a C head, but not vice versa. The presence of an infinitival complementizer does not mean wh-movement is possible. But the Rizzi framework might allow for us to build on Sabel's work.

Rizzi predicts that if a language allows topicalization in infinitives, such as Italian and Hebrew, then it should have wh-infinitives and infinitival complementizers. This is already borne out in Italian and Catalan according to Sabel and Villalba (2009) respectively. Furthermore, we would predict that Hungarian and Hebrew infinitives should allow topics, focalized elements, and wh-elements given that they allow *why*-infinitives. As we saw, this prediction was also borne out.

3.3.1 *Tough*-constructions

I now provide a survey of *tough*-constructions (TCs) crosslinguistically, to provide novel evidence of cartographic ordering within infinitives. I will argue the pattern in (28) is tightly connected to their distribution: I will argue that what we call TCs in PropP languages like German, CP1 languages like Swedish and WhP languages like English all have different properties.

But first, let us discuss Chomsky (1977b)'s arguments in favor of *tough*-movement involving a step of wh-movement. Here is an example of such a construction from English:

(59) a. It is easy to play sonatas on the violin. (without *tough*-movement)
b. The violin is easy to play sonatas on. (with *tough*-movement)

As it turns out, *tough*-movement and wh-movement at the same time out of the same infinitive is not possible, as seen in (60a)-(60d). The middle Spec,CP position was occupied by a Copy of *what sonatas* prevents *this violin* from moving up in (60d).

(60) a. It is easy to play these sonatas on this violin.
b. These sonatas are easy to play on this violin.
c. What sonatas are easy to play on this violin?
d. * What sonatas is this violin easy to play on?

Under a more modern understanding of the C domain, this means that *tough*-movement takes place to Spec,WhP in English infinitives. But movement of the embedded object to Spec,WhP, an A'-position, and then to matrix Spec,TP, would be a violation of Chomsky (1977a)'s Improper Movement constraint. Instead, the embedded object is a null operator that moves to Spec,WhP while the coreferring matrix subject is base-generated:

(61) Caitlin_i is tough $[_{\text{WhP}}\ \text{Op}_i\ [_{\text{TP}}\ \text{PRO}_{\text{arb}}$ to please $\text{t}_i.]]$

Crucially, as we will soon discuss, Chomsky's analysis predicts that there can be no reconstruction effects: Poole et al. (2016) provides strong arguments in favor of this conclusion in English.

But first, let us sharpen what exactly we mean by a *tough*-construction. We have seen that the maximally PropP-infinitive languages do not allow wh-infinitives at all, so they should not have *tough*-movement. And yet, according to Comrie (1997) among others, German, a maximally PropP language might prima facie appear to have TCs, along with both of the CP1 languages.

This is contradictory, given Chomsky (1977b)'s observation that *tough*-movement involves wh-movement. How is this possible if maximally PropP languages lack a WhP layer? I propose that in fact, the maximally PropP languages do not have *tough*-movement after all.

Wurmbrand (1994) argues that German does not in fact have TCs because it has different properties from TCs that we see in English. Out of four of her tests, I will include two. For example, they do not allow arguments intervening between the embedded object and matrix subject (62a) and do not license parasitic gaps (62b):

(62) a. * Dieses Buch ist schwer Hans zu überzeugen zu lesen.
this book is hard John to convince to read
'This book is hard to convince John to read.' German

b. * weil das Buch$_i$ [ohne vorher pg_i zu kaufen] schwer t_i zu lesen ist
because the book [without before to buy] hard to read is
(Intended) 'Because the book is hard to read without having bought beforehand.'

Following Wurmbrand, I propose that we call this kind of long A-movement in German *leicht*-movement, with the resultant construction a *leicht*-construction. By contrast, genuine *tough*-movement involves a step of A'-movement to the infinitival C domain prior to A-movement to the matrix subject position, as Chomsky proposes. Chomsky's observation predicts that the C domain must be present in the infinitives of TCs.[15] Thus, we would predict languages with TCs to have wh-infinitives and/or infinitival complementizers.

German is the odd one out in my sample: it is the only language which has been reported to have TCs but does not have an infinitival complementizer or wh-infinitives. Icelandic (Sigurðsson (2016)) and Swedish (Klingvall (2018)) have infinitival complementizers. Spanish, French, Italian (Hartman (2011)) and Dutch (van der Auwera and Noel (2011)) have TCs and wh-infinitives. Wh-infinitives exist in Tamil according to Schiffman (1999), and it too has English-style TCs (Selvanathan (2017)). I therefore make these generalizations:[16]

(63) a. If a language has tough-constructions, it has wh-infinitives or infinitival complementizers. (This generalization is simplified via Sabel (2006) below.)

b. If a language has tough-constructions, it has infinitival complementizers.

I now argue that Swedish and English TCs have different properties in terms of whether they allow reconstruction or not, providing more evidence for splitting WhP and CP1.

3.3.2 The relationship between WhP and reconstruction in *tough*-constructions

It may be possible to make further generalizations. What about CP1 languages like Swedish? Surprisingly, according to Klingvall (2018), Swedish *tough*-movement patterns somewhere in between English and German. Klingvall argues that there is a step of A'-movement in Swedish TCs and they don't just involve long A-movement like in German. For example, they pattern with English rather than German in licensing parasitic gaps (64a) and are not sensitive to arguments intervening between the embedded object and matrix subject (64b).[17]

[15] I concur with Wurmbrand and Lohninger (2019) that PropP is in the C domain, but this has purely semantic effects and has no syntactic relevance.

[16] Stefan Keine (p.c.) has pointed out to me that Hindi does not have TCs, which is fully expected. None of the TP-languages in my sample have been reported to have TCs. I suspect that Norwegian and Danish may pattern similarly to Swedish, in that they also have infinitival complementizers but no wh-infinitives. Selvanathan (2018) reports that Malay involves *leicht*-constructions. I believe this covers most, if not all, of the languages which have been reported to have TCs in the literature.

[17] For space reasons I've trimmed Klingvall's example. Also, her examples are subject to dialectical variation.

(64) a. [Den artikel-n]$_i$ är svår att övertala Lisa att be Johanna att läsa t_i.
that paper-CMN.DEF is hard.CMN to convince Lisa to ask Johanna to read
'That paper is hard to convince Lisa to ask Johanna to read.'

b. Bok-en$_i$ är lätt att kritisera t_i utan att ha läst pg_i
book-CMN.DEF is easy to criticize without to have read
'The book is easy to criticize without having read.' Swedish

But Swedish infinitives are also unlike English ones, lacking wh-infinitives. The lack of this seems to lead to a surprising difference: English TCs do not allow reconstruction, whereas Swedish ones do. I present three arguments in favor of this conclusion. The first difference between English and Swedish TCs is scope. Swedish has two types of TCs: adjectival (65a) with agreement–which only allows wide scope–verbal (65b) without agreement, which allow either narrow or wide scope, given below:

(65) a. Få personer$_i$ är lätt-a för Johan att prata med t_i.
few people are easy-PL for Johan to talk to
'Few people are easy for Johan to talk to.' *few* > *easy*, **easy* > *few*

b. Få personer$_i$ går lätt för Johan att prata med t_i.
few people go easily for Johan to talk to
'Few people are easy for Johan to talk to.' or 'It is easy for Johan to talk to few people. *few* > *easy*, *easy* > *few*

But Klingvall argues that the reason (65a) does not allow reconstruction is because of the ϕ-feature agreement between the subject and the predicate in Swedish. Agreement is missing in the (65b), and we see that both wide and narrow scope are licit.[18]

Let us now see independent evidence for this conclusion. Klingvall notes that although English TCs do not allow a pronoun inside a subject to be bound by something inside the embedded clause, Swedish does: this is because Swedish allows reconstruction whereas English doesn't:

(66) a. It was hard for John to tell every farmer$_i$ [the bad news about her$_i$ goat]$_k$.

b. * [The bad news about her$_i$ goat]$_k$ was hard for John to tell every farmer$_i$ t_k.

c. % Nog var [sin$_i$ (rättmätiga) lön]$_k$ svår (för oss) att ge varje anställd$_i$
surely was REFL rightful salary difficult.CMN for us to give every employee
t_k igår eftermiddag.
yesterday afternoon
'His/her rightful salary was surely difficult (for us) to give every employee yesterday afternoon.'

Eva Klingvall (p.c.) has helped me pinpoint a novel difference below. Starting with Postal (1971), linguists have pointed out that reflexives inside picture NPs may have logophoric properties. In the spirit of Charnavel and Sportiche (2016), we can eliminate this by making the referent of the reflexive inanimate. We find a surprising contrast: "reconstruction" is heavily degraded with an inanimate according to my consultants:[19]

[18]Susi Wurmbrand (p.c.) has brought to my attention that the lack of agreement in (65b) could in fact indicate that (65b) is not a TC, but rather involving topicalization and a Verb Second word order. In that case, this argument would not work.

[19]Poole et al. (2016) cite a blogpost by Benjamin Bruening to make a similar argument to mine. A sentence like

(67) Context: A rock is above a picture of itself. If it falls, the rock will break the picture.

a. (If it is windy), it is easy for the rock to break the picture of itself.

b. ?? (If it is windy), the picture of itself is easy for the rock to break.

Swedish reflexives do not behave logophorically as English reflexives do. As Klingvall points out, Swedish reflexives in subject position are acceptable only if the subject is derived. Thus, reconstruction is possible in the Swedish (68), but only if the context provided:

(68) % Bild-er-na på sig själv är möjlig-a för sten-en att ha sönder.
picture-PL-DEF of REFL self are possible-PL for stone-DEF to break
(lit.) 'The pictures of itself are possible for the rock to break.'

The possibility of reconstruction in Swedish TCs lead Klingvall to suggest that although Chomsky's original approach is right for English, it is not in Swedish. Some kind of long movement from inside the embedded clause is necessary to get the reconstruction facts in Swedish. Klingvall proposes that the null operator moves to an A'-position in the T domain, to avoid a violation of Improper Movement. I would thus like to propose the following generalization:

(69) If a language has *tough*-movement without reconstruction, then it has wh-infinitives.

A survey of the literature indicates some support for this conclusion. French has been reported to have TCs without reconstruction (Canac Marquis (1996)) while Icelandic appears to have some reconstruction Sigurðsson (2016). But more research is needed into the understudied possibility of reconstruction within TCs crosslinguistically. To conclude, these generalizations are important, as they show that one can make predictions on a given language's properties just by knowing that it has, for example, English-style *tough*-movement. Furthermore, they provide novel evidence that the Rizzi-style blueprint for the C domain is correct.[20]

3.4 Summary

This section has primarily been concerned with crosslinguistic generalizations on the size of infinitives. I have argued for the following empirical generalization: a high complementizer cannot co-occur with a nonfinite clause. I had a two-pronged approach: I first presented a survey on the maximal size of infinitives in several different languages that have been discussed in the literature, noting that none of them co-occur with a high complementizer.

---

this aspect of herself$_i$ was tough for Sarah Palin$_i$'s autobiography to present in a good light is acceptable. But it cannot involve reconstruction, because *Sarah Palin* does not c-command the purported trace of *this aspect of herself* at any stage of the derivation. This indicates that pragmatic or logophoric factors are at play rather than reconstruction. Furthermore, an obvious question at this point is why the presence of WhP leads to these differences. Although it would go beyond the scope of this paper, one possibility is that WhP is a phase: this would straightforwardly allow for the possibility of the reflexive being logophorically bound in English under Charnavel and Sportiche (2016).

[20] If this tentative generalization is true, it can be extended. Gärtner (2009) argues for the generalization in (ia), which can be extended to mine via elementary logic in (ib). A robust indefinite/interrogative ambiguity refers to languages like English which use different words for *who* vs. *someone* whereas German does not need to.

(i) a. If a language has wh-infinitives, then its pronominal system does not have a robust indefinite/interrogative ambiguity.

b. If a language has tough-constructions without reconstruction, then its pronominal system does not have a robust indefinite/interrogative ambiguity.

Yet, absence of evidence is not evidence of absence. It could be that such a language simply has yet to be reported. Therefore, I attempted to provide evidence of absence by presenting several different languages in which an element with a certain phonetic form behaves as a high complementizer in contexts we would consider finite, but never as a high complementizer in contexts we would consider nonfinite. A plausible explanation for this fact is that nonfinite clauses necessarily cannot co-occur with a high complementizer, *because* they are truncated in size.

This, I believe, gives us a foundation to create a theory of finiteness in terms of clause size. It allows us to make precise and falsifiable definitions for a clause which is finite and nonfinite.

(70) a. A clause is finite iff it is untruncated in the C domain.
b. A clause is nonfinite iff it its CP2 layer is truncated.

Notice that properties that have often been associated to finiteness in the literature such as tense, subject licensing and agreement are not a part of my definition. Such properties merely *correlate* with the presence of CP2 under my account–that is, greater clause size correlates with tense and agreement markings. This does not block the puzzling possibility of nonfinite forms which have more agreement than finite forms, which has been claimed to be instantiated in Icari Dargwa according to Kalinina and Sumbatova (2007), for example.

Indeed, we have seen examples of the complement of *try*–as small as vP as Wurmbrand et al. (2020) argues–bearing agreement in Serbian, and nonfinite clauses in Tamil licensing subjects and bearing even tense. None of this is contradictory under my theory, as it should be. We will now determine whether this definition of finiteness still holds once we consider a range of facts.

4 Implications

Many questions remain at the end of section 4, but the three that I focus on are the following:

(71) a. Do truncated finite clauses exist?
b. Is opacity a problem for clause size theories of finiteness?
c. What consequences does this theory have on derivational-style frameworks?

I discuss apparent counterexamples in 4.1, while 4.2 discusses Pesetsky (2021).

4.1 Potential counterexamples

4.1.1 *That*-less embedded clauses

The central empirical claim of this paper is that infinitives necessarily lack the ability to co-occur with high complementizers. But there is a great deal of controversy in the literature as to whether *that*-less embedded clauses have a CP2 layer or not, which could lead to a confound.[21] For example, Bošković and Lasnik (2003) notes the following contrast, in which (72d) cannot occur without the high complementizer but (72b) can:

[21] Of course, in the literature previous authors did not refer to CP2; they referred to CP. But to be in line with the rest of this paper I will refer to CP2 rather than CP. For accounts in which CP2 is present but null, the reader is referred to Pesetsky (1992), Pesetsky and Torrego (2001), Pesetsky and Torrego (2007) and Bošković and Lasnik (2003). For accounts in which CP2 is truncated, see Hegarty (1991), Webelhuth (1992), Doherty (2000), Svenonius (1994), Bošković (1997) and Wurmbrand (2014).

(72) a. It was widely believed [that he liked linguistics].
b. (?) It was widely believed [he liked linguistics].
c. [That he liked linguistics] was widely believed.
d. * [He liked linguistics] was widely believed.

Here is the problem. (72b) is uncontroversially finite, but if it truly lacks CP2, this is a counterexample to the definition of finiteness presented in 3.4.

Wurmbrand (2017) provides an interesting discussion of stripping phenomena–the elision of declarative TPs–that may be problematic for my theory of finiteness. Based on the contrast between (73a)-(73b) on one hand and (73c)-(73d) on the other, Wurmbrand (2017) claims that stripping of embedded clauses is only possible when the embedded clause lacks a CP2.

(73) a. * Abby claimed (that) Ben would ask her out, but she didn't think that Bill (too).
b. Abby claimed (that) Ben would ask her out, but she didn't think Bill (too).
c. * Jane loves to study rocks, and John says that geography too.
d. Jane loves to study rocks, and John says geography too.

For Wurmbrand, ellipsis is the option of not realizing a Spell-Out domain. To get the contrasts in (73a)-(73b) and (73c)-(73d), Wurmbrand assumes a hierarchy CP2 > FocP > TP. If CP2 is present, CP2 is phasal but not FocP, and when CP2 is not present FocP is phasal. The Spell-Out domain of CP2 is FocP, not TP, so it cannot be elided, because stripping is just the elision of TP. But if CP2 is not present, then TP can be elided, because FocP is phasal. This allows for a natural explanation of her Embedded Stripping Generalization: that stripping of embedded clauses is only possible if the embedded clause lacks TP. This might imply that CP2 really is missing, and not merely null, in instances of embedded stripping.

It is out of the scope of this paper to contribute to this debate. But it is essential to note that whether or not *that*-less embedded clauses have CP2 or not does not have any bearing on whether the ISG is true or not. If the generalization is true, it has to be explained. But here are two potential strategies to deal with Wurmbrand's generalization.

I could take for granted approaches in which CP2 is present but null in *that*-less embedded clauses, and no problem would arise. Alternatively, the simple definition of finiteness that I present in this paper can be revised slightly to accommodate approaches where CP2 is not present in *that*-less embedded clauses. That is, I would have to admit degrees of truncation. In other words, CP2 in finite clauses *can* be truncated, but nonfinite clauses deeper truncation than merely CP2. Truncation of CP2 is necessary but not sufficient. Here is an attempt to find this deeper size.

Recall that the language with the largest attested infinitives from 3.1 is Hebrew. There is at least one more independent reason from negative polarity item (NPI) licensing to believe that Hebrew infinitives are truncated, and that this is not due to the truncation of the CP2 layer. Matrix negation can license NPI licensing inside infinitive or subjunctive complements but not indicative ones, as first noted by Landau (2004). This is shown in (74a)-(74c) below; we see that the subjunctive is headed by the high complementizer *še* and still allows NPI licensing, so this restructuring property may be due to the truncation of some other projection in the C domain.[22]

[22]I have been unable to verify whether NPI licensing is possible across propositional infinitives in Hebrew. Subjunctives do not seem to have a propositional semantics. If it is not possible, that would indicate that the functional projection of the C domain responsible for this blocking is PropP. If it is possible, then this layer is something else. I

(74) a. Lo darašti me-Gil ledaber im af-exad.
not demanded.1SG from-Gil to-speak with anybody
'I didn't demand of Gil to speak to anybody.' Infinitive

b. Lo darašti me-Gil$_i$ še-pro$_i$ yedaber im af-exad.
not demanded.1SG from-Gil that-pro will-speak-3SG.M with anybody
'I didn't demand of Gil that he speak to anybody.' Subjunctive

c. * Lo he'emanti še-Gil yedaber im af-exad.
not believed.1SG that-Gil will-speak.3SG.M with anybody
'I didn't believe that Gil would speak to anybody.' Indicative, Hebrew

It is possible that there is (at least) one other functional projection together with CP2 that is truncated when a nonfinite clause is made. Let us call this layer XP. One possibility is that CP2 is necessarily deleted whenever XP is deleted. In other words, one could define nonfinite clauses as lacking both CP2 and XP, rather than just CP2 as I have done in this paper, to get the right results with *that*-less embedded clauses. But I must leave open to future research as to what XP is.

4.1.2 Selective opacity

One aspect of my theory that may seem counterintuitive is the fact that different structures vary crosslinguistically in terms of their opacity. According to Keine (2020), nonfinite clauses in Russian are transparent to A'-movement such as topicalization but opaque to A-movement such as subject-to-subject raising, as shown by the contrast in (75a)-(75b) below:

(75) a. Kažetsja [čto èti studenty znajut tri jazyka].
seem.3SG that these students know.3PL three languages
'It seems that these students know three languages.'

b. Èti studenty$_i$ kažutsja [t$_i$ učit' tri jazyka].
these students seem.3PL learn.INF three languages
(Intended) 'These students seem to be learning three languages.' Russian

It may seem prima facie puzzling that a Russian nonfinite clause with a truncated CP2 layer does not allow raising, unlike English. Equally puzzling is the operation of *hyperraising*–that is, raising from a finite clause–which does not seem to require structure removal to take place. As Wurmbrand (2019) notes, it is a common phenomenon crosslinguistically.

An illustrative example of hyperraising in Greek, which was first noted by Felix (1989), from Bird (1999) is given below. Greek systematically lacks infinitives and allows hyperraising from subjunctive complements co-occurring with an overt complementizer:

(76) Ta pedhia arxisan na trexoun.
the children.NOM started.3PL COMP.SBJV run.3PL
'The children started to run.' Greek

On one hand, we see that Russian nonfinite clauses do not allow raising. On the other, what seem to be finite clauses in Greek allow it. But I do not think these facts are problematic. I do not know

have to leave it open to future research as to what the functional projection between CP2 and IntP is that allows NPI licensing.

whether Keine (2020)'s theoretical tool of probes having different search domains is the right notion to capture selective opacity effects, but I do not need to take it for granted. What Keine convincingly shows that selective opacity is a pervasive phenomenon, which is all this theory needs. The lack of raising from Russian nonfinite clauses and possibility of hyperraising in languages like Greek, in my view, are merely an instance of selective opacity effects, and it is not the case that Russian nonfinite clauses are larger than Greek finite clauses.

4.1.3 Factives are not truncated

As has been noted extensively in the literature thus far, factives do not allow many of the properties of the C domain such as topicalization or focalization, as Hooper and Thompson (1973), Haegeman (2012) and others point out. An example with the complement of *regret* is below:

(77) * John regrets that this book Mary read.

This has led Miyagawa (2017) to claim that factives are in fact truncated in the C domain. This is at odds with my definition of a finite clause, which is fully untruncated in the C domain. As such, I adopt and defend Haegeman (2012)'s analysis of null operator movement in complements of factive predicates, rather than truncation. I present some corroborating evidence for her account in the form of infinitives, as well.

Haegeman is not the first to suggest null operator movement in factives. Hegarty (1992) points out that the complement clauses of factives are weak islands for extraction, whereas those of nonfactives are not, as seen below.

(78) a. How do you suppose that Maria$_i$ fixed the car t$_i$?
b. * How did you notice that Maria$_i$ fixed the car t$_i$?
c. Why does Mary$_i$ think that Bill left the company t$_i$?
d. * Why does Mary$_i$ regret that Bill left the company t$_i$?

As Haegeman (2012) points out, almost every property of the C domain that we have discussed thus far involves a step of A'-movement. Both null operators and a truncation analysis would get the desired result as both disallow movement. If it ever were possible to base-generate elements into a Spec position in the C-domain, for example Spec,TopP, then it would be possible to distinguish between the accounts, as they make different predictions.

Temporal adjuncts, in fact, seem to be base-generated into a Spec position of the articulated left periphery. Rizzi (1997) assumes they are Merged to Spec,TopP, although Rizzi (2001) distinguishes the position of topics from modifiers, positing a dedicated projection, ModP. However, for simplicity, I will continue assume that it is Merged onto Spec,TopP:

(79) [$_{\text{TopP}}$ Last week, [$_{\text{TP}}$ I was in Tokyo.]]

If temporal adjuncts are base-generated, then we would predict that they should be acceptable with factives. This prediction is borne out:

(80) John regrets that during dinner Mary read this book.

As mentioned, colloquial English appears to have double complementizer constructions for some:

(81) % She maintained **that** when they arrived **that** they would be welcomed.

According to my consultants, this sentence is equally acceptable with the factive *regret*, indicating the presence of CP2, TopP and CP1 layers and therefore a highly articulated structure:

(82) % She regretted **that** when they arrived **that** they weren't welcomed.

On the other hand, we would also predict that, as English infinitives are quite truncated, that they cannot take preverbal temporal adjuncts. This prediction is borne out, according to data from Shlonsky and Soare (2011). In the contrast below, the adjunct *at 5* cannot refer to the cooking of dinner; it must refer to the time of the promise–that is, it must be an adjunct to the matrix sentence rather than the infinitive. However, this is possible with the finite version of the sentence:

(83) a. * John promised us at 5 to cook dinner for his children.
b. John promised us that at 5 he would cook dinner for his children.

I conclude that factives are not truncated in the C domain.

4.2 Consequences on Exfoliation

In this subsection, I discuss the theory of finiteness most recently defended by Pesetsky (2021), which is the most well-developed theory of finiteness in terms of clause size. I conclude that although I believe the theory is on the right track, it does not have the theoretical tools needed to be able to derive the crosslinguistic variation in the maximal size of infinitives.

4.2.1 Background on clause size and finiteness

The core questions that Exfoliation seeks to address are: why do nonfinite clauses exist in the first place, and why do the properties of the subject position in nonfinite clauses differ from their finite counterparts? For Pesetsky, ultimately all nonfinite clauses are created via a process of subject extraction, even control constructions which do not prima facie involve subject extraction, putting aside movement theories of control like Hornstein (1999)'s. All clauses are born as full and finite CPs. Infinitives are made, not born, contra selectional accounts in which different predicates, like raising and control predicates, picked the size of their complement.

One piece of evidence for this is as follows. It has often been considered, since Vergnaud's letter to Chomsky and Lasnik, that the driving factor for raising-to-object constructions is Case assignment, and all nouns need Case. The distribution of DPs appears to be restricted:

(84) We are sure [$_{CP}$ that the world is round] vs. *[$_{DP}$ the world's roundness].

Under Case-driven accounts of raising-to-object constructions, the subject of the nonfinite clause in (85a) is not able to get Case in its base-generated position, so it needs to move up, perhaps to Spec,VP of the matrix verb. There, it is assigned accusative Case. A similar line of reasoning drives the assigning of nominative Case to the matrix subject in raising-to-subject constructions in (85b). In (85c)-(85f), we see that elements which cannot assign Case lead to unacceptability:

(85) a. Caitlin believes him$_i$ [t$_i$ to be smart]. *raising-to-object*
b. Caitlin$_i$ seems [t$_i$ to be smart]. *raising-to-subject*
c. * It seems Caitlin to have solved the problem. *unaccusative matrix verb*
d. * It was believed Caitlin to speak Irish well. *passive matrix verb*

e. * Caitlin is aware Madeline to be the cutest. *adjective*

f. * Caitlin's belief it to have been raining. *noun*

But this makes an incorrect prediction. If elements like CPs don't need Case, we would predict structures like the ones below to be grammatical. We obtain the same contrast with CPs:

(86) a. Caitlin considers [that the world is round] to be a tragedy. *raising-to-object*

b. [That the world is round] seems to be a tragedy. *raising-to-subject*

c. * It seems [that the world is round] to be a tragedy. *unaccusative matrix verb*

d. * It was believed [that the world is round] to be a tragedy. *passive matrix verb*

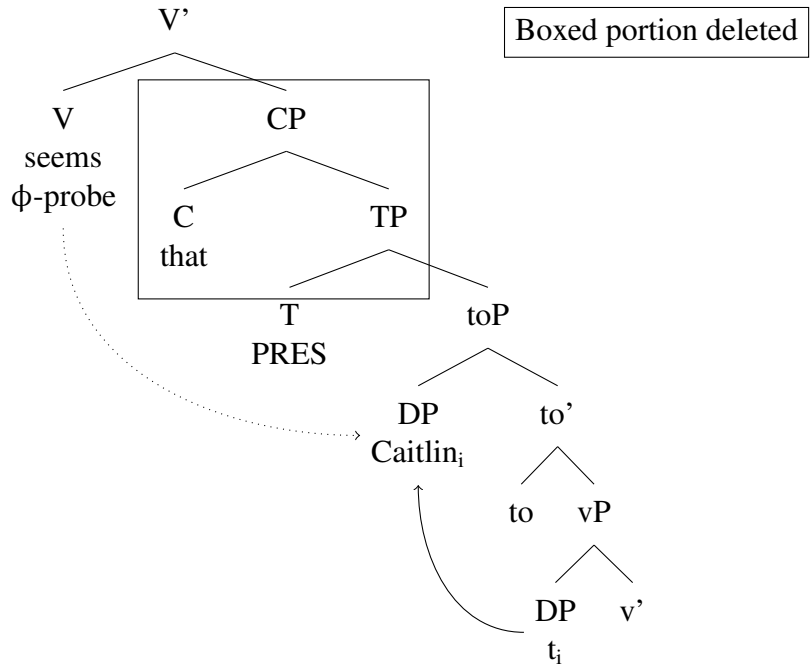
e. * Caitlin is aware [that the world is round] to be a tragedy. *adjective*

f. * Caitlin's belief [that the world is round] to have been raining. *noun*

Under Exfoliation, these examples follow do not follow from Case. All clauses are born finite and are reduced in structure to nonfinite via a process of subject extraction. While raising-to-object and -subject constructions allow (85a) and (85b) because they involve subject extraction, (85c)-(85f) are ruled out because they involve illegal infinivization, or subject extraction: these constructions simply do not have a subject extraction probe.

Let's see how a derivation of the sentence *Caitlin seems to be happy* would work. First, it is assumed that the embedded clause is born finite, so the embedded clause might look like *seems that Caitlin is happy* at a point in the derivation, as shown in the tree below. Further, all clauses are born with a toP, the relevance of which will be discussed shortly: it can only be pronounced post-Exfoliation.[23] Exfoliation removes structure to allow the probe on *V* to extract the subject:

(87)



[23] I have not discussed several technical details in Pesetsky's proposal for space; for example, the phase property of CP moves to toP after Exfoliation, and it is assumed that the DP *Caitlin* need not move to Spec,TP immediately.

The projection toP is present in all finite clauses, as well. Though it is present, to ensure that *to* is pronounced only with infinitives, Pesetsky adds a further condition–dubbed the Exposure Condition–on how certain elements can be pronounced if they head a phase:

(88) **The Exposure Condition**

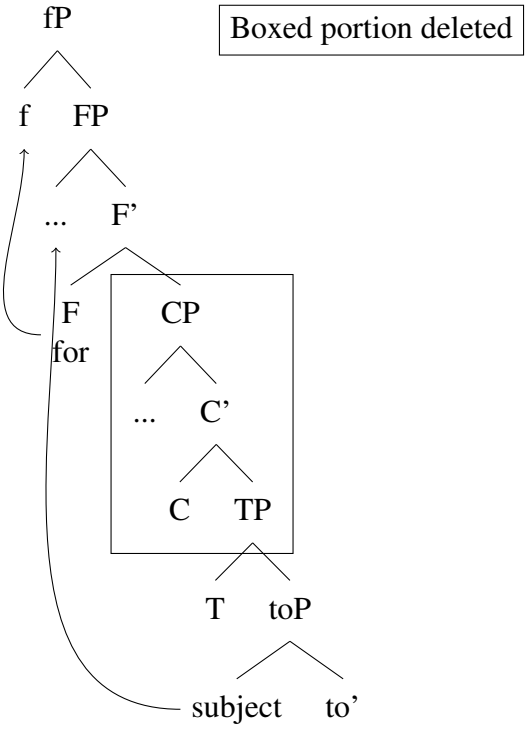
a. A is *exposed* iff it heads a phase and does not retain a specifier. (In other words, if it is the highest element in its phase.)

b. A functional head is overt iff it is exposed.

It's easy to see how derivation would apply to raising-to-subject and -object constructions. But under Exfoliation, sentences with *for*-infinitives like *Mary is eager for Caitlin to discuss the topic* involves subject extraction, as well. This seems prima facie counterintuitive given that *for* only occurs with infinitives to begin with: if infinitives are made and not born, how would *for* even come into play during a derivation? The answer is simple: *for*-infinitives have a similar syntax with raising-to-object constructions.

I will now discuss what I find to be the most controversial notion in this framework: the notion of a *superstructure*. *For* is not a complementizer, but rather an irrealis element that takes a CP as its complement.[24] This irrealis element is contained in a superstructure that Exfoliates and allows the embedded subject to raise to a position at which *for* can assign it with accusative Case. A simplified illustration of a derivation of a for-infinitive is provided below:

(89)



To get a structure for control infinitives, we have two options. First, we can either assume Hornstein (1999)'s movement theory of control, which would have a derivation identical to that of

[24]The reader is referred to Pesetsky (2021) for empirical evidence for this claim, which I will not be presenting in this paper. Under my account, *for* is a low complementizer in English.

(87), involving subject extraction in a very natural way. But if we don't assume Hornstein's theory, the subject extraction is not obvious. In that case, the derivation of a control infinitive would require a superstructure and an invisible *for*, as in (89).

Putting aside superstructures, we've seen that under Exfoliation, infinitives all come in the same size: toP, which is smaller than CP and TP but larger than vP. This is at odds with Wurmbrand and Lohninger (2019)'s (W&L) recent work which, in my view, conclusively show that infinitives can also come in different sizes. W&L provide empirical data that control complements can in fact have CP and TP layers. They propose that there are three kinds of control complements: propositional, which are CPs; situational, which are TPs; and events, which are vPs.

Propositional complements involve those which can be assigned a truth value, ex. *ESA claimed life to be on Venus, which seems true*. But situational ones cannot, ex. **Mary asked me to buy an apple, which is true*. One empirical test that they provide is given below; propositional infinitives behave like finite clauses in that they cannot occur in the non-progressive form when referring to a non-generic episodic event, but situational infinitives can:

(90) Clara decided to eat salad right now.

(91) Clara claimed to be eating/*eat salad right now.

Under Exfoliation, it is not straightforward to capture such contrasts, given that all infinitives–putting aside superstructures–are only as large as toP. But the most problematic issue is that the Exposure Condition cannot be used together with the arguments that infinitives can come in three different sizes. This would mean that a great deal of the framework would have to be altered.

Before concluding this section, I will note that this paper has much in common with W&L. We both show that infinitives can come in different sizes. For W&L, the maximal size for infinitives crosslinguistically is CP, but this is without splitting the C domain. Once we do so, we observe that infinitives can come in at least **eight** different sizes across languages: CP2 > IntP > FocP > TopP > WhP > CP1 > PropP > TP > vP.

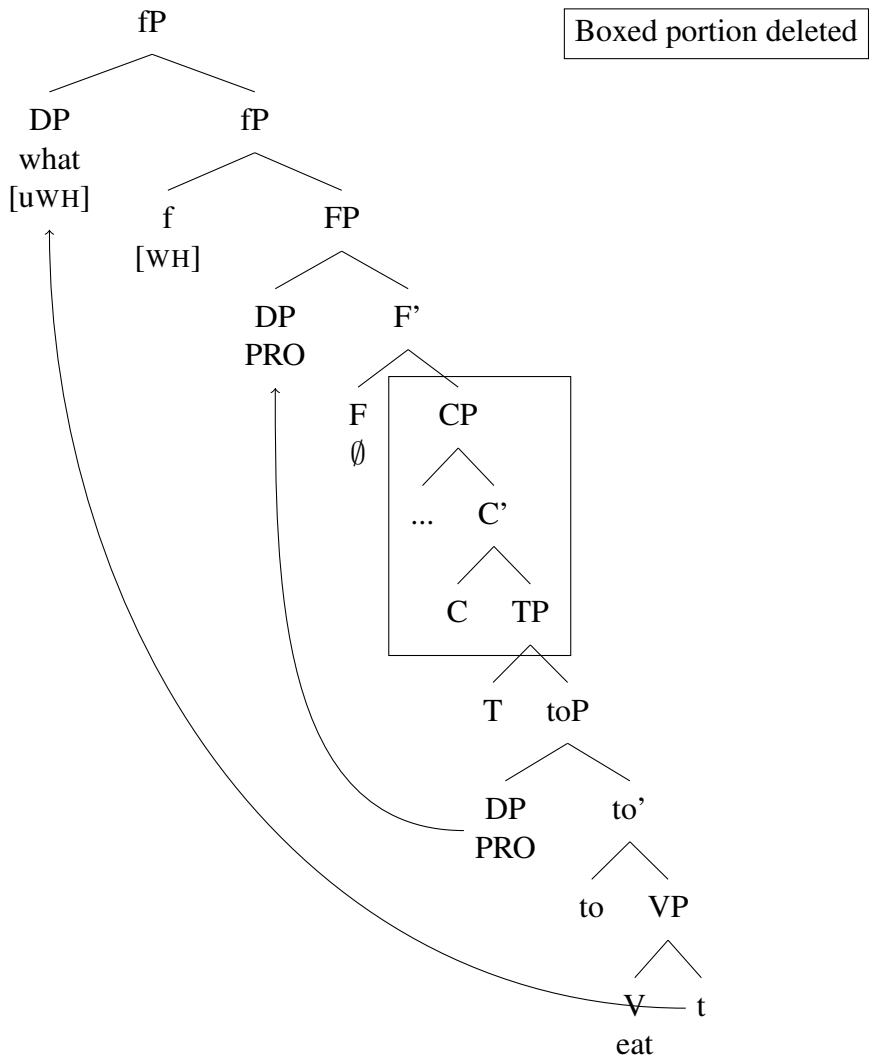
4.2.2 Discussion

On one hand, I believe that my empirical generalization–that nonfinite clauses by definition lack a CP2 layer, in which high complementizers are realized–constitutes evidence, in my view, that some kind of derivational process, perhaps Müller (2020)'s operation Remove or Pesetsky (2021)'s Exfoliation, is responsible for the truncated size of infinitives. Structure removal takes place because CP2 is a barrier for syntactic operations like subject extraction. Most importantly, it provides strong empirical support for the presupposition behind Muller and Pesetsky's work that finiteness is a matter of clause size. Derivational frameworks *predict* my generalization, while non-derivational frameworks of complementation do not.

On the other, the evidence from section 3 is at odds with Pesetsky's "one-size-fits-all" approach, where all infinitives have the same size: toP, apart from the superstructure that is sometimes added. To see where this goes wrong, let us see an attempt, under the Exfoliation framework, to derive a wh-infinitive such as *I know what to eat*. In this tree, f^0 has a WH-feature allowing the wh-infinitive to be formed.[25]

[25] I am omitting the movement of F^0 to f^0 for simplicity. One might object that this tree violates minimality conditions on movement. See, for example, Preminger (2014) on why it does not: the probe on f^0 looks specifically for WH-features even if PRO is a more local DP. It can skip past PRO because it does not have WH-features.

(92)



This sets the stage to present the first problem with the Exfoliation framework: it misses generalizations concerning the size of infinitives cross-linguistically. That is, it is not obvious under Exfoliation why wh-infinitives do not exist in languages like Hindi, German, Swedish and Icelandic, or why infinitival complementizers do not exist in Hindi and German, given that *all* of these languages have control constructions and hence, superstructures.

Under my account, the presence of infinitival complementizers in English is predicted from the presence of wh-infinitives. The differing properties of *tough*-movement in German, Swedish and English is also predicted: the infinitives of these languages come in three different maximal sizes, which are TP, CP1 and WhP. Superstructures do not allow such predictions to be made, because superstructures across languages are the same size: it is mysterious why English has infinitival complementizers, why maximal infinitive size correlates with the kind of *tough*-movement that is present, and why TP and CP1 languages do not have wh-infinitives, among other facts.

To start accounting for the lack of wh-infinitives in TP and CP1-languages, it is possible for David Pesetsky (p.c.) to claim these languages do not license WH-features on their superstructures whereas English does. But to see where this goes wrong, recall that there are even larger infinitives than those of English: Italian's infinitives go up to TopP, Hebrew's go up to IntP. Once

again, the presence of WhP and CP1 in Italian is predicted from the presence of TopP, which the Exfoliation framework does not predict; the presence of FocP, TopP, WhP and CP1 is predicted in Hebrew from the presence of IntP, none of which the Exfoliation framework predicts.

To account for these predictions, it seems impossible to not assume a Rizzi-style left periphery for superstructures. But at this point his superstructure has become indistinguishable from Rizzi's articulated C domain, and ultimately, he has to end up admitting that infinitives come in different sizes, as well. Though I cannot offer an alternative at this time, the notion of a superstructure thus seems redundant. Crosslinguistic variation between the sizes of infinitives has to be allowed in a non-derivational manner.

5 Conclusion

This paper has been an investigation on the size of infinitives. After laying the groundwork for this endeavor in sections 1-2, section 3 presented evidence that the size of infinitives can vary across languages. I have argued that finiteness is a matter of clause size, and defined finite clauses as those which are untruncated in the C domain, whereas nonfinite clauses are those which lack a CP2 layer. I have defended this approach, and discussed its consequences in section 4.

This paper has only investigated the size of finite and nonfinite *embedded* clauses, not root ones. As such, there are many open questions left at the conclusion of this paper. The most obvious one is that *that* can only appear with embedded clauses:

(93) (*That) Caitlin likes chocolate.

This is a question for all theories, and not mine specifically–perhaps there is just an independent requirement for *that* to be pronounced only with embedded clauses, but it is always present. But this is by no means a universal; as we will see, it seems to be attested in the Romance languages.

Furthermore, one prediction that my account makes involves projections above CP2, which appear to be attested in at least Romance. Cruschina and Remberger (2018) discusses constructions in Romance in which a complementizer is present in root clauses, and is preceded by an adjective or an adverb. It appears to be preceded by an adjective in (94):

(94) Certo che la capito!
certain that have.3SG understand.PST.PTCP
'Of course she understood!' Italian

For Cruschina and Remberger (2018), this indicates that there are a set of projections above CP2 which encode speaker-oriented and pragmatic features such as evaluative, evidential or epistemic values. The prediction is that no projection above CP2 should be present not just with infinitives, but with other nonfinite clauses like imperatives as well.

Of course, I have not discussed the nature of imperatives like "*Catch her!*" and how they come into being, which I leave to future work. If all embedded clauses are born finite as in Pesetsky (2021), it's not clear if the technology could be extended to root nonfinite clauses. But it is natural to suppose that they are missing many functional projections, leading to a truncated structure.

There are other kinds of structures which are often associated with nonfiniteness, like subjunctives and gerunds. It remains to be seen how this account can be extended to gerunds, which have a nominal nature, and structures like nominalized infinitives in Turkish–for which I presented evidence in section 3.1 that it is truncated.

What is more surprising, though, is the potential application of this theory of finiteness to subjunctive structures, which seem to be unambiguously headed by CP2–at least sometimes. For example, although subjunctives in English are headed by CP2 *that*, what is surprising is that subjunctives appear to be truncated in the C domain, in a manner very similar to infinitives:

(95) a. *No topicalization within infinitives*: *I suggested that this book he read.
b. *No focalization within infinitives*: *I suggested that THIS BOOK he read.
c. *No why-infinitives*: *I suggested why she eat salad.
d. *No if*: *I suggested that if he eat ice cream, then he exercise.
e. *No temporal adjunct*: ??I suggested that during dinner she eat salad.

This, coupled with Landau's observation in (74a)-(74c) above that Hebrew subjunctives more permeable than Hebrew indicatives, could be reason to believe that all subjunctives are truncated in the C domain in some regard. This could help explain the sense in the literature that subjunctives are borderline between finite and nonfinite, often co-occuring with morphology associated with finite clauses, and yet with controlled elements like PRO. The ultimate claim I would want to make is that all control complements are truncated in size, and hence never fully finite–though I must leave the details of this open to future research.

At the very least–no matter what one thinks of the analysis of finiteness in this paper–the goal of this paper has been to introduce the reader to novel empirical generalizations concerning nonfinite clauses. The inability for nonfinite clause to appear with high complementizers under the articulated C domain is a mystery worth investigating.

References

Adger, David. 2007. *Three domains of finiteness: a minimalist perspective*. OUP Oxford.

Al-Aqarbeh, Rania. 2011. Finiteness in Jordanian Arabic: a semantic and morphosyntactic approach. Doctoral Dissertation, University of Kansas.

van der Auwera, Johan, and Dirk Noel. 2011. Raising: Dutch between English and German 23:1–36.

Barbosa, Pilar. 2001. On inversion in wh-questions in Romance. In *Subject inversion in Romance and the theory of universal grammar*, ed. Aafke C. Hulk and Jean-Yves Pollock, 20–59. Oxford University Press.

Barrie, Michael. 2007. Control and wh-infinitivals. In *Studies in natural language and linguistic theory*, 263–279. Springer Netherlands.

Bhatt, Rajesh. 2006. *Covert modality in non-finite contexts*. Walter de Gruyter.

Bird, Steven. 1999. Multidimensional exploration of online linguistic field data. In *North East Linguistics Society*, ed. Tamanji Pius, Masako Hirotani, and Nancy Hall. University of Delaware: GLSA.

Bocci, Giuliano. 2007. Criterial positions and left periphery in Italian: Evidence for the syntactic encoding of contrastive focus. In *Papers from the consortium Workshops on Linguistic Theory*, volume 3.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Massachusetts: MIT Press.

Bošković, Željko, and Howard Lasnik. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34:527–546.
Bouchard, Denis. 1984. *On the content of empty categories*. Dordrecht.
Canac Marquis, R. 1996. The distribution of *a* and *de* in tough constructions in French. In *Grammatical theory and Romance languages*, ed. Karen Zagona, 35–46. John Benjamins.
Carstens, Vicki, and Michael Diercks. 2009. Parameterizing case and activity: Hyper-raising in Bantu. In *Proceedings of NELS 40*.
Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor binding – what French inanimate anaphors show. *Linguistic Inquiry* 47:35–87.
Chomsky, Noam. 1977a. Conditions on transformations. In *Essays on form and interpretation*, 81–162. New York, New York: Elsevier North-Holland, Inc.
Chomsky, Noam. 1977b. On wh-movement. In *Formal Syntax*, ed. P. Culicover, T. Wasow, and A. Akmajian. New York, New York: Academic Press.
Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 1–52. Cambridge, Massachusetts: MIT Press.
Comrie, Bernard. 1997. Tough movement and its analogs in germanic languages. In *Language and its ecology*. DE GRUYTER MOUTON.
Cruschina, Silvio, and Eva-Maria Remberger. 2018. Speaker-oriented syntax and root clause complementizers 18:336–358.
Dasgupta, Probal. 1982. Infinitives in Bangla and English: phrases or clauses. *Bulletin of the Deccan College Research Institute* 41:176–209.
Demonte, Violeta, and Olga Fernández-Soriano. 2009. Force and finiteness in the spanish complementizer system 21.
Doherty, Cathal. 2000. *Clauses without that: The case for bare sentential complementation in English*. Garland.
Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht, The Netherlands: D. Reidel Publishing Company.
Faarlund, Jan Terje. 2015. The Norwegian infinitive marker 1–10.
Felix, Sascha. 1989. Finite infinitives in modern greek. In *Linguistik aktuell/linguistics today*, 113–132. John Benjamins Publishing Company.
Gärtner, Hans-Martin. 2009. More on the indefinite-interrogative affinity: The view from embedded non-finite interrogatives 13:1–37.
van Gelderen, Elly. 1998. The future of for to 10:45–71.
Grano, Thomas. 2017. Finiteness contrasts without Tense? A view from Mandarin Chinese. *Journal of East Asian Linguistics* 26:259–299.
Haegeman, Liliane. 2006. Argument fronting in English, Romance CLLD, and the left periphery. In *Crosslinguistic research in syntax and semantics. negation, tense and clausal architecture*, ed. Rafaella Zanuttini, Héctor Campos, Elena Herburger, and Paul H. Portner, 27–52. Georgetown University Press.
Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford University Press.
Hartman, Jeremy. 2011. Intervention in tough-constructions. In *Proceedings of the 39th Meeting of the North East Linguistic Society (NELS 39)*, ed. Suzi Lima, Kevin Mullin, and Brian Smith, 387–398. GLSA.
Hegarty, Michael. 1991. Adjunct extraction and chain configurations. Doctoral Dissertation, MIT.

Hegarty, Michael. 1992. Adjunct extraction without traces. In *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, ed. Dawn Bates, 209–222. Stanford, California: Center for the Study of Language and Information.
Henry, Alison. 1995. *Belfast English and standard English*. Oxford: Oxford University Press.
Hooper, Joan B., and Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic inquiry* 4:465–499.
Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30:69–96.
Hornstein, Norbert, and David Lightfoot. 1987. Predication and PRO 63:23–52.
Hsu, Brian. 2015. Variation in Bangla complementizer order at the syntax-prosody interface. In *Proceedings of NELS 45*. GLSA.
Huang, Nick. 2018. Control complements in mandarin chinese: implications for restructuring and the chinese finiteness debate 27:347–376.
Kalinina, Elena, and Nina Sumbatova. 2007. Clause structure and verbal forms in Nakh-Daghestanian languages. In *Finiteness: Theoretical and Empirical Foundations*, ed. Irina Nikolaeva, 183–249. OUP Oxford.
Kalm, Mikael. 2016. Control infinitives and ECM-infinitives in Swedish: A synchronic and diachronic investigation. Doctoral Dissertation, Uppsala University.
Kayne, Richard S. 1981. ECP extensions. *Linguistic Inquiry* 12:93–133.
Keine, S. 2020. *Probes and their horizons*. Linguistic Inquiry Monographs. MIT Press.
Klingvall, Eva. 2018. Agreement and reconstruction correlate in Swedish 36:1165–1205.
Kornfilt, Jaklin. 1996. On some infinitival wh-constructions in Turkish.
Koster, Jan. 1984. On binding and control. *Linguistic Inquiry* 15:417–459.
Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811–877.
Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge, England: Cambridge University Press.
Larsson, Ida. 2017. Double complementizers. *Nordic Atlas of Language Structures Journal* 1.
Lujan, Marta. 1980. Clitic promotion and mood in Spanish verbal complements 381–484.
McFadden, Thomas, and Sandhya Sundaresan. 2014. Finiteness in south asian languages: an introduction. *Natural Language & Linguistic Theory* 32:1–27. URL `http://www.jstor.org/stable/43697712`.
Miyagawa, Shigeru. 2017. Topicalization 152:1–29.
Müller, Gereon. 2020. Rethinking restructuring. In *Syntactic architecture and its consequences II: Between syntax and morphology*, 149–190. Language Science Press.
Nikolaeva, Irina. 2007. *Finiteness: Theoretical and empirical foundations*. OUP Oxford.
Pesetsky, David. 1992. Zero syntax, volume 2.
Pesetsky, David. 2021. Exfoliation: towards a derivational theory of clause size. URL `https://ling.auf.net/lingbuzz/004440`.
Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 355–426. MIT Press.
Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. Amsterdam: John Benjamins Publishing Company.

Platzack, Christer. 1986. The structure of infinitive clauses in Danish and Swedish. In *Scandinavian syntax*, ed. Dahl sten and Anders Holmberg, 123–137. Stockholm Sweden: University of Stockholm.

Poole, Ethan, Jon Ander Mendia, and Stefan Keine. 2016. It's tough to reconstruct.

Postal, Paul. 1971. *Cross-over phenomena*. New York, New York: Holt, Rinehart and Winston.

Preminger, Omer. 2014. *Agreement and its failures*. Number 68 in Linguistic Inquiry Monographs. MIT Press.

Raposo, Eduardo. 1987. Case theory and Infl-to-Comp: the inflected infinitive in European Portuguese. *Linguistic Inquiry* 18:85–109.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2001. Reconstruction, weak island sensitivity and agreement. In *Semantic interfaces: reference, anaphora and aspect*, ed. Carlo Cecchetto, Gennaro Chierchia, and Maria Teresa Guasti, 145–176. Stanford, California: CSLI.

Sabel, J. 2006. Impossible infinitival interrogatives and relatives. In *Form, structure, and grammar*.

Schiffman, H.F. 1999. *A reference grammar of spoken tamil*. Reference Grammars. Cambridge University Press.

Selvanathan, Nagarajan. 2017. Subject theta roles and Tamil tough movement. URL `https://ling.auf.net/lingbuzz/003767`.

Selvanathan, Nagarajan. 2018. Malay tough movement, restructuring and the theory of phases. URL `https://ling.auf.net/lingbuzz/004029`.

Shlonsky, Ur. 2014. Topicalization and focalization: a preliminary exploration of the hebrew left periphery. In *Peripheries*.

Shlonsky, Ur, and Gabriela Soare. 2011. Where's 'why'? *Linguistic Inquiry* 42:651–669.

Sigurðsson, Einar Freyr. 2016. Circumventing the ban on improper movement in Icelandic. In *Proceedings of NELS 46*.

Sigurðsson, Halldór Ármann. 1989. Verbal syntax and Case in Icelandic. Doctoral Dissertation, Department of Scandinavian Languages, University of Lund.

Svenonius, Peter. 1994. Dependent nexus: Subordinate predication structures in English and the Scandinavian language. Doctoral Dissertation, University of California, Santa Cruz.

Szabolcsi, Anna. 2009. Overt nominative subjects in infinitival complements: Data, diagnostics, and preliminary analyses .

Szécsényi, Krisztina. 2009. On the double nature of Hungarian infinitival constructions 119:592–624.

Thráinsson, H. 2007. *The Syntax of Icelandic*. Cambridge Syntax Guides. Cambridge University Press. URL `https://books.google.com/books?id=70dbeoPHgzgC`.

Thraínsson, Höskuldur. 1993. On the structure of infinitival complements 3.

Todorović, Neda, and Susi Wurmbrand. 2016. *Finiteness across domains*.

Ussery, Cherlon, Lydia Ding, and Yining Rebecca Liu. 2016. The typology of mandarin infinitives. *Proceedings of the Linguistic Society of America* 1:28.

Villa-Garcia, Julio. 2012. The Spanish complementizer system: Consequences for the syntax of dislocations and subjects, locality of movement, and clausal structure. Doctoral Dissertation, University of Connecticut.

Villalba, Xavier. 2009. *The syntax and semantics of dislocations in catalan. a study on asymmetric syntax at the peripheries of sentence*. Lambert.

Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford University Press.

Wheelock, Kane. 2015. German wh-infinitives: restructuring and cyclic movement. Master's thesis, New York University.

Wurmbrand, Susanne, Magdalena Lohninger, Neda Todorovic, Iva Kovac, and Caroline Pajancic. 2020. Finiteness in south slavic complement clauses: Evidence for an implicational finiteness universal 119–137.

Wurmbrand, Susi. 1994. The German leicht-construction. In *Wiener Linguistische Gazette*.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Mouton de Gruyter.

Wurmbrand, Susi. 2014. The merge condition. In *Language faculty and beyond*, 130–166. John Benjamins Publishing Company.

Wurmbrand, Susi. 2017. Stripping and topless complements 48:341–366.

Wurmbrand, Susi. 2019. Cross-clausal A-dependencies. In *Proceedings of CLS 54*.

Wurmbrand, Susi, and Magdalena Lohninger. 2019. An implicational universal in complementation: Theoretical insights and empirical progress. In *Propositional arguments in cross-linguistic research: Theoretical and empirical issues*, ed. Jutta M. Hartmann and Angelika Wollstein. Mouton de Gruyter.