

# The Infinitive Size Generalization\*

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

This paper presents a systematic and detailed investigation into the infinitival left periphery in 17 languages. The survey indicates that languages can be split into five different classes based on the left peripheral properties of their infinitives. This leads to the following generalization on languages with infinitives: an infinitive cannot co-occur with a high complementizer (such as *that* in English). Although such an observation may seem trivial, assuming an articulated left periphery allows a redefinition of *that*. This paper combines arguments that finiteness is a matter of clause size together with truncation theories of infinitives to argue for a novel understanding of finiteness, proposing precise and falsifiable definitions for finite and nonfinite clauses. In addition, the survey presents evidence in favor of the existence of novel and fine-grained cartographic generalizations on the left peripheral properties of infinitives.

**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, as finiteness is seen as a property of the verb. 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.<sup>1</sup>

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 difícil [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: the relevant property is tense. But in Tamil, we see the opposite scenario according to McFadden and Sundaresan (2014). In (2) below, the embedded clause is marked with tense but not agreement, yet appears to be nonfinite given its inability to stand alone:

---

\*To be updated

<sup>1</sup>See Nikolaeva (2007) for a helpful introduction to finiteness in linguistic literature.

- (2) Raman<sub>i</sub> [EC<sub>i</sub> Seetha-vaē naaleeki paar-pp-adaagae] so-*nn*-aan.  
 Raman EC Seetha-ACC tomorrow see-FUT-GER-ACC say-PST-3MSG  
 ‘Raman<sub>i</sub> spoke of [EC<sub>i</sub> 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). A-movement out of a finite clause is not possible, as in (3a), but it is from a nonfinite clause, as in (3b).

- (3) a. \*David<sub>i</sub> seems [that t<sub>i</sub> likes exfoliation].  
 b. David<sub>i</sub> seems [t<sub>i</sub> to like exfoliation].

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: 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 provides novel evidence for these theories of finiteness, but with more fine-grained distinctions than just CP and TP.<sup>2</sup>

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 xihuan (\*ta) chi shoushi.  
 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 Tamil example in (5) below.

- (5) [Vasu poori porikk-a] Raman maavu vaangi-n-aan.  
 Vasu.NOM poori.ACC fry-INF Raman.NOM flour.ACC buy-PST-M.3SG  
 ‘Raman bought flour for Vasu to fry pooris.’

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 (2004b) and Szabolcsi (2009), who assume a more complex relationship between tense, agreement and subject licensing in clauses.

<sup>2</sup>I remain agnostic as to whether, as Pesetsky and Müller suggest, nonfinite clauses start out larger and are truncated during the derivation. The generalizations that will be proposed in this paper are independent of such theories. When I use the word *truncated*, I do not intend to presuppose that such clauses start out large and end up smaller by some operation. This is a separate empirical matter that is not within the scope of investigation of this paper.

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. 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.<sup>3</sup>

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 the notion 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 claimed (\*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 clarify 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 that is schematized below. Further details will be provided in section 2, 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)** > InterrogativeP > FocusP > TopicP > WhP > **CP1 (low)** > TP

I define a *high complementizer* as a complementizer that heads CP2. By definition, high complementizers always precede topic and focus-marked elements. My survey indicates 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, if it has not already been truncated beforehand—which is often the case. When topics and focalized elements are possible at all in an infinitive, low complementizers must follow them.

It is in fact possible to distinguish between these complementizers even in English: *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<sub>t<sub>i</sub></sub>, Chomsky wrote t<sub>i</sub>.

---

<sup>3</sup>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. 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.

*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.

(9) \* Chomsky claimed Manufacturing Consent<sub>i</sub>, to have written.

Languages differ in this regard. For example, 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]<sub>i</sub> lenasot t<sub>i</sub>.  
 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

This indicates that finiteness may be defined in terms of clause size. Based on my crosslinguistic survey, the main generalization that I will argue in this paper is in (12) below.<sup>4</sup>

(12) **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 cartographic generalizations beyond (12) above can be made. I present detailed and systematic evidence from 17 languages in the survey. The data was obtained through a mixture of gathering what has already been reported in the literature, plus consultation with at least one native speaker consultant per language. Table 1 in section 3 indicates that at the very least, the following generalizations are true:

- (13) a. *Sabel’s (2006) Generalization:* If a language has *wh*-infinitives, then it also has infinitival complementizers.  
 b. If a language allows topicalized elements within its infinitives, then it also has *wh*-infinitives and infinitival complementizers.  
 c. If a language allows *why* and *if* in its infinitives, then it has contrastive focus and topicalization within its infinitives, *wh*-infinitives and infinitival complementizers.

This paper is structured as follows. Section 2 introduces the reader to Rizzi (1997)’s structure for the C domain. Section 3 presents the crosslinguistic survey of the infinitival left periphery. Section 4 discusses the ISG in more detail, arguing that it is true, while Section 5 discusses potential counterexamples to the ISG. Section 6 discusses whether further generalizations beyond the ones in (13a)-(13c) above. Section 7 concludes.

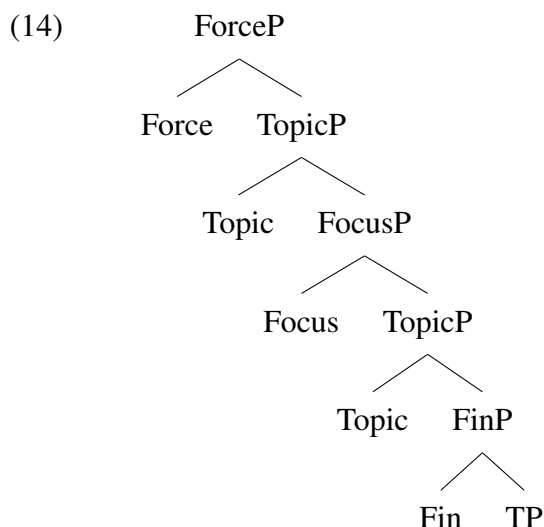
<sup>4</sup>The scope of this paper is to cover only the clausal size of infinitives; as such, I will only focus only on infinitives. For the most part, I will not discuss gerunds or other kinds of nonfinite clauses in this paper, leaving it to future research. However, it is likely that conclusions 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. 2.1 presents Rizzi (1997)'s arguments in favor of splitting up the C domain into several ordered functional projections. 2.2 provides evidence for there being high and low complementizers—two separate complementizers—in the C domain. 2.3 discusses existing accounts of the truncation of infinitives, and provides an update to Rizzi's structure, changing the labels of Rizzi's ForceP and FinP to CP2 and CP1 respectively.

### 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.



Rizzi (1997) argues that two complementizers in Italian, *che* and *di*, are realized in different positions in the C domain: in Force and Fin of (14) 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. For Rizzi, finiteness is to be understood as a very rudimentary specification of mood, tense and agreement in the IP domain. Fin is merely endowed with certain features that allow this aforementioned specification to take place, with no semantics of its own. This is what I will end up redefining.

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.                      b. Il tuo libro ho letto.  
           the your book, it I have.read                      the your book I have.read  
           ‘Your book, I have read it.’                      ‘Your book I have read.’ (but not his)

In this paper, I will assume that FocusP involves focus fronting *solely* for the purpose of contrastive focus, which appears to be the common state of affairs in the cartography literature. This does not preclude the possibility of a projection for focus lower than TopicP, such as WhP.

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.

What constitutes a phase head in (14)? It may vary by language. Given that wh-movement takes place to a position right above FinP, as I will argue later in this next section, it might be assumed that FinP is a phase as well. But there are many contrasts in clausal opacity that can only be captured if ForceP is a phase head but FinP is not. For instance, Carstens and Diercks (2009) observes that FinP is never phasal in Lubukusu. Ultimately, the argumentation in this paper would not be affected by what is phasal in the C domain.

## 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 left 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*; for Rizzi, *di* is a low complementizer in FinP whereas *che* is a high complementizer in ForceP.

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 outside of Romance. Haegeman (2012) notes two such examples below from colloquial English, which involve two instantiations of *that*.<sup>5</sup> For Haegeman, the adjuncts in ital-

<sup>5</sup>Because *that* never behaves as a low complementizer alone in English, it appears that *that* in FinP can only be licensed if *that* is also realized in ForceP.

ics are located in Spec,TopicP. The position of the adjunct sandwiched between two complementizers can be used to distinguish the two *that* in (19a)-(19b) below in a principled way.

- (19) a. She maintained [<sub>ForceP</sub> **that**<sub>high</sub> [<sub>TopicP</sub> *when they arrived* [<sub>FinP</sub> **that**<sub>low</sub> they would be welcomed]]].  
 b. He reminds me [<sub>ForceP</sub> **that**<sub>high</sub> [<sub>TopicP</sub> *in the days of Lloyd George* [<sub>FinP</sub> **that**<sub>low</sub> 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 FinP in this case:<sup>6</sup>

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

A source of confusion might be there is no principled way to distinguish between high and low complementizers. For instance, one might ask the following question: how can you tell apart high and low complementizers? As in (16) and many other examples provided in this paper, high and low complementizers are diagnosed *solely* by their position relative to topics and focalized elements in a principled way. That is how they are defined. Even if they had the same phonetic form, as in (19a)-(19b), we know that one *that* is high and the other is low. This just *is* the principled way of distinguishing between high and low complementizers in a clausal structure: looking at the position of a topic or focalized element relative to the complementizer.

The primary empirical contribution of this paper, then, is remarkably simple. The claim is that Force<sup>0</sup> cannot occur in nonfinite clauses, or infinitives. Section 4 provides further arguments that might help clarify and determine the relationship between high complementizers and finiteness. Topicalization is used to diagnose the presence of PRO vs. pro in Serbian *da*-complements in a principled and systematic way; I find that PRO cannot occur with a high complementizer.

The primary way of telling apart high and low complementizers is by looking at their position in a sentence. But there might be other hints, as well, depending on the syntactic opacity of an embedded clause. For instance, 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:

<sup>6</sup>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*.

- (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. Let us 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]<sub>i</sub>, to read t<sub>i</sub>.

Adger (2007) suggests that *for* is a low complementizer in FinP because it does not allow topics to its left or right. We will see that the reason why (24a) is ruled out is that, unlike Italian, the left periphery of English infinitives is slightly more truncated, lacking a TopicP:

- (24) a. \* I propose, [these books]<sub>i</sub>, for John to read t<sub>i</sub>  
 b. \* I propose for, [these books]<sub>i</sub>, John to read t<sub>i</sub>

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 at least to some extent. But as we will see, the degree of truncation can differ between languages.

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 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* and *if*.<sup>7</sup> Shlonsky and Soare (2011) provides a convincing argument that *why* is base-generated in a position lower than Spec,IntP but 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:<sup>8</sup>

<sup>7</sup>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 that precede the embedded subject must follow *whether* in finite embedded clauses. I take adjuncts in the C domain to be located in Spec,TopicP.

- (ii) Caitlin asked whether under any circumstances she should leave.  
 (iii) \* Caitlin asked whether under any circumstances to leave.

<sup>8</sup>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:



- (25) a. I asked Bill ??why/\*if to serve aubergines.  
 b. I asked Bill why/if 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.<sup>9</sup>

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,FocusP, at least in finite clauses. 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,FocusP.<sup>10</sup> Even in a language where fronted topics are possible in infinitives such as Italian, which also has wh-infinitives, Haegeman (2006) and Bocci (2007) note that focalization for the sake of contrastive focus 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

I take both FocusP and WhP to involve semantic focus, but in different ways. I assume that Spec,FocusP is occupied by moved elements for the purposes of contrastive focus, whereas the lower Spec,WhP is occupied by wh-words for the sake of introducing alternatives, which itself involves focus too, but of a different kind.

Infinitives, as I will argue in the rest of this paper, 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 if and only if its finiteness feature is encoded as + at FinP, following Adger (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. I will assume the structure in (27) throughout this paper:

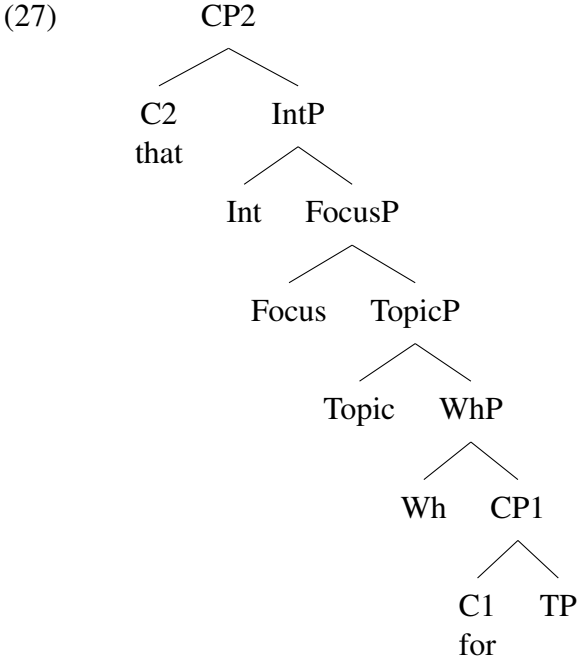
- 
- (i) Why did you ask her to resign?  
 a. What is the reason x, such that for x, you asked her to resign?  
 b. What is the reason x, such that you asked her to resign for that particular reason x?

<sup>9</sup>Shlonsky and Soare (2011) note that some native speakers of English accept (25a). This can be accounted for by assuming that for these speakers, *why* does not obligatorily move to Spec,IntP. *If* is ruled out because it is base-generated in Spec,IntP, and not moved from a lower position.

<sup>10</sup>Henry (1995) notes that Belfast English permits indirect questions introduced by a wh-element that isn't a subject, to the left of *that*:

- (i) I wonder which dish that they picked.

Larsson (2017) notes that several Scandinavian languages allow such constructions. It appears languages may optionally allow a WhP on top of ForceP.



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.

One important issue to address before presenting the survey is the structure in (27) from the perspective of language acquisition. I follow recent work by Satık (2022a,b), which discuss how Rizzi’s left periphery might have evolved, and how it might be acquired by children. Following Cinque and Rizzi (2009), he assumes that almost all of the ordered cartographic structure between C2 and C1 can be derived via general semantic principles, for instance, the relationship between Focus and Topic. But he takes the positioning of Force/C2 and Fin/C1 to be purely syntactic, and hence, not derivable. This, he believes, might in fact be a part of the faculty of language: in other words, the only part of the cartographic blueprint which is included in the innate endowment for natural language. I refer the reader to these papers for further discussion.

### 3 The size of infinitives and empirical generalizations

This section presents a survey of the infinitival left periphery. With the theoretical background established, we are now able to conduct a crosslinguistic survey. 3.1 introduces the reader to the methodology used to investigate the infinitival left periphery in different languages and the results of the survey. Sections 3.2-3.6 describe the different sizes of infinitives attested, while 3.7 provides an interim summary.

#### 3.1 Infinitives can differ in size, but are always truncated

Table 1 below presents the survey that will be discussed in detail in this section. The column PROP in Table 1 is not a traditional left peripheral property, but I will give some reasons for be-

lieving it is one in section 6.2. My goal in section 3 is to discuss only the classic properties.

**Table 1:** The summary of the various properties of the infinitival property in 17 languages.

HIGH C: ✓if the language allows high complementizers in infinitives.

INT: ✓if the language allows Int-properties (*why, if*) in infinitives.

FOCUS: ✓if the language allows contrastive focalized elements in infinitives.

TOPIC: ✓if the language allows topics in infinitives.

WH: ✓if the language allows wh-words in infinitives with embedded scope.

LOW C: ✓if the language allows low complementizers in infinitives.

PROP: ✓if the language has propositional infinitives (complement of *seem, claim* etc.).

! if a lack of the relevant property appears to be a prima facie counterexample.

LANGUAGE	HIGH C	INT	FOCUS	TOPIC	WH	LOW C	PROP
Hindi	X	X			X	X	X
Hungarian	X	X			X	X	X
Serbian	X	X			X	X	X
Turkish	X	X			X	X	X
German	X	X			X	X	✓
Icelandic	X	X	X	X	X	✓	✓
Swedish	X	X	X	X	X	✓	✓
Dutch	X	X	X	X	✓	✓	✓
English	X	X	X	X	✓	✓	✓
French	X	X	X	X	✓	✓	✓
Polish	X	X	X	X	✓	✓	✓
Spanish	X	X	X	X	✓	✓	✓
Catalan	X	X	X	✓	✓	✓	✓
Irish	X	X	X	✓	!	✓	✓
Italian	X	X	X	✓	✓	✓	✓
Hebrew	X	✓	✓	✓	✓	✓	✓
Russian	X	%			✓	!	X

To see the methodology used for this survey, I now provide a quick summary of the properties of the C domain of English infinitives:

- (28)
- 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, which the reader can discern from Table 1. This will be discussed again in 3.7.

There are limits to this methodology, however. Table 1 contains several entries for the Focus and Topic columns which are blank. An astute reader may have noticed that these languages are notable for having a free word order, and hence, scrambling phenomena. These languages have very free word order even within their infinitival clauses. Here is an example. According to Bošković (2004), Russian is a language with scrambling, and its infinitives allow both topicalization and contrastive focus fronting within infinitives, as in (29).

- (29) Ja choču [ZDES<sub>i</sub> byt t<sub>i</sub> (a ne tam)].  
I want here be.INF and not there Russian

Scrambling languages like Russian may have VP-internal focus and topic positions, so the topicalization and focus fronting tests seen in (28) above cannot straightforwardly be carried on to scrambling languages. How do we know whether *zdes* ‘here’ in (29) is truly located in the infinitival left periphery, or if it might involve scrambling within the VP instead? It is usually not possible to verify it. I cannot do so in German, Hindi, Hungarian, Russian, Serbian or Turkish.

However, there are a few scrambling languages—Dutch and Polish—in which it *is* possible to tease apart VP-internal scrambling from topicalization. In Dutch, this is via the infinitival complementizer *om*. In Polish, this is via the topic marker *to*, which has different properties from scrambling.

A few of the languages in this sample appear to have infinitives derived from nouns in some sense, but it does not appear to affect the survey. Turkish has nominalized infinitives, which can be case-marked, as in (30) with accusative case:

- (30) Ayşe [oku-ma-ya] karar ver-di.  
Ayşe read-INF-ACC decision make-PST.3SG  
‘Ayşe decided to read.’ Turkish

Kiss (2002) claims that *-ni*, the infinitive marker in Hungarian, is a nominalizing suffix, responsible for assigning dative case to the experiencer in constructions like (31). She takes it to be a nominalizer because it can appear with  $\phi$ -agreement with the experiencer:

- (31) Fontos [János-nak részt ven-ni\*(-e) a verseny-en].  
important John-DAT participate-INF-3SG the competition-SUP  
‘It is important for John to participate at the competition.’ Kiss (2002), Hungarian

Finally, according to McCloskey and Sells (1988), Irish infinitives are derived from verbal nouns together with the particle *a*, as in (32) below:

- (32) Caithfidh [foighid a bheith againn].  
we.must patience be.INF at.us  
‘We must be patient.’ McCloskey and Sells (1988), Irish

As such, the methodology here is systematic, detailed and careful. Let us begin.

### 3.2 Infinitival complementizers

I will start by pointing out examples of infinitival complementizers (in **bold**). As Sabel (2006) points out, what distinguishes these from true infinitival markers like English *to* is that they do not occur in all infinitival contexts. For instance, Pesetsky (2021) English *for*-infinitives have

their own irrealis semantics. Polish infinitives with *żeby* have a similar subjunctive, irrealis mood. Such elements are often ruled from raising constructions entirely and can only occur in control contexts, indicating that they are complementizers. These are seen in (33a)-(33h) below.

- (33) a. dat zij probeerde [(om) het boek te lezen].  
that she tried in-order the book to read Sabel (2006), Dutch
- b. En Joan ha intentat **de** cantar.  
the John has tried of sing.INF Villalba (2009), Catalan
- c. Il a oublié [**de** nettoyer la chambre].  
He has forgotten of to-clean the room Long (1976), French
- d. Dúirt sé [[duine ar bith a bhí bocht]; **gan** t<sub>i</sub> é a ligean isteach].  
say.PST he person any COMP be.PST poor COMP.NEG him let.INF in  
'He said not to let anybody in who was poor.' Chung and McCloskey (1987), Irish
- e. Maria pensa che Gianni abbia deciso **di** andare.  
Maria thinks that Gianni has decided of go.INF Kayne (1991), Italian
- f. Acabamos [**de** ofrecer se los].  
We-have-just of to-offer-him-them  
'We have just offered them to him.' Lujan (1980), Spanish
- g. Chciałem [**żeby** aprosić Kasię].  
want.1SG.PST COMP.SUBJ invite.INF Kasię Zabrocki (1981), Polish
- h. Rina<sub>i</sub> xadla [(me-) PRO<sub>i</sub> le'acben et Gil].  
Rina stopped (from-) PRO to.irritate ACC Gil  
'Rina stopped irritating Gil.' Landau (2013), Hebrew

Though it is commonly accepted that the languages presented above have infinitival complementizers, determining them for the Scandinavian languages is more controversial. Let us start with Icelandic. Though *að* is often called the infinitival marker in the literature, it does not appear in raising constructions (34a), among others, though it appears in most control constructions (34b).

- (34) a. Hesturinn virðist [hafa týnt knapanum].  
horse.DEF seems have lost jockey.DEF  
'The horse seems to have lost its jockey.' Thráinsson (2007)
- b. Risarnir lofa [að éta ríkisstjórnina á morgun].  
the-giants promise to eat the-government to-morrow  
'The giants promised to eat the government tomorrow.' Thraínsson (1993), Ice.

Following Rizzi's reasoning, this might indicate that they are complementizers instead. At the very least, it is more clear that infinitival *að* is above TP.<sup>11</sup> Sigurðsson (1989) notes that Icelandic has V-to-T (V-to-I in older frameworks) movement in infinitives, unlike English, and the verb still occurs after *að*. Notice that in (35a), the movement of the auxiliary to T precludes the movement of the embedded verb to T, but this is not the case in (35b), 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 (35c).

<sup>11</sup>Thraínsson (1993) cites the data in (35a)-(35c) to indicate that *að* is in AgrSP, a projection in the IP domain above TP. But this was before Rizzi's structure of the left periphery, and this data can be accounted for by assuming infinitival *að* is in CP1.

- (35) 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 stundum [VP t<sub>i</sub> ríkisstjórnir]].  
the-giants say that they eat sometimes governments  
'The giants say that they sometimes eat governments.'
- c. Risarnir lofa [að éta<sub>i</sub> oft [VP t<sub>i</sub> ríkisstjórnir]].  
the-giants promise to eat frequently governments  
'The giants promised to eat governments frequently.'
- Icelandic

Christensen (2007) provides further arguments from negation that infinitival *að* is located in C1. Though discussing these arguments in detail would go out of our scope, I follow Christensen in assuming that *að* and Swedish *att* (36) are both infinitival complementizers.<sup>12</sup>

- (36) Jag har försökt [(att) inte köpa boken].  
I have tried (to) not buy.INF the-book  
'I have tried to not buy the book.'
- Swedish

Moving on, to the best of my knowledge, Hindi, Serbian and Turkish have not been reported to have infinitival complementizers in the literature. Dékány (2017) reports that Hungarian does not have them at any point in its history. In the case of German, Sabel (2006) notes that the phonetically realized complementizer *um* is excluded in German infinitives:

- (37) \* dass sie versuchte [um das Buch zu lesen].  
that she.NOM tried COMP the book to read
- Sabel (2006), German

Russian is the only language in my sample that clearly does not have an overt infinitival complementizer. The subjunctive complementizer *čtoby* is possible in finite clauses, as demonstrated in (38a), but is ruled out completely from all infinitives which are complement clauses, as in (38b). This implies that it is not an infinitival complementizer:

- (38) a. Ivan xočet čtoby Maša pročitala/čitala [Vojnu i Mir]  
Ivan wants that.SUBJ Maša read.PST.PERF/.PST.IMPERF War and Peace  
'Ivan wants for Masha to read War and Peace.'
- Antonenko (2008)
- b. \* Ja choču [čtoby byt zdes].  
I want COMP.SUBJ be.INF here
- Russian

*Čtoby* can occur with infinitives, but only if the infinitival clause is an adjunct.<sup>13</sup> In this case, *čtoby* has a meaning akin to *in order to* in English. According to Jung (2009), these adjuncts can optionally have either an overt subject marked with dative case (39a), or PRO (39b):

- (39) a. On prišel [čtoby ej ne obedat' odnoj].  
he came in-order her.DAT NEG eat.INF alone.DAT  
'He came so that she would not have dinner alone.'

<sup>12</sup>For Christensen, Norwegian and Danish do not have infinitival complementizers, because their infinitival markers can occur in raising constructions unlike Icelandic and Swedish, among other reasons.

<sup>13</sup>By definition, *čtoby* in such cases is not a complementizer. It is therefore not problematic for my generalization, and indeed, my account of finiteness correctly rules it out from argumental infinitives. But it is still puzzling; if finiteness is a matter of clause size. This will be discussed further in the concluding section.

- b. On<sub>i</sub> zašel v magazin [čtoby PRO<sub>i</sub> kupit' maslo].  
 he stopped.by to store in-order buy.INF butter  
 'He stopped by the store in order to buy butter.' Jung (2009), Russian

Jung claims that the source of the dative case in (39a) is in fact a null prepositional complementizer occurring only in certain infinitival constructions, on a par with *for* in English, which can also appear with an accusative case-marked infinitive; when it is not present PRO is required.

- (40) I<sub>i</sub> stopped in order PRO<sub>i</sub> to smoke. vs. I stopped in order for Mary to smoke.

Jung takes this null prepositional element to be the head of Fin, C1 in our terminology, while *čtoby* is treated as an element that occupies the specifier position of a higher projection in the left periphery, on a par with *whether* in English. Given that these are adjuncts, the null element is strictly speaking not a complementizer, but a prepositional adjunct subordinator of some sort.

Dative case-marked arguments appear in other infinitival contexts, as well. It can appear in constructions that represent deontic modality, as demonstrated (41a), and can occur with the infinitival imperative (41b). This indicates that the null prepositional element really can be an infinitival complementizer, given that it need not only appear in adjunct infinitives:

- (41) a. Kuda mne bylo ujtí?                      b. Vsem vstat'!  
 where me.DAT be.PST leave.INF                      all.DAT stand-up.INF  
 'Where did I have to leave for?'                      '(You) all stand up!' Jung (2009)

Under Jung's analysis, the dative case-marked argument in these examples is base-generated as the embedded subject. It moves to Spec,CP1 via an EPP feature to get marked for dative case by the null complementizer, and finally moves to matrix Spec,TP to become the matrix subject. Summing up, it would be difficult to account for the distribution of dative case-marked arguments in Russian infinitives without positing the existence of null infinitival complementizers.

### 3.3 Wh-infinitives

We now move onto the second left peripheral property: whether a language can have a *wh*-word in its infinitive with embedded scope and not matrix scope. For example, one can say *I know what to do* in English, meaning something like 'there is an X such that I know to do X'. This is not possible in a language like Hindi, which does not have true *wh*-infinitives. The English equivalent, in which the *wh*-word has embedded scope is ungrammatical:

- (42) a. \* 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.' Keine (2020), Hindi

Wh-in-situ languages like Hindi do allow *wh*-words in their infinitive, but only if the *wh*-word has matrix scope. As such, (42a) can be acceptable, but only if the *wh*-word has matrix scope, which would be equivalent to *What do you know to do?* in English, which is expected given that Hindi is *wh*-in-situ.

Turkish, Icelandic, Swedish, Irish, Serbian and Hungarian are like Hindi in this regard, as (43a)-(43g) below demonstrate. Given that Icelandic, Swedish and Irish have infinitival complementizers, this may appear surprising. But this just means that the infinitival left periphery is truncated to a greater degree than in English or Italian, but less so compared to Hindi or Turkish in these languages. A structure for the different maximal infinitival sizes will be provided in 3.7.

- (43) a. \* Ahmet Ayşe-ye [PRO ne-yi oku-ma-yı] söyle-di.  
 Ahmet Ayşe-DAT what-ACC read-INF-ACC say-PST.3SG  
 ‘Ahmet told Ayşe what to read.’ Kornfilt (1996), Turkish
- b. \* Ég veit hvað að gera.  
 I know what to do.INF Icelandic
- c. \* Ich weiß nicht [was zu kaufen].  
 I know not [what to buy.INF] German
- d. \* Jag vet inte [vart att gå].  
 I know not where to go Holmberg (1983), Swedish
- e. \* Tá a fhios agam [cad a ithe].  
 I know.PST of.1SG what eat.INF Irish
- f. \* Ne znam šta jesti.  
 NEG eat.1SG what eat.INF Serbian
- g. \* János meg kérdezte mit enni.  
 John VM asked what eat.INF Hungarian

On the other hand, Catalan, Dutch, French, Italian, Polish, Russian, Spanish and Hebrew, as seen in (44a)-(44h) pattern with English.

- (44) a. No sé, d’aquest pernil, on comprar-ne.  
 not know.1SG of.this ham where buy-of.it  
 ‘I don’t know where to buy this ham.’ Villalba (2009), Catalan
- b. Ik weet niet [wie te bezoeken].  
 I know not who to visit.INF Sabel (2006), Dutch
- c. Je lui ai dit [où aller].  
 I him have said where to-go Kayne (1984), French
- d. Gli ho detto [dove andare].  
 Him I told [where go.INF] Kayne (1981), Italian
- e. Janek nie wie [gdzie szukać Marka].  
 Janek not know where seek.INF Marka Zabrocki (1981), Polish
- f. Ja sprosil Ivana kuda bezhat  
 I asked Ivan.ACC where run.INF Russian
- g. No se [qué decirle].  
 not I-know what to-say-him LaPolla (1988), Spanish
- h. ani lo yode’a efo la’avor dira.  
 I not know where to.move apartment Hebrew

There is more to be said about Irish. It appears that wh-infinitives are ruled out from Irish non-finite clauses for independent reasons. Oisín Ó Muirthile (p.c.) has pointed out to me that Irish does not have "pure" wh-words, in the sense that they can stand alone without some kind of copula. A preliminary example of *cé* ‘who’ is given in (45), which is fused with a copula:

- (45) Cé hí?  
 who.COP.PRES she.ACC  
 ‘Who is she?’ Irish





- b. ?? Je pense, ton livre, pouvoir le comprendre.  
I think your book be-able.INF it understand Rizzi (1997), French
- c. 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.' Bocci (2007), Italian
- d. \* Juan niega a María haber-le dado el premio.  
John deny.3SG to Mary to.have.to.her given the prize  
(Intended) 'John denies having given the prize to Mary.' Villalba (2009), Spanish

The possibility of topicalization within infinitives is not limited to CLLD. It is also possible in Irish and Hebrew.<sup>16</sup> In Irish (50a), we see that topicalization must take place to a position preceding the infinitival complementizer, indicating that it is low and not high, as the ISG predicts.

- (50) a. Dúirt sé [[duine ar bith a bhí bocht]<sub>i</sub> gan *t<sub>i</sub>* é a ligean isteach].  
say.PST he person any COMP be.PST poor COMP.NEG him let.INF in  
'He said not to let anybody in who was poor.' Chung and McCloskey (1987), Irish
- b. ani roce [et ugat ha pereg]<sub>i</sub> lenasot *t<sub>i</sub>*.  
I want DOM cake the poppyseed to.try  
'I want to try the poppyseed cake.' Shlonsky (2014), Hebrew

It is not possible in Dutch, Icelandic, Polish or Swedish:

- (51) a. \* dat zij probeerde [[het boek]<sub>i</sub> (om) *t<sub>i</sub>* te lezen].  
that she tried the book in-order to read Sabel (2006), Dutch
- b. Risarnir lofa [(\*[á morgun]<sub>i</sub>) að éta ríkisstjórnina *t<sub>i</sub>*].  
the-giants promise to-morrow to eat the-government Icelandic
- c. \* Jag har försökt [boken<sub>i</sub> (att) inte köpa *t<sub>i</sub>*].  
I have tried the-book (to) not buy.INF Swedish
- d. \* Chciałem [Kasię<sub>i</sub>, to żeby aprosić *t<sub>i</sub>*].  
want.1SG.PST Kasię TOP COMP.SUBJ invite.INF Polish

These data confirm the generalization in (48) above, given that all of the languages in which topicalization is allowed, except Irish, have *wh*-infinitives and infinitival complementizers. I take this to be evidence for the ordering TopicP > WhP > CP1 under Rizzi's framework.

### 3.5 Contrastive focus fronting within infinitives

Although CLLD and contrastive focus fronting may seem like similar phenomena, surprisingly languages seem to distinguish between the two. In almost all of the languages we have seen in 3.4, focus fronting within infinitives is impossible or at least very degraded:

- (52) a. \* Je pense pouvoir TON LIVRE comprendre (pas ton magazine).  
I think be-able.INF your book understand not your magazine French
- b. \* Luis quiere CERVEZA beber (y no sidra).  
Luis wants BEER to.drink (and not cider) Fernández-Sánchez (2016), Spanish

<sup>16</sup>Landau (2004a) points out that Hebrew does not have productive scrambling, so the possibility of this example cannot be attributed to VP-internal topic or focus positions.

- c. \* En Joan es pregunta, EL SOPAR, on fer (no el dinar).  
 the John to himself asks the dinner where to.make not the lunch  
 (Intended) ‘John is wondering where to eat dinner (not eat lunch).’ Catalan
- d. ?? 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
- e. ?? Dúirt sé [[DUINE AR BITH A BHÍ BOCHT]<sub>i</sub> gan t<sub>i</sub> é a ligan  
 say.PST he person any COMP be.PST poor COMP.NEG him let.INF  
 isteach], ach [duine ar bith a bhí saibhir] a ligan.  
 in but person any COMP be.PAST rich let.INF  
 ‘He said not to let anybody in who was poor, but to let anyone in who was rich.’

The only language in my sample which allows contrastive focus fronting that is not a scrambling language is Hebrew:

- (53) ani roce [ET UGAT HA PEREG]<sub>i</sub> lenasot t<sub>i</sub> (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).’ Shlonsky (2014), Hebrew

The possibility of focus fronting in Hebrew, together with its impossibility in every other language without scrambling, provides some evidence for the ordering FocusP > TopicP > WhP > CP1. Though it is not as strong as one would like, given that I only have one piece of evidence for this in Hebrew. For this reason, I will move on to the next part of the survey.

### 3.6 Why and if within infinitives

The second to last step is to determine whether Rizzi’s ordering IntP > FocusP > TopicP > WhP > CP1 is correct. This would imply the truth of the following generalization:

- (54) If a language allows *why* and *if* in its infinitives, then it has contrastive focus and topicalization within its infinitives, wh-infinitives and infinitival complementizers.

I will primarily present *why*-infinitives here. But if a *why*-infinitive is possible in a given language, I will also present data from *if*-infinitives. Because an example of a *why*-infinitive alone is not sufficient evidence to show that *why* has moved to Spec,IntP.<sup>17</sup> This because it could in fact be located in a low position, as Shlonsky and Soare (2011) note, this is possible even with some native speakers of English. To help eliminate this possibility, I will also show that *if*-infinitives are possible in such languages because *if* is base-generated in Spec,IntP, at least in English.

The vast majority of languages in my survey disallow *why*-infinitives.<sup>18</sup> In my experience, the judgments are clearer than in English, likely because of the dialectal variation Shlonsky & Soare note. Some examples are given below in (55a)-(55f):

- (55) a. \* En Joan es pregunta perquè fer el sopar  
 the John to himself asks why to.make the dinner Catalan

<sup>17</sup>Thanks to Susi Wurmbrand for pointing this out.

<sup>18</sup>For space, I will only show data from languages which already allow wh-infinitives, given that *why* would already be ruled out from non-wh-infinitive languages. But I have verified this for all languages studied in the survey.

- |    |  |         |
|----|--|---------|
| b. | * Maria vroeg [waarom om pizza te eten].<br>Maria asked why in-order pizza to eat    | Dutch   |
| c. | * Je lui ai dit pourquoi manger une pizza<br>I her AUX asked why eat.INF a pizza     | French  |
| d. | * Maria ha chiesto perché andare.<br>Maria AUX asked why go.INF                      | Italian |
| e. | * Ana pidió por qué comer pizza.<br>Ana asked why to.eat pizza                       | Spanish |
| f. | * Janek nie wie dlaczego/jeśli szukać Marka]<br>Janek not know why/if seek.INF Marka | Polish  |

Hebrew is the only language that unambiguously allows both *why* and *if*-infinitives. Though (56a) is from Shlonsky (2014), he merely states *if*-infinitives are possible in Hebrew. I verified his claim in (56b) below:

- |      |    |   |                         |
|------|----|---|-------------------------|
| (56) | a. | ani lo mevin lama la'avor dira.<br>I not understand why to.move apartment<br>'I don't understand why to move apartments.' | Shlonsky (2014), Hebrew |
|      | b. | ani lo yode'a im la'avor dira.<br>I not know if to.move apartment<br>(Literally) 'I don't know if to move apartments.'    | Hebrew                  |

The possibility of IntP in Russian infinitives appears to be subject to dialectical variation. I have had one native speaker of Russian and 2 native speakers of Ukrainian accept both of the sentences in (57a)-(57b) below, but another did not at all. For this reason, I marked Russian with % on Table 1.

- |      |    |  |    |   |
|------|----|--|----|---|
| (57) | a. | % Ja sprosil Ivana zachem bezhat<br>I asked Ivan.ACC why run.INF | b. | % Ja sprosil Ivana esli bezhat<br>I asked Ivan.ACC if run.INF |
|------|----|--|----|---|

Though Russian has scrambling and we cannot directly verify the presence of TopicP and FocusP, Russian unambiguously allows *wh*-infinitives, in addition to the constructions in which dative case-marked arguments appear in infinitives, which likely involve a null infinitival complementizer. Altogether, we have enough evidence for (54), and hence evidence for Rizzi's cartographic ordering presented in section 2, example (27) above.

### 3.7 Interim Summary

We've thus far seen that, with some language-specific exceptions, the infinitival left peripheral properties of a language with infinitives are predictable based on the highest possible left peripheral property in the cartographic structure I presented in (27) above:

- |      |    |  |
|------|----|--|
| (58) | a. | <i>Sabel's (2006) Generalization</i> : If a language has <i>wh</i> -infinitives, then it also has infinitival complementizers.             |
|      | b. | If a language allows topicalized elements within its infinitives, then it also has <i>wh</i> -infinitives and infinitival complementizers. |

- c. If a language allows *why* and *if* in its infinitives, then it has contrastive focus and topicalization within its infinitives, wh-infinitives and infinitival complementizers.

Apart from two cases, these generalizations were maintained without issue. These two cases are less clear, but in my view not problematic. First, Irish does not have wh-infinitives despite having infinitives with topicalized objects, but I argued that this was due to the independent nature of wh-elements in Irish. Second, though Russian does not have an overt infinitival complementizer, I argued it has a null one.

I would like to propose to capture these generalizations by assuming that infinitives can come in different *maximal* sizes across languages which have infinitives at all. For instance, given that English allows wh-infinitives but not all the other properties in the hierarchy, the maximal size of an English infinitive would be WhP. I therefore classify the languages presented in Table 1 into the groups presented in (59) below.

- (59) **Hierarchy:** CP2 > IntP > FocP > TopP > WhP > CP1 > TP
- a. **Maximally TP Infinitives:** Hindi, Hungarian, Serbian, Turkish, German
  - b. **Maximally CP1 Infinitives:** Icelandic, Swedish
  - c. **Maximally WhP Infinitives:** Dutch, English, French, Polish, %Russian, Spanish
  - d. **Maximally TopP Infinitives:** Catalan, Irish, Italian
  - e. **Maximally IntP Infinitives:** Hebrew, %Russian
  - f. **Maximally CP2 Infinitives:**  $\emptyset$

What is crucial is that we never see infinitives occurring with high complementizers—the Infinitive Size Generalization (ISG). Is this really the case? This is what we will now discuss.

## 4 Can infinitives ever have high complementizers?

In this section, I discuss further evidence for the Infinitive Size Generalization. In the literature, complementizers like *that* in English and *che* in Italian are called *finite* complementizers, and this presupposition precludes them from occurring in infinitives. This already accounts for the vast majority of languages in Table 1. But not all of them. I will now discuss apparent counterexamples from Middle English and Scandinavian which are not, in fact, problematic to my account.

I then present further evidence in favor of the ISG from Serbian *da*-constructions and Mandarin; these examples do not involve infinitives strictly speaking, but may involve a finiteness distinction. This supports the idea that embedded clauses without high complementizers are truncated in size, and have properties one would expect nonfinite clauses to have, such as requiring a PRO subject or allowing restructuring phenomena.

The only direct contradiction of the ISG I am aware of is presented by van Gelderen (1998), who claims that Middle English infinitives project ForceP. According to van Gelderen (1998), it is possible for *ai* in (60) 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:

- (60) Til [all our 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 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.

The only case I am aware of in which elements that appear to have the same phonetic form as a high complementizer can occur in infinitives involves the Scandinavian languages. In Icelandic orthography, for instance, both the finite complementizer and infinitival marker share the same phonetic form *að*. Is this a problem for my account?

I think not, for two reasons. First, the two *að* have very different properties. In (61b) below, we find that finite complement clauses with *að* allow internal topicalization. But infinitives with *að* do not allow internal topicalization at all, as (61a) shows. neither to the left or right of *að*:

- (61) a. Risarnir segja [að [á morgun]<sub>i</sub> éti þeir ríkisstjórnina t<sub>i</sub>].  
 the-giants say that tomorrow eat they the-government  
 'The giants said that they will eat the government tomorrow.'
- b. Risarnir lofa [(\*)[á morgun]<sub>i</sub>] að [(\*)[á morgun]<sub>i</sub>] éta ríkisstjórnina t<sub>i</sub>.  
 the-giants promise to-morrow to to-morrow eat the-government  
 'The giants promised to eat the government tomorrow.' Thraínsson (1993)

This indicates that *að* in finite clauses is a high complementizer, but not in infinitives. It must instead be a low complementizer (or something else) in infinitives.

The second piece of evidence is that this similarity is in fact the fault of the orthography of the Scandinavian languages. Holmberg (1986) points out that the infinitival marker is not pronounced the same as the finite complementizer in any of the Scandinavian languages, except in slow and formal speech. *Att* in Swedish is pronounced /o/ for instance, while in finite clauses the complementizer is pronounced /at/. The infinitival marker is instead derived from a preposition.

Let us look at some constructions which aren't infinitives. I would expect that, given my definition of finiteness, it would be possible to extend observations on truncated clause size to nonfinite constructions in general, and not just infinitives. And this does appear to be the case. Perhaps the strongest evidence in favor of using topicalization to diagnose non-finiteness comes from Serbian *da*-constructions. Though we've seen that Serbian infinitives are highly truncated, Serbian complement clauses also allow another construction with a complementizer-like element *da*, with agreement on the embedded verb.

Both the infinitival form of the verb in addition to the *da*-form are allowed in the complement of *decide*, as demonstrated in (62a). The infinitival form of the verb is not allowed in the complement of *claim*, as seen in (62b), indicating that the complement must be finite.

- (62) a. Odlučila sam {da čitam / čitati} ovu knjigu.  
 decided.SG.F AUX.1SG DA read.1SG / read.INF.IPFV this book  
 'I decided 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

Although the subject in (62a) is null, it need not be. As (63) shows, the complement of *decide* may allow an overt embedded subject.

- (63) Jovan je odlučio da  $\emptyset$ /Petar/on ode.  
 Jovan AUX decided DA  $\emptyset$ /Petar/he leaves  
 (Potential reading 1) ‘Jovan decided to leave.’  
 (Potential reading 2) ‘Jovan decided that Peter/he would leave.’ Serbian

At this point, we do not have enough information to determine whether the possible empty category in (63) is *pro* or *PRO*. To figure this out, I have determined that complement of *decide*, in fact, in certain cases does not allow overt subjects. This can be teased apart via clause-internal topicalization—the key empirical test of this paper.

One possibility might be that there are two locations of *da*. It may be a high complementizer, or it may be a low complementizer:<sup>19</sup>

- (64) CP2 (the location of *da* in (62b) > TopicP > CP1 (the potential location *da* in (62a))

It is the topmost C2 head in which *da* is located an example like (62b), so we would expect it to be required for it to precede clause-internal topics. This is the case:

- (65) a. \*Tvrdim sam [ovu knjigu]<sub>i</sub> da čitam t<sub>i</sub>.  
 claim.SG.F AUX.1SG this book DA read.1SG  
 ‘I claimed to be reading this book.’  
 b. Tvrdim sam da [ovu knjigu]<sub>i</sub> čitam t<sub>i</sub>. Serbian

Similarly, if *da* is located in C1 in (62a), we would expect it always be preceded by clause-internal topics. This prediction is partly borne out. According to Todorović and Wurmbrand (2016), *decide*-complements allow topicalization both before and after *da*. I take this to show that *da* can be located either in C1 or C2; in (66a) it is located in C1 while in (66b) it is in C2:

- (66) a. Odučila sam [ovu knjigu]<sub>i</sub> da čitam t<sub>i</sub>.  
 decided.SG.F AUX.1SG this book DA read.1SG  
 ‘I decided to read this book.’  
 b. Odučila sam da [ovu knjigu]<sub>i</sub> čitam t<sub>i</sub>. Serbian

This being the case, I predict that (66a) can only allow *PRO*, because it is truncated in size, while (66b) can allow overt subjects. This is borne out. It turns out that when a topicalized element precedes *da* as in (66a), overt NPs are disallowed, as shown in (67a). Only a null and obligatorily controlled subject is allowed. And as predicted, when a topicalized element follows *da*, it allows for an overt NP, as in (67b).<sup>20</sup>

- (67) a. \*Odučila sam ovu knjigu da čita Ivan.  
 decided.SG.F AUX.1SG this book DA read.3SG Ivan  
 (Intended reading) ‘I decided for Ivan to read this book.’  $\text{CP} > \text{TopicP} > \text{TP}(\text{da})$

<sup>19</sup>This is simplified from the analysis proposed by Wurmbrand et al. (2020), because *da* is analyzed as a lower clausal head, such as the head of T or v, rather than a low complementizer. Assuming that *da* can be a low complementizer is sufficient for the purposes here and does not change the result. I suspect that a similar analysis can be applied to *shuo* in Mandarin.

<sup>20</sup>One complication in the data in (67a)-(67b) is that my speakers preferred to topicalize the verb above the embedded subject *Ivan*. I take this to involve some kind of verb-medial focalization or topicalization. I am not sure in what way this would affect the data, if at all.

- b. Odlučila sam da ovu knjigu čita Ivan.  
 decided.SG.F AUX.1SG DA this book read.3SG Ivan  
 ‘I decided for Ivan to read this book.’ CP (da) > TopicP > TP

In other words, in (67a), when the clause is deficient in size as a result of *da* being located in C1, no overt subjects are allowed. However, when *da* is located in C2—indicating that the clause is not deficient—overt subjects are once again allowed. This, again, strongly indicates that there is a relationship between subject licensing and clause size, diagnosed via topicalization:

- (68) *Serbian Generalization:*  
*Da* not in C2 → PRO obligatory  
*Da* in C2 → PRO not permitted, overt subjects or pro required

This is strong evidence for my account in terms of finiteness being a matter of clause size.

Mandarin has a similar pattern to what we see in Serbian *da*-complements. Huang (2018) makes a similar argument in Mandarin, and his analysis can be straightforwardly translated into mine. As Huang (2018) convincingly shows, *shuo* behaves as a finite complementizer (in our terminology *high*) when it heads a finite embedded clause. In (69), topicalization is only allowed within the embedded clause, because the complement of *believe* must be finite.

- (69) a. Wo xiangxin [shuo Lisi [zhe-pian baogao]<sub>i</sub> xie-wan-le t<sub>i</sub>].  
 I believe SHUO Lisi this-CL report write-finish-PFV  
 ‘I believe that Lisi has written this report.’  
 b. \* Wo [zhe-pian baogao]<sub>i</sub> xiangxin [shuo Lisi xie-wan-le t<sub>i</sub>].

But *shuo* behaves as a lower complementizer when it heads a nonfinite embedded clause, such as the complement of *try*, with which the pattern in (69b) is possible. The complement of *try* in (70), 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:

- (70) Wo [zhe-pian baogao]<sub>i</sub> hui shefa [shuo jinkuai xie-wan t<sub>i</sub>].  
 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 (69b). But restructuring—removal of the CP2 layer—allows for this movement to take place, as in (70). 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:

- (71) a. Xiaoming<sub>i</sub> xihuan (\*ta<sub>i/j</sub>) chi shousi. b. Xiaoming<sub>i</sub> xiwang (ta<sub>j</sub>) chi shousi.  
 Xiaoming like he eat sushi Xiaoming hope he eat sushi  
 ‘Xiaoming likes to eat sushi.’ ‘Xiaoming hopes to eat sushi.’

We can draw this conclusion regarding Mandarin:

- (72) *Mandarin Generalization:*  
*Shuo* not in C2 → PRO obligatory, restructuring permitted  
*Shuo* in C2 → PRO not permitted, restructuring not permitted



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

Of course, one might 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 the following potential finiteness generalization:

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

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.

- (74) a. A clause is finite iff it is untruncated in the C domain.  
b. A clause is nonfinite iff 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 merely 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 Icarí Dargwa according to Kalinina and Sumbatova (2007), for example.

In the next section, we will determine whether this definition of finiteness still holds once we consider a range of facts.

## 5 Potential counterexamples to the Infinitive Size Generalization

Although I have presented arguments in favor of the ISG, it is not surprising that many apparent counterexamples of it already exist in the literature. Some of these possibilities are listed below:

- (75) a. *That*-less embedded clauses have been argued to be truncated in the C domain.  
b. Languages like Russian have infinitives that are fully opaque to cross-clausal syntactic operations.  
c. Factive embedded clauses have been claimed to be truncated in the C domain like infinitives.

The goal of this section is to argue that none of these constitute true counterexamples for the generalization. 5.1 offers two solutions to the problem in (75a): infinitives may be more truncated than merely CP2, or *that* may simply be null in them. In 5.2, I assume Keine (2020)'s observation of selective opacity to explain the possibility in (75b). 5.3 concludes that factives are in fact not truncated in the C domain.

## 5.1 *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 whether *that*-less embedded clauses have a CP2/ForceP layer or not, which could lead to a confound.<sup>21</sup> For example, Bošković and Lasnik (2003) notes the following contrast, in which (76d) cannot occur without the high complementizer but (76b) can:

- (76) 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. (76b) is uncontroversially finite, but if it truly lacks CP2, this is a counterexample to my definition of finiteness.

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 (77a)-(77b) on one hand and (77c)-(77d) on the other, Wurmbrand (2017) claims that stripping of embedded clauses is only possible when the embedded clause lacks a CP2.

- (77) 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 (77a)-(77b) and (77c)-(77d), 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. As such, 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.

Here are two potential paths one can take. 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

---

<sup>21</sup>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).

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 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 (2004b). This is shown in (78a)-(78c) 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.

- (78) 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<sub>i</sub> še-pro<sub>i</sub> 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 projection together with CP2 that is truncated when a nonfinite clause is made. Let us call this layer IndicativeP (IndP), given the indicative syntax and semantics of (78c). Admittedly, this is nothing more than merely restating the pattern in (78a)-(78c), but for my purposes this is sufficient. One possibility is that CP2 is necessarily deleted whenever IndP is deleted. In other words, one could define nonfinite clauses as follows:

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

Given that IndP is a mere stipulation, I must leave open to future research as to what this layer really is.

## 5.2 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 (80a)-(80b) below:

- (80) 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<sub>i</sub> kažutsja [t<sub>i</sub> 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:

- (81) 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. Keine convincingly shows that selective opacity is a pervasive phenomenon. 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. This will be discussed further in 6.2.

### 5.3 Factives are not truncated in the C domain

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:

- (82) \* 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 also present some novel evidence for her account.

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 non-factives are not, as seen below.

- (83) a. How do you suppose that Maria<sub>i</sub> fixed the car t<sub>i</sub>?  
 b. \* How did you notice that Maria<sub>i</sub> fixed the car t<sub>i</sub>?  
 c. Why does Mary<sub>i</sub> think that Bill left the company t<sub>i</sub>?  
 d. \* Why does Mary<sub>i</sub> regret that Bill left the company t<sub>i</sub>?

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 to assume that it is Merged onto Spec,TopP:

(84) [TopP Last week, [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:

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

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

(86) 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:

(87) 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:

- (88) 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.

## 6 Further cartographic generalizations

Though section 3 presented a few cartographic generalizations, I believe that a few more promising ones remain. I discuss one based on *tough*-constructions in 6.1 and another one on the possibility of infinitives with propositional semantics in 6.2.

### 6.1 *Tough*-constructions

In the literature, *tough*-movement in English and *wh*-movement have been claimed to take place to a position in the infinitival left periphery. Given this, one would expect something like the following generalization to be true:

- (89) If a language has a *tough*-construction, then it has *wh*-infinitives or infinitival complementizers.

I will propose in this section that this generalization is likely to be true, though an in-depth survey in future work is necessary to ensure that it is.

What is a *tough*-construction? Here are examples from English:

- (90) 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)

Chomsky (1977b) convincingly shows that *tough*-movement involves a step of *wh*-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 (91a)-(91d). The middle Spec,CP position was occupied by a Copy of *what sonatas* prevents *this violin* from moving up in (91d).

- (91) 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 my 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:

- (92) Caitlin<sub>i</sub> is tough [<sub>WhP</sub> Op<sub>i</sub> [<sub>TP</sub> PRO<sub>arb</sub> to please t<sub>i</sub>.]]

We would expect languages like German to not have *tough*-constructions (TCs), given they do not contain an infinitival left periphery.<sup>22</sup> But according to Comrie (1997) among others, German appears to have TCs. This is contradictory for Chomsky's account. Where would it move to?

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 (93a) and do not license parasitic gaps (93b). This is because German "*tough*"-constructions do not involve A'-movement, unlike in English.

- (93) 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<sub>i</sub> [ohne vorher *pg*<sub>i</sub> zu kaufen] schwer t<sub>i</sub> 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. Thus, we would predict languages with TCs to have *wh*-infinitives and/or infinitival complementizers, as I proposed at the start of this subsection.

German is the odd one out: it is the only language that has been reported to have TCs but does not have an infinitival complementizer or *wh*-infinitives. Out of the other languages, Icelandic (Sigurðsson (2016)), Swedish (Klingvall (2018)) Spanish, French, Italian (Hartman (2011)) and Dutch (van der Auwera and Noel (2011)) are all reported to have true *tough*-constructions, which do not involve long A-movement.

<sup>22</sup>Although the next subsection argues that this is strictly speaking false, I do not take PropP to have a specifier position to which A'-movement is possible. It is present purely for semantic purposes.

Outside of my survey, I have verified that wh-infinitives exist in Tamil according to Schiffman (1999), and it too has English-style TCs (Selvanathan (2017)). 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 believe this covers most, if not all, of the languages which have been reported to have English-style TCs in the literature.

Future work could determine whether Norwegian and Danish pattern like German *leicht*-constructions, given that according to Christensen (2007) they have true infinitival markers, unlike Swedish. Finally, Selvanathan (2018) reports that Malay involves *leicht*-constructions, so I would predict that it should not project the infinitival left periphery. These are all encouraging lines of inquiry for future work to look into.

## 6.2 Propositional semantics in the left periphery

In this section, I will discuss whether the following cartographic generalization is tenable.

- (94) If a language has infinitival complementizers, then it has propositional infinitives.

I will conclude that it is likely for it to be true, though future research will be needed to once again explain why Russian appears to lead us astray.

First, I will present a definition of what it means for an infinitive to have propositional semantics. Following Wurmbrand and Lohninger (2019) (W&L), I assume that infinitival complements can come in four sizes: vP, TP and CP. I will show that the ISG is compatible with this approach. 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.

I do not need get into the semantics of vP-infinitives. But CP-complements involve those which can be assigned a truth value, while TP-complements do not (95a)-(95b). On the other hand, TP-complements have a future-irrealis reading with respect to matrix tense, whereas CP-complements are read with tense that is simultaneous to the matrix tense (95c)-(95d):

- (95) a. Caitlin claimed to have eaten salad, which is true.  
 b. # Caitlin decided to eat salad tomorrow, which is true.  
 c. Caitlin decided to fly tomorrow.  
 d. \* Caitlin claimed to be happy tomorrow.

Examples of infinitives with a propositional semantics in English are the complements of *claim* seen above, in addition to the complement of many raising predicates such as *seem*.

Now, we would like to know why W&L place the propositional semantics of such infinitives into the C domain in particular, rather than somewhere lower. Wurmbrand (2001) notes a distinction between vP- and TP-infinitives like *try* and *decide* on one hand, and *claim* on the other, regarding the restructuring phenomenon known as *splitting* in German. In cases of splitting, a matrix element (the matrix subject in bold in (96a)-(96b)) can occur between material from the embedded complement. This is seen in (96a)-(96b) below.

The infinitival complement can occur to the left of the matrix verb, so pronoun fronting across the matrix subject is used in (96a). We see that in (96a) the matrix subject is sandwiched between

the embedded object and the infinitival verb. Although the complement of propositional complement usually occurs after the matrix verb, splitting isn't possible at all in (96b), in either position for the infinitival verb:

- (96) a. weil ihn **der Hans** zu reparieren versuchte/beschloss  
 since it.ACC the John to repair tried/decided  
 'since John tried/decided to repair it'
- b. \* weil sie **der Hans** {zu mögen} behauptete {zu mögen}.  
 since her the John {to like} claimed {to like}  
 'since John claimed to like her' German

Wurmbrand takes this to indicate that the complement of *claim* has more structure than that of *decide* or *try*, which are TP- and vP-infinitives respectively. A natural suggestion is to suppose that the complement of *claim* is a CP, blocking this kind of restructuring phenomenon.

Indeed, German is unique in my sample in that it is the only language without infinitival complementizers that can have infinitives with propositional semantics. Raising is seen in (97):

- (97) Er scheint [intelligent zu sein].  
 he seems intelligent to be German

On one hand, German lacks any of the classic properties of the C domain. On the other, the lack of some restructuring in propositional infinitives in German indicates that there might be some of the C domain present. This is puzzling for my account. My proposal is to assume the presence of a very low projection in the left periphery, PropositionP (PropP) to explain how the contrast in (96a)-(96b) is possible.

- (98) **Hierarchy:** CP2 > IntP > FocP > TopP > WhP > CP1 > PropP > TP
- a. **Maximally TP Infinitives:** Hindi, Hungarian, Serbian, Turkish
- b. **Maximally PropP Infinitives:** German

Should we classify Hindi, Hungarian, Serbian and Turkish as PropP languages as well? They do not seem to have propositional infinitives of any kind. Wurmbrand et al. (2020) demonstrate this for Serbian, as we saw in section 4. According to Bhatt (2006), Hindi lacks raising predicates like that of English entirely. Although Szabolcsi (2009) demonstrates that Hungarian has raising predicates, *seem* can only take finite complements. Finally, in Turkish, propositional complements must be gerunds, as in (99), and cannot be infinitival:

- (99) Deniz [kapı-yı ac-tiğ-i-ni] iddia et-ti.  
 Deniz door-ACC open-GER-POSS.3SG-ACC claim AUX-PST.3SG  
 'Deniz claimed that he opened the door.' Turkish

We are now ready to go back to our purported generalization. I would now like to show that each of the other languages, with the exception of Russian, have propositional infinitives, with the raising examples seen in (100a)-(100j) below:

- (100) a. En Joan sembla [estar content].  
 the John seems to.be happy Catalan
- b. Maria lijkt blij te zijn.  
 Maria seems happy to be Dutch



- c. Chloé semble être heureuse  
Chloé seems to.be happy French
- d. Caithfimid<sub>i</sub> [foighid a bheith againn<sub>i</sub>]  
We-must patience be-INF at-us  
'We must be patient.' McCloskey and Sells (1988), Irish
- e. Mi sembra di essere felice.  
I seem of be.INF happy Italian
- f. ha-hafgana omedet lehitkayem be-yom šiši  
the-demonstration.SF stands.SF to.occur in-day sixth  
'The demonstration is about to take place on Friday.' Melnik (2015), Hebrew
- g. Ana parecía [beber demasiado].  
Ana seemed to.drink too.much Spanish
- h. Hesturinn virðist [hafa týnt knapanum].  
horse.DEF seems have lost jockey.DEF  
'The horse seems to have lost its jockey.' Thráinsson (2007), Icelandic
- i. Piotr wydawał się być niespokojny.  
Piotr.NOM seemed be.INF uneasy Przepiórkowski and Rosen (2004), Polish
- j. Jag verkar vara glad.  
I seem be.INF happy Swedish

Once again, the sore thumb in my sample is Russian. To the best of my ability, I have not been able to find any examples of propositional control complements. As noted before in (80a)-(80b), Russian has no raising out of infinitives at all. At the very least, there appears to be a strong correlation between the presence of the C domain and propositional infinitives. But one wonders if the generalization might be saved.

I believe that this is possible, following Keine (2020)'s work on selective opacity. Here is how the solution would work. In German, PropP appears to be transparent for A-movement (raising), control and opaque for splitting. But given that different projections can vary in their opacity for different operations across languages, there is no need to suppose that the opacity of PropP remains the same across all languages.

One could assume that PropP could in principle be present in Russian infinitives, but cannot be because PropP is opaque to everything in Russian: both A-movement and control. It is more difficult to see control as an operation on a par with A-movement, but one could follow Landau (2001, 2015) in assuming that control is preceded by the syntactic operation Agree, which PropP is opaque to in Russian. This is sufficient to explain the case in Russian. But this must be explored in future work; my goal has been to discuss something that appears to be promising.

## 7 Conclusion

This paper has been an investigation on the size of infinitives. After laying the groundwork for this endeavor in sections 1-2, sections 3 presented evidence that the size of infinitives can vary across languages. Section 4 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/ForceP layer. I have defended this approach in further detail in section 5 against potential objections, and presented further avenues of inquiry in section 6.

But much remains open for future investigation. I have not discussed the nature of imperatives like "*Catch her!*" and how they come into being. But it is natural to suppose that they are missing many functional projections, leading to a truncated structure. Indeed, there are other kinds of structures that are often associated with nonfiniteness and/or a truncated structure, 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, which I have shown are highly truncated.

What is more promising, though, is the potential application of this theory of finiteness to subjunctive structures, which seem to be unambiguously headed by CP2—at least sometimes. Though English subjunctives are headed by *that*, they still appear to be truncated in the C domain:

- (101) 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 (78a)-(78c) 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-occurring 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.

Another line of research that would be worth pursuing is looking at the clausal size of adjunct infinitives. Another reason why subjunctives are puzzling is because of the Russian subjunctive complementizer *čtoby*. I noted in section 3.1 that it is ruled out from infinitival complements entirely (102a), but it can occur in infinitives which are adjuncts (102b):

- (102) a. \*Ja choču [čtoby byt zdes].  
 I want COMP.SUBJ be.INF here Russian  
 b. On<sub>i</sub> zašel v magazin [čtoby PRO<sub>i</sub> kupit' maslo].  
 he stopped.by to store in-order buy.INF butter  
 'He stopped by the store in order to buy butter.' Jung (2009), Russian

The ISG does make the correct prediction here. By definition, a complementizer can only occur in complement clauses; *čtoby* in adjuncts is not strictly speaking a complementizer. But it is still puzzling for my account of finiteness, as one would expect it to be ruled out regardless, given that finiteness is a matter of clause size. For this reason, it would be useful to look at the syntax of *čtoby* in more detail. Perhaps it has the same syntax as English *in order* in adjunct infinitives. Jung (2009) assumes a solution that is compatible with my account: *čtoby* is not a true complementizer, but rather an element that occupies a specifier position in the higher left periphery. Further evidence would help determine what the right analysis for *čtoby* is.

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 non-finite clauses. It does not seem coincidental that the cartographic generalizations noted in this paper appear to be attested in the vast majority, if not all, of the cases in the survey. The inability of infinitives 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. In *Finiteness: empirical and theoretical foundations*, ed. Irina Nikolaeva. Oxford: Oxford University Press.
- Antonenko, Andrei. 2008. Puzzles of Russian Subjunctives. In *UPenn Working Papers in Linguistics*, volume 14.2. Philadelphia: Penn Graduate Linguistics Society.
- Arregi, Karlos. 2003. Clausal pied-piping. *Natural Language Semantics* 11:115–143.
- van der Auwera, Johan, and Dirk Noel. 2011. Raising: Dutch between English and German. *Journal of Germanic Linguistics* 23:136.
- 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*. Berlin: Walter de Gruyter.
- Bird, Steven. 1999. Multidimensional exploration of online linguistic field data. In *North East Linguistics Society*, ed. Tamanji Pius, Masako Hirotsu, 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. *Nanzan Linguistics* 3.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Massachusetts: MIT Press.
- Bošković, Željko. 2004. Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry* 35:613–638.
- 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*. Foris: Dordrecht.
- Carstens, Vicki, and Michael Diercks. 2009. Parameterizing case and activity: Hyper-raising in Bantu. In *Proceedings of North Eastern Linguistic Society 40*. Amherst, MA: GLSA Publications.
- Chomsky, Noam. 1977a. Conditions on transformations. In *Essays on form and interpretation*, 81–162. New York, NY: 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.
- Christensen, Ken Ramshøj. 2007. The infinitive marker across Scandinavian. *Nordlyd* 34.
- Chung, Sandra, and James McCloskey. 1987. Government, barriers, and small clauses in Modern Irish. *Linguistic Inquiry* 18:173–237.
- Cinque, Guglielmo, and Luigi Rizzi. 2009. The Cartography of Syntactic Structures. *Manuscript*

- Comrie, Bernard. 1997. Tough movement and its analogs in Germanic languages. In *Language and its Ecology*. Berlin: Mouton de Gruyter.
- Dékány, Éva. 2017. Anti-agreeing infinitives in Old Hungarian. In *Infinitives at the syntax- semantics interface*, 193–222. Berlin: De Gruyter.
- Doherty, Cathal. 2000. *Clauses without that: The case for bare sentential complementation in English*. New York: Garland.
- Felix, Sascha. 1989. Finite infinitives in Modern Greek. In *Linguistik Aktuell/Linguistics Today*, 113–132. Amsterdam: John Benjamins Publishing Company.
- Fernández-Sánchez, Javier. 2016. Topics at the left edge of infinitive clauses in Spanish and Catalan. *Borealis – An International Journal of Hispanic Linguistics* 5:111.
- Gärtner, Hans-Martin. 2009. More on the indefinite-interrogative affinity: The view from embedded non-finite interrogatives. *Linguistic Typology* 13:1–37.
- van Gelderen, Elly. 1998. The future of for to. *American Journal of Germanic Linguistics and Literatures* 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: 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.
- Holmberg, Anders. 1983. Infinita fragebisatser i finlanfssvenk dialekt. Doctoral Dissertation, University of Stockholm.
- Holmberg, Anders. 1986. Word order and syntactic features in the Scandinavian languages and English. Doctoral Dissertation, University of Stockholm.
- Hooper, Joan B., and Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465–499.
- Hornstein, Norbert, and David Lightfoot. 1987. Predication and PRO. *Language* 63:23–52.
- Huang, Nick. 2018. Control complements in Mandarin Chinese: implications for restructuring and the Chinese finiteness debate. *Journal of East Asian Linguistics* 27:347–376.
- Jung, Hakyung. 2009. Null Prepositional Complementizers and the Russian Dative of Obligation. In *Proceedings of Formal Approaches to Slavic Linguistics*, volume 17, 64–80. Ann Arbor: Michigan Slavic Publications.
- Kalinina, Elena, and Nina Sumbatova. 2007. Clause structure and verbal forms in Nakh-Dagestani languages. In *Finiteness: Theoretical and Empirical Foundations*, ed. Irina Nikolaeva, 183–249. OUP Oxford.

- Kayne, Richard S. 1981. ECP extensions. *Linguistic Inquiry* 12:93–133.
- Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht, Holland: Foris Publications.
- Kayne, Richard S. 1991. Romance clitics, Verb movement, and PRO. *Linguistic Inquiry* 22:647–686.
- Keine, S. 2020. *Probes and their horizons*. Linguistic Inquiry Monographs. MIT Press.
- Kiss, Katalin. 2002. *The Syntax of Hungarian*. Cambridge Syntax Guides. Cambridge: Cambridge University Press.
- Klingvall, Eva. 2018. Agreement and reconstruction correlate in Swedish. *Natural Language & Linguistic Theory* 36:1165–1205.
- Kornfilt, Jaklin. 1996. On some infinitival wh-constructions in Turkish. In *Dilbilim Araştırmaları*, 192–215. Ankara.
- Koster, Jan. 1984. On binding and control. *Linguistic Inquiry* 15:417–459.
- Landau, Idan. 2001. Control and extraposition: the case of super-equi. *Natural Language and Linguistic Theory* 19:109–152.
- Landau, Idan. 2004a. Chain resolution in Hebrew V(P)-fronting. In *Proceedings of the North East Linguistic Society 34*, ed. Keir Moulton and Matthew Wolf, volume 2, 357–372. University of Massachusetts at Amherst: Graduate Linguistic Student Association.
- Landau, Idan. 2004b. 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.
- Landau, Idan. 2015. *A two-tiered theory of control*. MIT Press.
- LaPolla, M. V. 1988. Clitic Movement in Spanish and the Projection Principle. In *Advances in Romance Linguistics*, ed. D Birdsong and J-P Montreuil, 217–231. Dordrecht: Foris.
- Larsson, Ida. 2017. Double complementizers. *Nordic Atlas of Language Structures Journal* 1.
- Long, M. E. 1976. French infinitival complementizers and their place in Generative Grammar. In *Current studies in romance linguistics*, ed. M. Luján and F. G. Hensey, 205–220. Washington, DC: Georgetown University Press.
- Lujan, Marta. 1980. Clitic promotion and mood in Spanish verbal complements. *Linguistics* 381–484.
- McCloskey, James, and Peter Sells. 1988. Control and A-chains in Modern Irish. *Natural Language and Linguistic Theory* 6:143–189.
- McFadden, Thomas, and Sandhya Sundaresan. 2014. Finiteness in South Asian languages: an introduction. *Natural Language & Linguistic Theory* 32:1–27.
- Melnik, Nurit. 2015. Raising, inversion and agreement in modern Hebrew. *Journal of Linguistics* 53:147–179.
- Miyagawa, Shigeru. 2017. Topicalization. *GENGO KENKYU (Journal of the Linguistic Society of Japan)* 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. Unpublished manuscript. Cambridge, Massachusetts.

- 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.
- Przepiórkowski, Adam, and Alexandr Rosen. 2004. Czech and Polish Raising/Control with or without Structure Sharing. *Research in Language* 3.
- 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*, ed. Patrick Brandt and Eric Fuss. Berlin: Akademie Verlag.
- Satik, Deniz. 2022a. Cartography: Innateness or Convergent Cultural Evolution? *Frontiers in Psychology* 13.
- Satik, Deniz. 2022b. The Strong Minimalist Thesis is too strong: syntax is more than just Merge. *Biolinguistics* 16.
- Schiffman, H.F. 1999. *A Reference Grammar of Spoken Tamil*. Reference Grammars. Cambridge: 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*, 327–341. Tokyo: H. Syobo.
- 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, University of Lund, Sweden.
- 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. *NYU Working Papers in Linguistics* .
- Thráinsson, H. 2007. *The Syntax of Icelandic*. Cambridge Syntax Guides. Cambridge: Cambridge University Press.
- Thráinsson, Höskuldur. 1993. On the structure of infinitival complements. *Harvard Working Papers in Linguistics* 3.

- Todorović, Neda, and Susi Wurmbrand. 2016. Finiteness across domains. In *Selected proceedings from FDSL*, ed. Teodora Radeva-Bork and Peter Kosta. Frankfurt am Main: Peter Lang.
- Ussery, Cherlon, Lydia Ding, and Yining Rebecca Liu. 2016. The typology of Mandarin infinitives. *Proceedings of the Linguistic Society of America* 1:28.
- Villalba, Xavier. 2009. *The syntax and semantics of dislocations in Catalan. a study on asymmetric syntax at the peripheries of sentence*. Cologne: Lambert.
- Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford: Oxford University Press.
- Wurmbrand, Susanne, Magdalena Lohninger, Neda Todorovic, Iva Kovac, and Caroline Pajancic. 2020. Finiteness in South Slavic Complement Clauses: Evidence for an implicational finiteness universal. *Linguistica* 60:119–137.
- Wurmbrand, Susi. 1994. The German leicht-construction. In *Wiener Linguistische Gazette*, volume 48-50, 87–100.
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.
- Wurmbrand, Susi. 2014. The Merge Condition. In *Language faculty and beyond*, 130–166. Amsterdam: John Benjamins Publishing Company.
- Wurmbrand, Susi. 2017. Stripping and Topless Complements 48:341–366.
- Wurmbrand, Susi. 2019. Cross-clausal A-dependencies. In *Proceedings of the Fifty-fourth Annual Meeting of the Chicago Linguistic Society*. Chicago, IL: Chicago Linguistic Society.
- 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.
- Zabrocki, T. 1981. *Lexical Rules of Semantic Interpretation. Control and NP-Movement in English and Polish*. Poznan: Poznan.