Article

More on the indefinite-interrogative affinity:
The view from embedded non-finite interrogatives

HANS-MARTIN GÄRTNER

Abstract

*German disallows embedded infinitival (*wh-*)interrogatives. Although cross-linguistic comparison has not been undertaken to any broader extent, most accounts attribute this gap to properties of complementizers and the C-system. By contrast, this gap will here be linked to a correlation between non-finite (*wh-*)interrogatives and the indefinite-interrogative affinity present in the inventory of* wh-*pronouns of many languages. The claim will be defended that, if a language has embedded non-finite (*wh-*)interrogatives, then its pronominal system does not have a robust indefinite/interrogative ambiguity. In addition to a crosslinguistic survey, it will be sketched – with particular reference to German – how to relate these two domains in terms of clausal typing and the illocutionary force of non-finite (root) interrogatives.*

Keywords: clause types, complementation, finiteness, illocutionary force, indefinite, infinitive, interrogative, pronouns, syntax

1.   Introduction

It is a well-known fact about the syntax of English and German that they differ with respect to the acceptability of embedded infinitival interrogatives (see, for example, Tappe 1984). Thus, consider the contrast in (1), where (1b) is a direct translation of (1a).[1]

(1)   a.   English (Indo-European)
              Mary suddenly remembered [*where to find the keys*]

---

1. Readers who trust corpus examples more than made up ones should consult Duffley & Enns (1996: 238), who give examples like ... *whose employer* ... *told her what to answer if anyone called* ...

> b.   German (Indo-European)
>   *Maria erinnerte sich plötzlich* [*wo die Schlüssel (zu) finden*]

Although various accounts for the unacceptability of (1b) have been put forward, this phenomenon has rarely been discussed in crosslinguistic perspective, Sabel (1996: 275–301, 2006) being a notable exception. The present article aims to shed new light on the contrast in (1) by linking it to a crosslinguistic correlation involving the inventory of interrogative and indefinite pronouns of English and German (Section 2). Section 3 is devoted to (removing some obstacles to) a broader typological survey, and Section 4 outlines how to relate presence vs. absence of embedded non-finite interrogatives to the indefinite-interrogative affinity; the focus will be on German facts and on assumptions about clausal typing and illocutionary force. Section 5 provides a summary and outlook.

2.   Two hypotheses about embedded infinitival interrogatives

To my knowledge, Sabel (1996: 275–301, 2006) has so far been the only serious attempt to view the contrast in (1) from a crosslinguistic perspective.[2] Building on a study of Germanic, Romance, and Slavic languages, Sabel (1996: 295) suggests the following generalization:

(2)   If a language L possesses *Wh*-movement to Spec,CP in infinitives, then L possesses the (independent) option of filling the infinitival C-system with a base-generated overt element.

The existence of items like *for* in the English C-system, as shown in (3), and the lack of counterpart elements in German, is interpreted as substantiation of (2).[3]

(3)   English (Indo-European)
       They would prefer [$_{CP}$ *for* [$_{IP}$ *Mary to put the keys in the safe*]]

Sabel goes on to implement (2) in terms of the "strength" of the (infinitival) head $C°$ interacting with $X°$-to-$C°$-movement and the *Wh*-Criterion (Rizzi 1996). This allows him to derive the contrast in (1). A sketch of the account can be found in Appendix A.

For the sake of exposition, languages possessing embedded non-finite interrogatives – of which those possessing embedded infinitival interrogatives are a

---

2. See Sabel 1996 for a survey of earlier formal analyses ruling out (1b).

3. I will not go into discussing whether Sabel's assumptions about particular complementizers in particular languages are correct, as this issue is orthogonal to the concerns of this article.

subclass (see Section 3) – will be called [+enfi]-languages and languages possessing overtly base-generated infinitival complementizers [+obc]-languages. Sabel's generalization can then be abbreviated as in (4).

(4)   [+enfi] $\Rightarrow$ [+obc]

(5) represents the (small) typology of languages on which Sabel's study is based.[4]

(5)   a.   [+enfi, +obc]:   *English, French, Italian, Polish, (European) Portuguese, Spanish*
       b.   [+enfi, −obc]:   Ø
       c.   [−enfi, +obc]:   ?
       d.   [−enfi, −obc]:   *Danish, German, Norwegian, Swedish*

Now, the point of the present article is to argue that (5) masks another generalization, which – to my knowledge – has so far gone unnoticed. Thus, note that many German word strings involving *wh*-indefinites and *to*-infinitivals are acceptable as soon as an interpretation is available that takes them to be non-interrogative. This is shown in (6).[5]

(6)   German (Indo-European)
       a.   *Ich habe vor, was zu tun*
             'I intend to do something.'
       b.   *Ich erinnere mich, wohin zu fahren (, wo ich lange nicht gewesen war)*
             'I remember going somewhere (specific) (where I hadn't been for a long time).'

Quite strikingly, none of Sabel's [+enfi, +obc]-languages allows such (re-)interpretations, which is because their sets of interrogative and indefinite pronouns are strictly disjoint. (7) shows this for (counterparts of) *who*.

---

4. Sabel's generalization is not a biconditional. Nevertheless it would be useful to know more about languages that make up class (5c), i.e., [−enfi, +obc]-languages. These languages are predicted to be able to develop into [+enfi, +obc]-languages. In fact, both Icelandic and Swedish (*pace* Sabel) are often analyzed as [−enfi, +obc] (Beukema & den Dikken 1989, Jónsson 1996, Platzack 1986; but see also Thráinsson 1998).

5. The same effect can be replicated for *to*-less ("bare") infinitives: *Wir wollten wen besuchen* ('We wanted to visit someone'). Some speakers of German seem to require the additional relative clause in (6b), in order to accept the indefinite reading of *wohin* (lit.,'where to'). See Fodor & Sag 1982 for a discussion of the influence of rich descriptive content on the availability of specific (or "referential") readings. For further remarks on specificity see below.

(7) a. English: *who*   vs. *someone*
     b. French: *qui*   vs. *quelqu'un*
     c. Italian: *chi*   vs. *qualcuno*
     d. Polish: *kto*   vs. *ktoś*
     e. Portuguese: *quem* vs. *alguém*
     f. Spanish: *quién* vs. *alguien*

If this is not just sheer coincidence – and I will argue it is not – it warrants the following (preliminary) hypothesis:

(8) If a language *L* possesses embedded non-finite (*wh-*)interrogatives, then the pronominal system of *L* does not possess any indefinite/interrogative ambiguity.

Indeed, Basque and Modern Hebrew are two more languages – not investigated by Sabel – confirming (8) the same way as the languages in (5a) do. The pronominal systems of Basque and Modern Hebrew strictly distinguish interrogatives like *nor* and *mi* ('who'), respectively, from pure indefinites like *nor-bait* and *mishehu* ('someone'), respectively (Amir Coffin & Bolozky 2005: 173–175, Haspelmath 1997: 324). (9) shows a Basque and a Modern Hebrew embedded non-finite interrogative (henceforth ENFI):

(9) a. Basque (isolate)
         [*Nora joan*] *erabaki dute*
          where go.PTCP decided they.have
         'They decided where to go.' (Ricardo Etxepare, personal communication)
     b. Modern Hebrew (Semitic)
         Mary shaxexa [*lean la-lexet la-reayon shela*]
         Mary forgot   where to-go   for-interview hers
         'Mary forgot where to go for her interview.' (Hadas Kotek, personal communication)

Let us call languages showing the indefinite/interrogative ambiguity [i=i]-languages, the ones lacking it [i≠i]-languages. (8) can then be abbreviated as in (10).

(10) [+enfi] ⇒ [i≠i]

On the basis of generalization (8), (5) is recast into a small typology in which German assumes an isolated position.

(11) a. [+enfi, i=i]:   Ø
      b. [+enfi, i≠i]:   *Basque, English, French, Modern Hebrew, Italian, Polish, (European) Portuguese, Spanish*

c. [−enfi, i=i]: *German*
d. [−enfi, i≠i]: *Danish, Norwegian, Swedish*

However, as we are going to see later, German appears to be indicative of a larger typological trend here. Before developing that broader picture, though, generalization (8) has to be further sharpened. Thus, Dutch, one of German's closest neighbors, is a language that follows Sabel's generalization in being [+enfi, +obc] (Sabel 1996: 294–295) yet does not obey (8) (cf. Postma 1994). This is shown in (12).

(12)   Dutch (Indo-European)
a. *Ik weet niet* [*wie te bezoeken*]
I know not who to visit
'I don't know who to visit.' (Sabel 1996: 295)
b. *dat zij probeerde* [CP *om* [IP *het boek te lezen*]]
that she tried   COMP   the book to read
'that she tried to read the book' (Sabel 1996: 294)
c. *Jan heeft wat gedaan*
Jan has what done
'Jan has done something.' (Postma 1994: 187)

(12a) is evidence for Dutch being [+enfi];[6] *om* in (12b) has been argued to be an infinitival complementizer (cf., among others, Beukema & den Dikken 1989, Tappe 1984), which warrants the categorization of Dutch as [+obc]; and the indefinite reading of *wat* in (12c) indicates that Dutch is also a [i=i]-language. However, a closer look reveals that the [i=i]-status of Dutch is of a limited kind. In presenting Dutch indefinites, Haspelmath (1997: 246) notes that "[a] colloquial variant of *iets* is *wat*, i.e., the bare interrogative (but the other interrogatives cannot be used as indefinites)" (cf. Postma 1994: 188). (13) provides one example for this difference between German and Dutch.

(13)   German: *Ich habe wo gelesen, dass der Sommer schön wird*
Dutch: *Ik heb *waar / ergens gelezen, dat de zomer mooi wordt*
'I read somewhere that summer is going to be nice.' (Michael Cysouw, personal communication)

Clearly, Dutch lacks a general, or – as I will call it – ROBUST, [i=i]-ambiguity.[7] Let us therefore revise (8) accordingly.

---

6. Tonjes Veenstra (personal communication) informs me that, at least for some speakers, ENFIs in Dutch are only fine with interrogative pronoun *wat* ('what').

7. While the [i=i]-property of German is strong, it is not exceptionless. Thus, the counterparts of *how* and *why* (*wie* and *wieso/warum*) do not allow the pure indefinite reading. Exceptional behavior of these two pronouns is familiar from the study of *wh*-dependencies and has some-times been linked to their particular denotational domain (cf. Szabolcsi & Zwarts 1997). See

(14)   If a language *L* possesses embedded non-finite (*wh-*)interrogatives, then the pronominal system of *L* does not possess any robust indefinite/interrogative ambiguity.

   Apart from the paradigmatic non-robustness of Dutch, there are two further types, namely, distributional and semantic non-robustness. The former is exemplified here by Latin, the latter by Lithuanian, Russian, and Slovene.

   For Latin, we can again rely on Haspelmath, who notes that "[...] bare interrogatives can be used as indefinites only when they are enclitic upon an element (such as *si* 'if', *num* 'question particle') early in the sentence [...]" (Haspelmath 1997: 254). This restriction is shown in (15).

(15)   Latin (Indo-European)
   a.   *Si   quid   petieritis   me   in   nomine   meo,   hoc*
         if   what   ask.SBJV.FUT.2SG   me   in   name   my   that
         faciam
         do.SBJV.1SG
         'If ye shall ask anything in my name, I will do it.' (Haspelmath 1997: 255)
   b.   *Licet   mihi   loqui   *quid   /   ali-quid   ad   te?*
         allowed   me.DAT   say.INF   what   /   some-what   to   you
         'May I say something to you?' (Haspelmath 1997: 254)

The distributional non-robustness of Latin is of particular interest here, given that Latin may potentially be analyzed as a [+enfi]-language. Thus, what looks like an ENFI occurs in the following piece of reported speech.[8, 9]

Zaefferer 1991 for a survey of *wh-*pronouns in German. I disagree with Zaefferer's classification of *wann* ('when') as lacking a pure indefinite reading (Zaefferer 1991: 84), since I accept sentences like *Könntest Du mich mal wann anrufen* ('Could you give me a call sometime').

8. Thanks to Peter Staudacher (personal communication) for bringing this to my attention. Without a look at the wider context of (16), I am not fully convinced that we are dealing with subordination here. Bolkestein's translation actually contains a question mark after 'discreditable'. Likewise it is pointed out there that (16) reports on a rhetorical question, something that is not well captured by use of 'informed' as reportive predicate. (16) could alternatively be taken as a case of free indirect discourse and thus some kind of "embedded root phenomenon" (cf. Banfield 1973, Heycock 2006, Hooper & Thompson 1973).
   Similar remarks apply to Latin interrogative narrative infinitives, which according to Nikolaeva (2007b: 158) marginally occur in dependent clauses. Their use in main clauses is pragmatically restricted in that they are "not really acceptable in dialogues and question-answer pairs". The example discussed in the text, *Qui mori timore nisi ego?* ('Who died from terror more than me?'), appears to function as a rhetorical question.

9. It has been noted in the literature (e.g., Gärtner & Steinbach 1994: 37, Zaefferer 1991: 78) that German *wh-*indefinites obey a weaker kind of distributional restriction, insofar as they cannot occur alone in pre-V2 position ("Vorfeld", Spec,CP). This is shown in (ia).

(16) Latin (Indo-European)
docebant … postremo [*quid esse … turpius*]
informed   finally   what be   more.discreditable
'They finally informed him of what was more discreditable.'
(Bolkestein 1996: 127)

Semantic non-robustness is a feature of Lithuanian (Axel Holvoet, personal communication), Russian (Andrej Malchukov, personal communication), and Slovene (Haspelmath 1997, Herrity 2000). These languages restrict their bare interrogatives to non-specific readings.

(17) Slovene (Indo-European)
a. *Ali me je kdó iskál?*
   Q me AUX who look.for
   'Has anyone been looking for me?' (Herrity 2000: 109)
b. *Včeraj me je nekdo/*kdó poklical*
   yesterday me AUX someone/who called
   'Yesterday someone called me.' (Boštjan Dvořák, personal communication)

This fits in with observations by Haspelmath (1997: 173), who notes that in "western Indo-European" [i=i]-languages, "bare interrogatives may occur practically in all non-emphatic non-specific functions. They are generally excluded from past or current present affirmative declarative clauses, where indefinites must be specific." (17b), which clearly meets Haspelmath's criterion for the specific use of indefinites, could be represented by an existentially quantified statement an assertion of which commits the speaker to the existence of an individual satisfying the description in question. This is provided in (18) (sp = speaker).

(18) $\exists x$ [PERSON(x) $\land$ CALLED(x, sp)]

According to such a specificity criterion, German *wh*-pronouns can clearly be specific, since (17b) could be translated into German by (19), involving *wer* ('someone').

(i)   a.   *Wer hat angerufen*
       b.   *Jemand hat angerufen*

(ia) can only mean 'who called?' but not 'someone called', the latter having to be expressed as in (ib). For some speakers the effect seems to be weaker, especially if the indefinite bears an L*H ("topic") accent (Reis 1991: 45). The restriction in (i) is sometimes mistakenly interpreted as a ban on *wh*-indefinites in initial, or procliticized, position (Bhat 2004: 241, Haspelmath 1997: 170). As has been shown by Reis (1991: 44), such a characterization is incorrect, as is illustrated by *Wen Nettes haben wir hier noch nie gesehen*. This can be translated as 'We have never seen anyone (who is) nice here'.

(19)   German (Indo-European)
        Gestern hat mich wer angerufen

There are stronger views on the specificity of indefinites, according to which speaker beliefs about the individual in question have to be more specific, for example, allowing identification (in principle) of the individual if confronted with it. This is not necessarily implied by an utterance of (17b)/(19), which could be continued by 'but due to a noisy connection I couldn't find out who it was'. Fodor & Sag (1982) appeal to the criterion of "rich descriptive content" for eliciting this stronger interpretation of indefinites. German *wh*-pronouns also pass this stronger specificity requirement, as (20) shows (cf. (6b)).

(20)   German (Indo-European)
        Gestern   hat mich wer   angerufen, den   kenne ich noch
        yesterday has me   who   called   DEM   know I   still
        aus   der Schule
        from the school
        'Yesterday, someone called me who I know from school.'

There are also weaker views on specificity, according to which the indefinite's taking wide scope with respect to an operator is a sufficient criterion (cf. Ruys 1993).[10]

    The point of going into some detail here is to get some feeling for specificity as a third robustness factor for [i=i]-languages. To the extent that Haspelmath is correct, not only Lithuanian, Russian, and Slovene, but also Belorussian, Gothic, Latin, colloquial Polish, Old Church Slavonic, and Ukrainian would be semantically non-robust [i=i]-languages. At least these are the languages for which evidence in that direction is available.[11] In fact I will tentatively assume that this is correct, except for the case of Gothic. Thus, consider (21).

(21)   Gothic (Indo-European)
        Skal   þus   hwa   qiþan
        I.must to.you what say
        'I must tell you something.' (Haspelmath 1997: 173)

It seems that for Haspelmath the presence of modal *skal* is sufficient to consider *hwa* a non-specific indefinite. By the semantic criteria outlined above it is much more likely though that we are dealing with a specific reading. Thus, (22b),

---

10. This leaves open the option for the indefinite to be in the scope of other operators and assume so-called "intermediate readings".

11. For colloquial Polish this is independently confirmed by Cheng (1991: 104) and Joanna Błaszczak (personal communication).

paraphrasable as 'there is something I must tell you', is the much more likely rendering of (21) than (22a) (ad = addressee, □ = necessity).

(22)   a.   □ [∃x [THING(x) ∧ TELL(sp, ad, x)]]
         b.   ∃x [THING(x) ∧ □ [TELL(sp, ad, x)]]

(22b) would imply specificity of the indefinite on the weak requirement of taking scope with respect to to an operator, the modal operator in the case at hand. At the same time (22b) commits the speaker to the existence of something – a piece of information – that he or she is to communicate to the addressee. Under these circumstances, it is also very likely that some more intimate acquaintance of the speaker with that piece of information is implied.[12] Thus, all three criteria of specificity would be fulfilled, *pace* Haspelmath.[13] I am therefore inclined to count Gothic among the robust [i=i]-languages, at least as far as semantics goes.

As a matter of notational precision, [i=i]-languages will henceforth be split up into languages with a robust [i=i]-ambiguity (abbreviated as [+i=i]) and languages with a non-robust [i=i]-ambiguity (abbreviated as [%i=i]). Where the distinction does not matter, I will speak of [i=i]-languages. For the sake of notational parallelism, [−i=i] will henceforth be used instead of [i≠i]. Generalization (14) must therefore be rendered as in (23).

(23)   [+enfi] ⇒ ¬[+i=i]

Among the [%i=i]-languages lacking robustness according to the specificity criterion, Lithuanian, Russian, and Slovene are important cases at this point because they are known to be [+enfi]-languages. This is shown in (24).[14]

(24)   a.   Lithuanian (Indo-European)
           Jis nežino [*kada skaityti knyga*]
           he not.know when to.read book.ACC
           'He doesn't know when to read the book.' (Franks & Lavine 2006: 240)
        b.   Russian (Indo-European)
           Ja ne znaju [*kuda idti*]
           I not know where go.INF
           'I don't know where to go.'

---

12. A look at the biblical source (Luke 7, 40–42) confirms this strongly specific interpretation of *hwa*.

13. See also Dahl 1999 for further related discussion.

14. The Russian example can be found at https://lingweb.eva.mpg.de/linguipedia/index.php/Wh-infinitive.

c. Slovene (Indo-European)
Odločil sem se [kje zgraditi hišo]
decided AUX REFL where build.INF house
'I decided where to build a house.' (Marušic 2005: 118)

The putative [+enfi]-language Latin lacks semantic as well as distributional robustness, as we have seen above. According to Postma (1994), the [+enfi]-language Dutch even lacks robustness in all three dimensions, paradigmatic (as we have seen above), distributional (*wat* used as a pure indefinite being confined to its VP-internal base-position), and semantic.

I thus submit that (14) is the correct generalization, banning the coexistence in a single language of a robust [i=i]-ambiguity and ENFIs.

Let me turn to another rather interesting piece of evidence. It has been discovered that certain dialects of German do possess ENFIs. These are Reichenau German (a variety of Alemannic, Brandner 2004) and Pennsylvania German. Examples are given in (25).[15]

(25) a. Reichenau German (Indo-European)
I zoag dir schnell [welle Socke
I.NOM show you.DAT quickly which socks.ACC
schtopfe]
mend.INF
'I show you quickly which socks to mend.' (Brandner 2004: 9)
b. Pennsylvania German (Indo-European)
Ich hab vergesse [wo fer annegehe]
I.NOM have forgotten where for there.go.INF
'I forgot where to go.' (Mark Louden, personal communication)

Brandner (2004) argues that Reichenau German is actually a [+obc]-language, the inflected preposition *zum* ('to.the.DAT') having turned into an infinitival complementizer.[16] To the extent that *fer* in Pennsylvania German can be analyzed as counterpart of English *for*, similar conclusions about the [+obc]-status

---

15. Giusti (1986: 134) reports on a South Tyrolian dialect of German that may possess ENFIs too.
16. Crucial data for this assumption are given in (i).

(i) a. *Die sell Wies war amel schwer [zum vu Hand maie]*
the this meadow was sometimes difficult to.the by hand mow
'This meadow was sometimes difficult to mow by hand.' (Brandner 2004: 4)
b. **Die sell Wies war amel schwer [vu Hand zum maie]*

According to Brandner (2004), the placement of *zum* in Reichenau German differs from its counterpart in Bavarian where it must be adjacent to the infinitival verb. The overall distribution of infinitival markers in Reichenau German is rather intricate. Thus, *zum* is missing from ENFIs. As far as I can see, this has no bearing on Sabel's and my generalization.

of Pennsylvania German might be drawn. Both variants of German would then confirm Sabel's generalization in (2).

Interestingly, Reichenau German and Pennsylvania German equally confirm (8)/(14), i.e., they are not [+i=i]. For assuming Pennsylvania German to be [−i=i], I rely on Mark Louden (personal communication). For Reichenau German, crucial evidence can be gathered from Bayer & Brandner 2004 where, systematically, Bavarian indefinite *was* ('what') corresponds to *ebbes* ('something', cf. German *etwas*) in Reichenau German.[17] This points toward paradigmatic and/or semantic non-robustness, if not absence, of any [i=i]-ambiguity. Reichenau German should therefore be considered at least [%i=i]. I thus take (14) to be a generalization worth exploring further. (26) summarizes our findings so far.

(26) a.   [+enfi, +i=i]:   Ø
      a′. [+enfi, %i=i]:   *Dutch, (Reichenau German), (Latin), Lithuanian, Russian, Slovene*
      b.   [+enfi, −i=i]:   *Basque, English, French, Pennsylvania German, Modern Hebrew, Italian, Polish, (European) Portuguese, Spanish*
      c.   [−enfi, +i=i]:   *German*
      c′. [−enfi, %i=i]:   *(Latin)*
      d.   [−enfi, −i=i]:   *Danish, Norwegian, Swedish*

3. A crosslinguistic survey

The following section is devoted to exploring generalization (14) crosslinguistically. Pure logic demands that one find the [+enfi]-languages and check their [i=i]-status or that one find the [+i=i]-languages and check them for ENFIs. As it turns out – given the current lack of published sources that would provide sufficient detail – none of this is easy, so this section is of a rather tentative nature with very preliminary results and many open questions. In fact, I am forced to take a mixed approach. First, I will list [i=i]-languages. Then – avoiding the difficult further check for their [+i=i] vs. [%i=i] nature – I will look at non-finite forms and make an assessment as to which of the [i=i]-languages could in principle display ENFIs built from which kinds of non-finites. In order to stay close enough to what I believe is involved in contrast (1), I will privilege a notion of non-finiteness that requires morphosyntactic asymmetry/reduction (Bisang 2001) as a formal reflex. Feasibility, however, demands that I trust published sources and naively take what these label as "infinitives", "converbs",

---

17. Jürg Fleischer (personal communication) informs me that this is generally true of the Swiss varieties of Alemannic.

and "action nominals" as my prime candidates for non-finite sources of EN-FIs.[18]

Also, it is assumed that crosslinguistic identification of embedded interrogatives (sometimes also referred to as "indirect questions") must look to occurrences of interrogatives as arguments of predicates like 'wonder', 'know', and 'tell', etc.[19]

Let us thus begin with the indefinite/interrogative ambiguity. A survey based on Cheng 1991, Haspelmath 1997 and 2005, Bhat 2000 and 2004, and Bruening 2007 yielded the following 62 [i=i]-languages (out of a sample of roughly 150).[20]

(27)   [i=i]-languages

Aghem, Assuriní, Atayal, Belorussian, Burushaski, Chinese, Diyari, Djaru, Dutch, Dyirbal, Old English, Finnish Sign Language, Galibi, German, Goajiro, Gooniyandi, Gothic, Classical Greek, Guaraní, Hmong Njua, Hopi, Jaminjung, Kaingang, Kamaiurá, Khmer, Klamath, Koasati, Korean, Lakhota, Latin, Lithuanian, Mangarayi, Mapuche, Maricopa, Martuthunira, Masalit, Mising, Mundari, Newari, Panare, Panyjima, Pashto, Passamaquoddy, Eastern Pomo, Ancash Quechua, Russian, Vedic Sanskrit, Santali, Shoshone, Siuslaw, Old Church Slavonic, Slovene, Takelma, Thai, Ukrainian, Warndarang, Wintu, Xinh Mul, Yaqui, Yidini, Yindjibarndi, Yup'ik

From our discussion in Section 2 we already know that Dutch, Latin, Lithuanian, Russian, and Slovene are [%i=i], and we conjectured the same for Belorussian, Old Church Slavonic, and Ukrainian. Newari can be classified as

---

18. "Participles" are left out of the picture. Except for the Basque example in (9a), I am not currently aware of any relevant non-finite structures in the languages investigated that are based on them. Eventually, participles will have to be investigated more thoroughly too, of course.

19. Karttunen (1977: 6) gives a list of such predicates. Embedded interrogatives may derivatively – depending on the analysis – give rise to free relatives. Curiously, English ENFIs do not allow this, as P. Jacobson (1995: 478) shows: *I haven't yet read* [*what(ever) to read*]. In the domain of adjuncts, embedded interrogatives arguably give rise to (concessive) conditionals (cf. König & van der Auwera 1988). I believe, though, that it is safe to assume that these are derivative, not exclusive, usages and that therefore adjuncts can be excluded from the search for ENFIs.

20. The study by Ultan (1978) does not contain any specific information about the [i=i]-property of languages. I owe the information that Aghem belongs here to Larry Hyman and Maria Polinsky (personal communication). For Old English, see Quirk & Wrenn 1989: 39 and Fischer et al. 2000. For Finnish Sign Language, see Zeshan 2004: 26. Bhat (2000: 383, 2004: 241) incorrectly classifies Bagandji as a [i=i]-language. The error is due to misleading terminology used by Hercus (1982: 171), who dubs usages of interrogative pronouns in indirect as opposed to direct questions (e.g., *I did not see who chopped down that tree*) "indefinite".

a [%i=i]-language, given that its non-interrogatively used bare *wh*-pronouns are reported to be confined to the scope of negation, meaning 'nobody' and 'nothing' only (Haspelmath 1997: 170). Another [%i=i]-language is Eastern Pomo, where the only [i=i]-ambiguity arises for *čʰí* in its meaning 'how' or 'somehow' (Haspelmath 1997: 328, McLendon 1975: 130). (28) summarizes this.

(28) [%i=i]-languages (so far)
Belorussian, Chinese,[21] Dutch, Latin, Lithuanian, Newari, Eastern Pomo, Russian, Old Church Slavonic, Slovene, Ukrainian

Now, since the available published sources make it rather difficult to establish paradigmatic, distributional, and semantic robustness for the remaining 51 languages in (27),[22] it would seem worthwhile trying to show instead that they do not possess ENFIs. However, many grammars are silent about the kind of reduced interrogatives that would count as evidence in this matter. Furthermore, there seem to be no monographs about them.[23] What can be investigated, though, is the respective inventories of non-finite forms that could plausibly underlie ENFIs. This is the direction pursued in the following.

As already indicated above, in order to stay close enough to contrast (1) without ruling out languages prematurely, I follow Bisang (2001: 1404–1408), who makes the suggestion that a necessary condition for possession of (counterparts of) infinitives requires a language *L* to be m-ASYMMETRIC as defined below.[24]

---

21. My information about specificity in (Mandarin) Chinese is conflicting. While both Cheng (1991) and Haspelmath (1997) report that interrogative pronouns can only be used as non-specific indefinites, Haspelmath (1997: 171) and Bhat (2000: 379) give examples that can be interpreted as specific indefinite uses of *shenme* ('what'). Chinese will temporarily be counted among the [%i=i]-languages, though.

22. I am inclined to count Passamaquoddy as a [+i=i]-language, given the possibility of *wh*-indefinites to occur in sentences translated as 'While they were dragging the deer they heard something' (Bruening 2007: 150). On the other hand, unlike German and Gothic (see Section 2), Passamaquoddy *wh*-indefinites are reported to be confined to narrow scope with respect to operators (Bruening 2007: 160). For the [+i=i]-status of Yup'ik, evidence is provided by S. Jacobson (1995: 185).

23. The study by Bhatt (2006) is almost exclusively concerned with English. It is briefly mentioned, though, that Hindi/Urdu does not possess ENFIs (Bhatt 2006: 108). Rajesh Bhatt (personal communication) informs me that Hindi/Urdu is a [−i=i]-language.

24. This definition allows languages to be p-asymmetric and m-asymmetric at the same time. Joseph (1983: 6–36) devises a more restrictive general notion of infinitive as a subclass of "nonfinite verbs" fulfilling additional functional constraints. For further pertinent discussion see Ylikoski 2003 and Nikolaeva 2007a.

(29)   If a language *L* has clause types with obligatory marking of relevant features in the extended projection of the verb, then
    a.   *L* is *m(inus)-asymmetric* if that marking disappears in dependent clauses, and
    b.   *L* is *p(lus)-asymmetric* if that marking disappears in independent clauses.

(29a), of course, conveniently relates to a notion of non-finites "understood in its traditional sense; i.e., in contrast to finite forms, non-finites are not usually marked for such categories as tense, mood, aspect, person or number, and they do not function as only predicates of independent sentences" (Ylikoski 2003: 186).

As an immediate consequence of requiring ENFIs to be m-asymmetric, it is possible to discard languages for which there is evidence that "they are neutral to the finite/non finite distinction" (Bisang 2001: 1408). These are languages of the "Far East type" (i.e., Chinese, Thai, Vietnamese, Cambodian, Hmong).[25]

Now, given that – as we have seen in Section 2 – quite a number of European languages have ENFIs based on infinitives, it makes sense to look for infinitives among the remaining languages of (27) first. Relying in part on Cristofaro 2003, I assume that at least the languages in (30) possess infinitives.[26]

(30)   Burushaski, Old English, Gothic, Pashto, Ancash Quechua, Vedic Sanskrit, Shoshone

The comprehensive study of Sanskrit by Whitney (1950) and the absence of embedded interrogatives in the study of Shoshone by Dayley (1989: 344–386) allow a reasonably strong conjecture that both are [−enfi]-languages.[27] For

---

25. For Thai, see also Hudak 1990: 771. It remains to be investigated whether a sharpened notion of "non-finiteness", capable of integrating these languages, can be designed in terms of obligatory subject-drop and/or obligatory control. For discussion, see Nikolaeva 2007a: 6, Huang 1984, Landau 2004, and Barrie 2007.

26. See Cole 1984 (Ancash Quechua), Berger 1998: 103 (Burushaski), Quirk & Wrenn 1989: 41 (Old English), Lehmann 1994: 32 and Streitberg 1981: 53–57 (Gothic), Lorenz 1979: 36 and MacKenzie 1990: 559 (Pashto), and Dayley 1989: 375 (Shoshone).
    Among the languages already eliminated and not discussed in the main text there is evidence for infinitives in Classical Greek (Joseph 1983: Chapter 3), Russian (Comrie 1990: 340, Perlmutter 2007), Old Church Slavonic (Lunt 1959: 96, 139; Ylikoski 2003: 209), and Ukrainian (Amir-Babenko & Pfiegl 2005: 49).

27. One has to be careful not to misinterpret certain structures on the basis of their translations. Thus, Whitney (1950: 352) renders an example as 'thou knowest how to loosen all bonds', where the Sanskrit original clearly does not contain any counterpart of *how* but rather a variant of French *savoir faire* structures. A comparable case can be found for Passamaquoddy (Bruening 2007: 150). It is equally important not to mix up ENFIs with the "modal existential construction" (MEC). Ambrazas (1997: 728) shows a structure glossed as *will-be-where-keep*.INF-*cow* and translated as 'We'll have somewhere to keep the cow'. This, however, seems

Shoshone this is crucial, since additional evidence suggests that it is a [+i=i]-language (Dayley 1989: 488). The [−enfi]-status of Old English, which I will come back to below (see Section 4), is attested by Fischer et al. (2000: 94). The discussion of indirect questions by Streitberg (1981: 88–89) makes no mention of ENFIs in Gothic. Instead, a good candidate for an ENFI, translatable as 'I know what to do', is shown to appear with *taujau*, a finite version of *do* bearing optative mood: *andþahta mik* [*hva taujau*]. Given our previous discussion, we can conclude that Gothic is a [−enfi, +i=i]-language like German.[28]

Let us turn next to two well-known sources of m-asymmetry, namely, "converbs" and "action nominals". It has been noted that "[c]onverbs are formally and functionally closely related to participles (verbal adjectives), infinitives, and verbal nouns or gerunds (in the sense of the English gerund)" (Tikkanen 2001: 1114). Among the [i=i]-languages, there appear to be at least five languages possessing (special) converbs, namely, those in (31) (cf. Koptjevskaja-Tamm 1994: 1247).

(31)   Burushaski, Korean, Mundari, Ancash Quechua, Vedic Sanskrit

However, canonically, converbs are used for modification and coordination, not complementation. Haspelmath (1995a: 3) characterizes a converb as "a non-

---

to be an instance of the MEC (Axel Holvoet, personal communication), which should be analyzed as some kind of free relative restricted to being an argument of predicates asserting existence or "coming into being, view, or availability" (Grosu 2004: 406). Such a view also helps reducing the number of [i=i]-languages. Hungarian, for example, allows MECs (Grosu 2004) that involve *bona fide* interrogative pronouns. The same pronouns cannot be used as pure indefinites. At the same time, Hungarian is [−enfi], as (i) shows (cf. Grosu 2004: 421). The exposition by É. Kiss (2002: 202) is misleading in this respect.

(i) a. *János   elfelejtette   keresni   a   kenyeret*
         János   forgot   look.for.INF   the bread.ACC
         'János forgot to look for the bread' (Beáta Gyuris, personal communication)
     b. **János   elfelejtette,   (hogy)   hol   keresni   a   kenyeret*
         Já   forgot   that   where   look.for.INF   the bread.ACC
         'János forgot where to look for the bread' (Beáta Gyuris, personal communication)

28. I have been unable to establish the [enfi]-status of Burushaski and Pashto. Ancash Quechua appears to be a [+i=i]-language (Haspelmath 1997: 310). Currently, I have no evidence that it has ENFIs. The related language Huallaga Quechua does possess ENFIs, as (i) shows.

(i) *Mana   musyaran-chu   [may-man   aywa-y-ta-pis*]
     not   he.knew-NEG   where-GOAL   go-INF-ACC-INDF
     'He didn't know where to go' (Weber 1983: 92)

At the same time, *wh*-pronouns used as "pure" indefinites seem to require the indefinite suffix *-pis* (Weber 1996: 78–79). Huallaga Quechua would thus not count among the [i=i]-languages.

finite verb whose main function is to mark adverbial subordination". (32) provides an example from Burushaski, where the converb *nu-qás* functions as a secondary predicate.[29]

(32)   Burushaski (isolate)
    In(-e)   nu-qás   ásimi
    he(-ERG)   CVB-laugh   he.told.me
    'He said to me laughing . . . ' (Tikkanen 2001: 1117)

Since this functional limitation is a general feature of converb constructions, we could discard languages for which this is the only source of m-asymmetry from the list of potential counterexamples to (14).

    A look at action nominals raises additional questions. From an (Indo-)European perspective it may seem unlikely that they could function as interrogative complements. Note, for example, that Scottish Gaelic, which lacks standard infinitives, does not allow its so-called "verbal noun clauses" as ENFIs. This is shown in (33).[30]

(33)   Scottish Gaelic (Indo-European)
    *Cha robh fhios   agam [*dè   a   dhèanamh*]
    NEG was   knowledge   to.me   what   PRT   do.NMLZ
    'I did not know what to do.' (Adger 2007: 49)

Irish, on the other hand, does allow ENFIs based on "verbal nouns", as (34) shows.

---

29. It is unclear how to analyze the initial pronoun *ine* ('he'). Presumably it is a topic rather than part of the constituent headed by the converb. Lezgian allows converbs as complements to a small set of predicates, among which perception verbs seem to be the only class relevant to our purposes. Consider (i).

    (i)   *Cükwera-z   čpi-n   pataw sa   žehil qwe-z   aku-na*
        Cükwer-DAT   selves-GEN   to   one   youth come-CVB.IPFV see-AOR
        'Cükwer saw a young man coming toward them' (Haspelmath 1995b: 425)

    Here *qwe-z* is an imperfective converb. A similar case from Tamil is provided by Bisang (1995: 157). However, Haspelmath (1995b: 425) notes that "factive complements as in *see that* [. . .], are expressed differently". I assume that questions, being of type (set_of_)proposition(s), induce a similar "epistemic" reading of perception verbs. The prediction therefore is that converbs cannot head embedded interrogatives as complements of perception verbs in Lezgian. This has been confirmed by Martin Haspelmath (personal communication). For further discussion of the semantics of perception verb complements, see Barwise & Perry 1983.

30. Scottish Gaelic is a [−i=i]-language (Lamb 2001: 37).

(34)   Irish (Indo-European)
      Fuaireadar amach [*conas potaí a   dhéanamh*] *ó*
      find.PST.3   out   how   pots PRT make.NMLZ from
      bheirt Shasanach
      two   Englishmen
      'They found out how to make pots from two Englishmen.' (Jim Mc-
      Closkey, personal communication)

At the same time, Irish is a [−i=i]-language, as documented by Haspelmath
(1997: 278–280).

   A thorough look at Koptjevskaja-Tamm 1993 confirms the scarcity of in-
terrogatives built from action nominals. The main exception is the following
example from West Greenlandic.[31]

(35)   West Greenlandic (Eskimo-Aleut)
      Umiarsu-up qassi-nut   tikin-ni-ssa-a
      ship-REL   how.many-ALL arrive-AN-FUT-3SG.POSS
      nalunngil-ara
      know-1SG.3SG.IND
      'I know when [*sic*!] the ship will arrive.' (Koptjevskaja-Tamm 1993:
      115)

Of course, the richly inflected AN-form *tikin-ni-ssa-a* does not strike one as a
particularly convincing exponent of m-asymmetry. However, among the [i=i]-
languages in (27) there are cases worth closer attention. The following is a
plausible example of an action nominal-based ENFI from Mapuche.[32]

(36)   Mapuche (Araucanian)
      Chew ñi   amu-n   kim-nge-la-y
      where 3SG.POSS go-NMLZ know-PASS-NEG-IND
      'It is not known where she went.' (Zúñiga 2000: 65)

More literally, the ENFI could be translated as 'her going where'. At the same
time, Mapuche is a [%i=i]-language, if the [i=i]-phenomenon reduces – as
it appears – to the fact that "[*c*]*hew* 'what' is used together with a negated
predicate to convey the meaning of 'nothing' [. . .]" (Zúñiga 2000: 20).

   Maricopa is another language whose action nominals should be considered.
First of all, Maricopa shows signs of a [+i=i]-language as (39) illustrates.

---

31. Another example is an AN-based polar question from Mongolian (Koptjevskaja-Tamm 1993:
    286).
32. For further examples, see Smeets 2008: 106.

(37)   Maricopa (Yuman)
    a. *Mki-sh   hav-ii*
        who-SBJ   enter-Q
        'Who came in?', 'Did someone come in?' (Gordon 1986: 62)
    b. *Mki-sh   hav-sh*
        who-SBJ   enter-PRF
        'Someone came in.' (Gordon 1986: 62)

At the same time, an action nominal-based embedded polar interrogative is reported on by Gordon (1986) (VAUG = augment vowel on nouns).

(38)   Maricopa (Yuman)
    Pam-sh   [*Bonnie   uuchuy-h-ny-a*]   *ny-kshkwe-k*
    Pam-SBJ   Bonnie   marry.NMLZ-IRR-DEM-VAUG   3/1-ask-REAL
    'Pam asked me if Bonnie was getting married.' (Gordon 1986: 228)

Yet, given the absence of any Q-marker in (38), an analysis of the complement as a "concealed question" suggests itself. The proper translation of (38) would therefore be 'Pam asked me (about) Bonnie's marriage-to-be'.[33] The upshot of this is that Maricopa can tentatively be added to the [−enfi, +i=i]-languages.

Clausal nominalization in Lakhota, achieved by the addition of an article, does not involve any m-asymmetry, if Comrie & Thompson (1985: 393) are right in stating that "there is nothing noun-like about the verb in this nominalized clause; it undergoes no change whatsoever from its form in a finite sentence [...]".

Let us turn to less clearly classifiable cases. Similar to the languages of the "Far East type" mentioned above, there appear to be languages that are not m-asymmetric at all, or at least lack m-asymmetric interrogatives. Among the languages in (27), one interesting case in point is Panyjima. By Haspelmath's criterion of occurrence in "past or current present affirmative declarative clauses" (see Section 2 above), Panyjima must be considered a [+i=i]-language as (39) shows.[34]

(39)   Panyjima (Australian)
    Ngatha   ngananha-lu   nhantha-nnguli-nha
    1SG.NOM   something-A   bite-PASS-PST
    'I was bitten by something.' (Dench 1991: 164)

33. A genitival analysis of *Bonnie* is clearly correct given that in Maricopa "[t]he noun which expresses the possessor has no case suffix and immediately precedes the possessed noun" (Gordon 1986: 31). For a case from Kamaiurá amenable to a similar treatment, see Brandon & Seki 1984: 98.

34. The [i=i]-ambiguity affects a number of different pronouns in Panyjima and no distributional restrictions are reported (Dench 1991: 164–166).

Crucially, Panyjima is almost completely non-m-asymmetric – and thus a member of the set of [−enfi, +i=i]-languages – as the following statement by Dench (1991: 197) indicates:[35]

> Subordinate clauses are not reduced. They take the same range of arguments, and with the exception of the non-finite relative clauses, present the same range of tense/aspect/mood information as similarly inflected main clauses.

The same applies to (Sonora) Yaqui, as the following pair of root and embedded interrogatives illustrates.

(40)   Sonora Yaqui (Uto-Aztecan)
    a.   *Hakwóo-sa   yebí-nee*
          when-Q   arrive-FUT
          'When will he arrive?' (Dedrick & Casad 1999: 93)
    b.   *Hunáman   'émpo   téhwaa-tu-nee   hitása   'ém   hóo-nee-'u*
          there   you   tell-PASS-FUT   what   you   do-FUT-NMLZ
          'There you will be told what you are to do.' (Dedrick & Casad 1999: 378)

The "gerundial" marker *'u* is added onto the tensed verb in (40b).[36]

Lack of m-asymmetry seems to be a characteristic of Santali too, according to Neukom (2001: 61).

A positive case of m-asymmetry can be found in Atayal, which has ENFIs, as shown in (41) (Rau 1992: 58), built on "subjunctive" verbal forms. These forms differ from "independent" verbal forms in disallowing temporal inflection (DPASS = "direct passive", LPASS = "local passive").

---

35. Martuthunira (Dench 1988) seems to have some m-asymmetric constructions, namely, non-finite relatives, "*lest* clauses", i.e., clauses describing "expected unwelcome consequences" of situations described in their main clauses, and purpose clauses. To the extent that purpose clauses may be involved in complementation (Haspelmath 1989, Los 2005), it cannot be ruled out that ENFIs in Martuthunira are built from them, but no examples are attested. Likewise, Gooniyandi has "non-finite clauses [that] form a relatively small class of rather impoverished clauses, characterized by an obligatory non-finite verb" (McGregor 1988: 45). Again, the existence of ENFIs remains to be established. "*Lest* clauses" are also attested for Diyari (Evans 2007: 380).

36. The treatment of the *sa*-marker by Dedrick & Casad (1999) is not fully satisfactory. Because of its allegedly being "bleached in meaning" (Dedrick & Casad 1999: 95) it is neither written as a separate morpheme nor glossed with interrogative pronouns such as *hitása* in embedded clauses. At the same time it is separated and glossed as INDF (!) in (i), i.e., a candidate for showing the [+i=i]-status of (Sonora) Yaqui.

(i)   *Hitá-'apo-sa   bibá-ta   yáa-k*
       thing-in-INDF   cigarette-ACC   make-PRF
       'In some way he made a cigarette' (Dedrick & Casad 1999: 250)

(41)   Atayal (Austronesian)
         [*Nanu' s-aw*]   *iŋat   qbaqan*
          what   go-DPASS.SBJV NEG   know.LPASS
         'They don't know what to do.' (Rau 1992: 53)

Importantly, there is evidence that Atayal is a [%i=i]-language, that is, "pure" indefinite uses of *wh*-words are confined to the scope of operators (Haspelmath 1997: 324–325).

Finally, according to Larry Hyman and Maria Polinsky (personal communication), Aghem lacks genuine clausal subordination. It thus lacks ENFIs and is counted as [−enfi] accordingly.[37]

Let me summarize the results of this admittedly brief and tentative crosslinguistic survey. We have found evidence that among the 62 languages in (27) the following 13 are non-robust [i=i]-languages.

(42)   [%i=i]-languages from sample (27)
         Atayal, Belorussian, Chinese, Dutch, Latin, Lithuanian, Mapuche, Newari, Eastern Pomo, Russian, Old Church Slavonic, Slovene, Ukrainian

These languages comply with generalization (14) by instantiating the consequent of implication (23) ([+enfi] $\Rightarrow$ ¬[+i=i]). Thus, irrespective of whether they have ENFIs or not, they will make (23) true.

Next, we have found evidence that among the 62 languages in (27) the following 16 lack ENFIs.

(43)   [−enfi]-languages from sample (27)
         Aghem, Chinese°, Old English, Classical Greek, German, Gothic, Hmong Njua°, Khmer°, Maricopa, Panyjima°, Vedic Sanskrit, Santali°, Shoshone, Thai°, Xinh Mul°, Yaqui°

The languages lacking ENFIs "trivially", i.e., because they are not m-asymmetric (at all or in the relevant clausal domains), have been marked by superscript °. The [−enfi]-languages comply with generalization (14) by falsifying the antecedent of implication (23).

Together this gives us 28 languages out of the 62 [i=i]-languages for which we have reason to believe that they comply with generalization (14). I take this as encouragement for pursuing the issue further.

In addition, we have found evidence for eight robust [i=i]-languages.

---

37. However, Anderson & Keenan (1985: 302–303) provide conflicting evidence. For the [−enfi]-status of Classical Greek I rely on Peter Staudacher (personal communication).

(44) [+i=i]-languages from sample (27)
German, Gothic, Maricopa, Panyjima, Passamaquoddy, Ancash Quechua, Shoshone, Yup'ik

Crucially, five of these have already been shown to lack ENFIs, as summarized in (43). Checking Passamaquoddy, Ancash Quechua, and Yup'ik for ENFIs remains high on the agenda for further research.

Finally, disregarding the languages introduced in Section 2 that are not in sample (27) because they are [−i=i], there is evidence for eight languages possessing ENFIs.

(45) [+enfi]-languages
Atayal, Dutch, Irish, Lithuanian, Mapuche, Huallaga Quechua, Russian, Slovene

Crucially, six of these have already been shown to be [%i=i]-languages, as summarized in (42). The remaining two, i.e., Irish and Huallaga Quechua, even turn out to be [−i=i]-languages.

Appendix B presents a full summary of our findings in form of a table. Reluctantly, I have to leave filling the remaining gaps for further research.

4. Relating the ENFI-gap to the indefinite-interrogative affinity

The remainder of this article will be devoted to a speculation on how to relate ENFIs to the indefinite-interrogative affinity, in particular, how to relate the absence of ENFIs to the presence of the [+i=i]-property. The main object of study here will be German.

One core "functionalist" intuition is that the development of ENFIs is blocked in languages where these structures would be "hard to recognize". The German pattern could thus be an instance of the tendency toward the avoidance of ambiguities.[38] The task of recognition can be divided into two parts, namely, formal and interpretive.

In order to FORMALLY identify an expression as an interrogative clause, one would have to detect what syntacticians have come to call its "clausal typing" (cf. Brandner 2000, Cheng 1991). This is roughly equivalent to the strategies discussed by Bhat (2000) for turning declaratives containing indefinite pro-

---

38. See Blutner 2000 and articles in Blutner & Zeevat (eds.) 2004 for recent discussion of such "blocking effects". The issue of disambiguation has also been raised in connection with [i=i]-languages by Haspelmath (1997: 170–171) and Rissanen (1987: 419). Arguments based on ambiguity avoidance are notoriously slippery, so a "deeper" account – possibly seeking unification with the work of Sabel (1996, 2006) – will ultimately be called for.

nouns into interrogatives.[39] Strategy 1 consists in using specialized pronouns. This is shown for French in (46).

(46)   French (Indo-European)
    a.   *Marie a   mis le   savon quelque part*
           Mary   has put the   soap some   place
           'Mary put the soap somewhere.'
    b.   *Marie a   mis le   savon oú?*
           Mary   has put the   soap where
           'Where did Mary put the soap?'

Strategy 2 consists in putting the indefinite in a prominent position. This usually means fronting, as shown for German in (47).[40]

(47)   German (Indo-European)
    a.   *Hans hat wen beleidigt*
           Hans has who insulted
           'Hans insulted somebody.'
    b.   *Wen hat Hans beleidigt?*
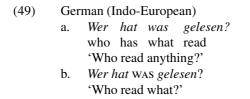   who has Hans insulted
           'Who did Hans insult?'

Strategy 3 achieves prominence for the indefinite by accentuation. This is shown for Chinese in (48).

(48)   Chinese (Sino-Tibetan)
    a.   *Zheli* QUE-LE *shenme*
           here   miss-ASP something
           'Something is missing here.' (Bhat 2000: 379)
    b.   *Zheli que-le* SHENME*?*
           'What is missing here?' (Bhat 2000: 379)

The same strategy is used to disambiguate *in-situ wh*-indefinites in German multiple interrogatives, as shown in (49).

---

39. I sidestep the use of Q-particles or verbal Q-mood (see Bhat 2000). Among the [i=i]-languages in (27) the following have such marking strategies (cf. Bruening 2004): Assuriní, Galibi, Guaraní, Klamath, Koasati, Korean, Lakhota, Lithuanian, Maricopa, Mising, Takelma, Thai, Wintu, Yup'ik. Also, I do not go into the issue of purely *wh-in-situ*-languages. Among the languages considered in this article, at least Chinese, Hopi, Korean, and Thai belong to that class (cf. Bruening 2004, Cheng 1991).

40. According to Cheng (1991: 108–109), Martuthunira and Panyjima use this strategy for disambiguation while Diyari fronted *wh*-pronouns are still ambiguous.

(49) German (Indo-European)
    a. *Wer hat was gelesen?*
        who has what read
        'Who read anything?'
    b. *Wer hat* WAS *gelesen*?
        'Who read what?'

Now, trivially, the most reliable strategy, i.e., (lexical) strategy 1, is unavailable in [+i=i]-languages. To mark *wh*-interrogatives, German uses fronting of one *wh*-indefinite and accentuation for further *wh*-expressions in multiple interrogatives. However, the fronting strategy, i.e., strategy 2, is not reliable when it comes to the attempt of forming ENFIs, given the fact that German is an OV-language with scrambling. (Specific) *wh*-indefinites will easily end up on the left edge of an infinitival constituent independently. This has already been shown above in (6).[41]

Finally, fronted *wh*-indefinites in interrogative clauses are not stressed unless some additional focusing is intended. They are thus not intrinsically identifiable as interrogative pronouns, which means that strategy 3 is equally insufficient for clausal typing in the case of putative German ENFIs.[42]

Let us turn to INTERPRETIVE identification. It is plausible to assume that interrogatives are most easily recognized as such if they possess their standard "erotetic" force, i.e., if they are interpreted as request by the speaker to get some information from the addressee. However, infinitival interrogatives do not seem to allow for such default forces.

Consider the following two scenarios for the possible rise of ENFIs in German. According to scenario 1, they could arise as complements to verbs allowing both infinitival and interrogative complements. A small set of such verbs can be collected on the basis of work by Sabel (1996: 278) and Bhatt (2006: 111–117): *vergessen* ('forget'), *entscheiden* ('decide'), *erklären* ('state'), *lernen* ('learn'), *mitteilen* ('inform'), *(sich) überlegen* ('ponder'). None of these

---

41. It goes without saying that the same effect arises in *to*-less ("bare") infinitives. One systematic exception is relative order with respect to weak personal pronouns. Thus, (ib) is only acceptable if *ihr* is narrowly focused.

(i) a. *Er hat versucht* [*ihr was zu sagen*]
         he has tried her.DAT what.ACC to tell
         'He tried to tell her something'
     b. *Er hat versucht* [*was* ??*ihr* / IHR *zu sagen*]

Structures like (i) would thus be a factor pushing German in the direction of developing ENFIs.

42. Lisa Selkirk (personal communication) cautions me that this claim has to be backed up by experimental evidence. It can be added that intuitions about "stress" are easily confounded with intuitions about intonational phrasing.

verbs denote default acts of seeking information. Thus their complements would not easily be recognizable as interrogatives from interpretation alone.[43]

According to scenario 2, ENFIs would start out as embedded root phenomena, i.e., as root clauses that undergo some form of pseudo-embedding (cf. Heycock 2006, Hooper & Thompson 1973). However, for German it has been shown that non-V2-interrogatives (i.e., V-final or infinitival ones) possess rather peculiar speech act values when in root position (Reis 2003, Truckenbrodt 2004). Reis (2003: 189–190) discusses a context in which a customer visits a travel agency inquiring at the reception desk who to consult. This can felicitously be expressed by a V2-interrogative (50a) but not by an infinitival one (50b).[44]

(50)   German (Indo-European)
       a. *An   wen   soll   ich   mich   wenden?*
           at   who   shall   I   me   turn.INF
           'Who should I talk to?'
       b. #*An   wen   sich   wenden?*
           at   who   oneself   turn.INF

Non-V2-interrogatives in root position only function as " 'uncertainty' questions, with the uncertainty implying a deliberative attitude toward the question raised, and thus inducing self-directedness" (Reis 2003: 191). Again, this is not the kind of canonical force of seeking information that would help identify a structure as interrogative from interpretation alone.[45]

To summarize, formal as well as interpretive aspects conspire against the use of *wh*-infinitives as embedded non-finite interrogatives in German. None of the available strategies of clausal typing, semantic interpretation, or pragmatic assignment of illocutionary force allows a straightforward identification of the relevant structures as interrogatives.

This tentative perspective on the relation between ENFIs and the [+i=i]-property in German is a useful starting point for further research. Let me sketch

---

43. English *wh*-infinitives have been shown to possess a peculiar modal component which reduces the environments they can be embedded into (Bhatt 2006: 105): **John predicted who to invite*, **It is important who to talk to at the party*, **What to do depends on where to be.* The behavior of "conjecture-predicates" like *predict* is more complicated, though, as the acceptability of *It is usually difficult to predict how much money to invest in order to make some profit* shows (confirmed by Laura Downing, Bryan Jurish, and Greg Kobele, personal communication).

44. Note that German only allows bare infinitives in such a function. This distinguishes it from Dutch (and English) where similar reflective questions are expressed by *to*-infinitives, as shown in *Wat *(te) doen?* [what to do] (Zwart 1993: 102).

45. The very existence of root *wh*-infinitives like (51b) in German has to be made compatible with this blocking account of the absence of German ENFIs, as Marga Reis (personal communication) rightly points out.

two areas where results are already beginning to emerge. First, historical evidence is available for the development of ENFIs in English. This can help clarify the two developmental scenarios previously sketched. Secondly, the clausal typing profile of Korean, another [+i=i]-language in sample (27), differs from the one in German. This makes interesting predictions for the possibility of Korean ENFIs.

As already indicated in (27), Old English is an [i=i]-language and there is evidence that in fact it is [+i=i]. (51) shows this (cf. Rissanen 1987: 416).

(51) Old English (Indo-European)
  Nu wille we eow hwæt lytles be him gereccan
  now want we you what of.little about him tell.INF
  'Now we want to relate to you a little thing concerning him.' (Fischer et al. 2000: 142)

At the same time, it has been observed that Old English lacks ENFIs (Fischer et al. 2000: 95). Their functional counterparts are subjunctives, as we have seen for Gothic (Section 3). An example is given in (52).[46]

(52) Old English (Indo-European)
  þæt hy ne bodian ælcon men [hwæt him
  that they not preach.SBJV each man.DAT what him
  sy to donne]
  is.SBJV to do
  'to tell anyone what they should do' (Los 2005: 113)

Quite strikingly, the [−enfi, +i=i]-language Old English turns into the [+enfi, −i=i]-language Middle English. (53) provides an early ENFI.[47]

(53) Middle English (Indo-European)
  ant nuste [hwet seggen]
  and knew.not what say.INF
  'and didn't know what to say' (Fischer et al. 2000: 96)

A PPCME (Penn-Helsinki Parsed Corpus of Middle English) corpus search yielded 20 ENFIs for the period between the years 1225 and 1450, 17 of which are complements to the negation of the verb *know*; (53) is one instance. The English facts are thus compatible with scenario 1, i.e., rise of ENFIs in the complement of predicates allowing both infinitival and interrogative arguments. (54) gives an infinitival complement of *know* from Old English.

---

46. The negation *ne* marks concord with a higher negation.

47. Establishing the [−i=i]-nature of Middle English requires further work. So far I have to go by absence of counterexamples. For relevant discussion see Mustanoja 1960.

(54)   Old English (Indo-European)
        forþon   ic   wat   me   to   beonne   unscadwis   on   swa
        because   I   know   me   to   be.INF   unskilful   in   such
        deorwurðra   spræca
        worthy   languages
        'because I know I am not skilful in such worthy languages' (Fischer
        et al. 2000: 222)

Let me turn to Korean, which poses the most serious challenge I am aware
of to generalization (14). First of all it is a robust [+i=i]-language as (55a)
indicates (cf. Haspelmath 1997: 314). Also, Korean allows what may look like
ENFIs, as shown in (55b).

(55)   Korean (isolate)
        a.   *Pakk-ey   nwu-ka   wa-ss-ta*
              outside-at   who-NOM   come-PST-DECL
              'Somebody has come outside.' (Sohn 1994: 21)
        b.   *Mary-nun   hangsang   [mwues-ul   ilk-eya   ha-l-ci]*
              Mary-TOP   always   what-ACC   read-FUT   have.to-PROS-Q
              a-n-ta
              know-PRS-DECL
              'Mary always knows what to read.' (Shin-Sook Kim, personal
              communication)

On the basis of the discussion by Sohn (1994: 61–68), the putative ENFI in
(55b) must be analyzed as an "adjectival clause" headed by *ha-l* in apposition
to the "defective noun" *ci* translated as 'whether'. The verb *ha-l* contains the
prospective mood marker *l* followed by a zero "modifier suffix". Whether or
not m-asymmetry can be postulated here is a difficult question. If Sohn (1994:
355) is correct in stating that a "linguistically significant finite-nonfinite divi-
sion must be based on the presence and absence of sentence enders that con-
tain a speech level and a sentence-type", everything hinges on the status of *ci*.
Whether or not *ci* can be assumed to be such a sentence ender will depend
on further research.[48] In any case, *ci* contributes to the clausal typing of the
putative ENFI in (55b) and thus "recognition" as interrogative. Also, it seems
that Korean *wh*-pronouns have to be stressed if they are used interrogatively

---

48. See also Bisang 2007 for some discussion of finiteness in Korean. Another open issue is the
    internal structure of the putative ENFI in (55b), in particular the position of the interrogative
    pronoun. Also, it should be noted that an overt nominative subject, e.g., *Bill-i* would be li-
    censed in (55b), i.e., *Mary-nun hangsang [Bill-i mwues-ul ilk-eya ha-l-ci] a-n-ta*, changing
    the interpretation to 'Mary always knows what Bill has to read'. Thanks to Shin-Sook Kim
    (personal communication) for pointing this out.

(Shin-Sook Kim, personal communication). Korean thus differs from German in providing sufficiently clear clausal typing for ENFIs, which – from the perspective developed above – makes their co-existence with a [+i=i]-property more plausible.

5. Summary and outlook

This article has been meant as an attempt to shed new light on the well-known syntactic difference between English (*Mary suddenly remembered* [*where to find the keys*]) and German (**Maria erinnerte sich plötzlich* [*wo die Schlüssel (zu) finden*]) as exemplified in (1) above. It was pointed out – for the first time, so far as I know – that the languages that are usually grouped together with English as possessing such embedded non-finite interrogatives (ENFIs) (Sabel 1996, 2006) also share another property: they strictly distinguish interrogative (*who*) from pure indefinite (*someone*) pronouns. These are the languages in (56).

(56)   Basque, English, French, Modern Hebrew, Italian, Polish, (European) Portuguese, Spanish

German differs from these in both dimensions: it lacks ENFIs and it allows interrogative and pure indefinite uses of (most of) its *wh*-pronouns.

In a second step, another group of languages was identified that allows ENFIs like the languages in (56) but, at the same time, limited dual use of *wh*-pronouns is also attested. The limitations – not shared by German – are either (i) paradigmatic: very few items possess the indefinite/interrogative ambiguity, (ii) distributional: *wh*-pronouns in their pure indefinite use are constrained to narrowly defined syntactic contexts, or (iii) semantic: *wh*-pronouns in their pure indefinite use are limited to non-specific readings. This group of languages is given in (57).[49]

(57)   Dutch, (Latin), Lithuanian, Russian, Slovene

On the basis of these observations, generalization (14), which links the presence or absence of ENFIs in particular languages to their inventory of indefinite and interrogative pronouns, was formulated: If a language *L* possesses embedded non-finite (*wh-*)interrogatives, then the pronominal system of *L* does not possess any robust indefinite/interrogative ambiguity. The three-way split induced by the indefinite/interrogative ([i=i]-)ambiguity thus recognizes languages that lack the ambiguity completely like the ones in (56), ([−i=i]), languages that possess a "non-robust" ambiguity like the ones in (57), ([%i=i]),

---

49. It could not be conclusively established whether Latin possesses ENFIs.

and languages with a robust ambiguity like German, ([+i=i]). In abbreviated form, generalization (14) then reads as in (23): [+enfi] $\Rightarrow$ ¬[+i=i].

Independent evidence for the generalization was provided from German dialects (Section 2) and the historical development of English (Section 4). Thus, evidence was presented that Pennsylvania German and Reichenau German both possess ENFIs while at the same time the former is [−i=i] and the latter at least not [+i=i]. Also, it could be shown that Old English was [−enfi, +i=i] like German while Middle English changed into [+enfi, −i=i] like Modern English.

Section 3 undertook a broader crosslinguistic survey starting from 62 [i=i]-languages (i.e., [+i=i] or [%i=i]) listed in (27). A detailed summary of the search and its results is provided at the end of Section 3. Pulling things together from all sections we end up with the following crucial results. (See Appendix B for a full chart of languages discussed.)

There is evidence for 19 languages, listed in (58), that possess ENFIs. These can be exhaustively grouped into [−i=i]-languages, (58a), and [%i=i]-languages, (58b). Thus, they all comply with generalization (14) by instantiating the consequent of implication (23).

(58)   [+enfi]-languages
    a.   [−i=i]:   Basque, English, Middle English, French, Modern Hebrew, Irish, Italian, Pennsylvania German, Polish, (European) Portuguese, Huallaga Quechua, *Spanish*
    b.   [%i=i]:   Atayal, Dutch, Reichenau German,[50] Lithuanian, Mapuche, Russian, Slovene

Also, there is evidence for 23 languages, listed in (59), that lack ENFIs. They all comply with generalization (14) by falsifying the antecedent of implication (23).

(59)   [−enfi]-languages
    a.   [+i=i]:   Old English, German, Gothic, Maricopa, Panyjima, Shoshone
    b.   [+i=i]/[%i=i]:   Aghem, Classical Greek, Hmong Njua, Khmer, Vedic Sanskrit, Santali, Thai, Xinh Mul, Yaqui
    c.   [%i=i]:   Chinese
    d.   [−i=i]:   Danish, Scottish Gaelic, Hindi/Urdu, Hungarian, Icelandic, Norwegian, Swedish

---

50. It is still open whether Reichenau German is [%i=i] or [−i=i].

The most interesting next step for further research suggested by (59) is to establish the exact [i=i]-status of the nine languages in (59b). If they turn out to possess robust [i=i]-ambiguities, this would be compatible with generalization (14).

In addition, it was possible to provide evidence concerning the [i=i]-status of nine languages with undecided [enfi]-status. They are listed in (60).

(60)   [±enfi]-languages
      a.   [+i=i]:   Passamaquoddy, Ancash Quechua, Yup'ik
      b.   [%i=i]:   Belorussian, Latin, Newari, Eastern Pomo, Old Church Slavonic, Ukrainian

Obviously, clarifying whether or not Passamaquoddy, Ancash Quechua, and Yup'ik possess ENFIs is highest on the agenda for further research.

Finally, Section 4 was devoted to finding reasons for linking the absence of ENFIs with the presence of a robust [i=i]-ambiguity. It was argued that in the case of German, ENFIs would be hard to recognize as interrogatives due to (i) lack of unambiguous clausal typing – lack of specialized interrogative pronouns being one of the factors responsible for that lack – and (ii) lack of unambiguous identification as interrogatives in terms of interpretational properties, given (i) the peculiar semantics of predicates selecting both infinitivals and interrogatives and (ii) the peculiar illocutionary force potential of infinitival interrogatives when used as independent clauses.

In the context of this discussion, the most convincing candidate for being a counterexample to generalization (14), Korean, was analyzed. Korean is a [+i=i]-language for which the ENFI-like structure in (55b) has been reported on. Crucially, this structure possesses unambiguous clausal typing, given that (i) the indefinite prounoun has to be accented on its interrogative use and (ii) there is a clause-final Q-particle. However, it remains to be seen whether (55b) must be classified as non-finite or finite, given that finiteness in Korean is encoded by "sentence enders" (Bisang 2007, Sohn 1994) and given that it cannot be ruled out that the Q-particle *ci*, involved in (55b), belongs to that class of markers.

In sum, I hope to have shown that generalization (14) not only sheds interesting further light on the original contrast between English and German in (1), but that it has considerable crosslinguistic potential. I have to leave the investigation of the remaining languages in (27) as well as the unification of this approach with the one by Sabel (1996, 2006), which covers non-finite relatives and polar interrogatives in addition, for further research.

Appendix A: The account of Sabel 1996 summarized

Sabel (1996: 296) assumes the following parametrization of infinitival C°.

(A1)   C°_{inf} is [−/+strong], if the infinitival C-system can/cannot be overtly filled with a base-generated element

In addition, Sabel assumes that C°_{[+strong]} triggers X°-to-C°, i.e., AgrS°-to-C° in his system. Also, the *Wh*-Criterion is taken to hold.

(A2)   *Wh*-Criterion (Rizzi 1996: 64)
    a.   Each X°_{+wh} must be in a Spec-Head relation with a *Wh*-phrase
    b.   Each *Wh*-phrase must be in a Spec-Head relation with an X°_{+wh}

In embedded (selected) interrogatives [+*wh*] is introduced via choice of an appropriate C°_{+wh} (licensed by government). Following Rizzi (1996: 66), Sabel assumes that AgrS°-to-C°_{+wh} overwrites the feature [+*wh*]. Thus, *wh*-infinitives are incompatible with AgrS°-to-C°, i.e., incompatible with a strong C°, because of a *Wh*-Criterion violation.

The initial stages of deriving (1a) in English, whose C°_{inf} is [−strong], would look like in (A3).

(A3)   a.   [_{C'} C°_{inf,wh,agr} [_{AGRP} PRO [_{AGR'} AGR°_{agr} [_{TP} *to find the keys where*]]]]

b. $[_{C'} C°_{inf,wh,\text{agr}} [_{AGRP} \text{PRO} [_{AGR'} AGR°_{\text{agr}} [_{TP} \textit{to find the keys where}]]]]$

Following Chomsky (2000) one can interpret "weak" C° as involving a weak [*agr*]-feature. Thus, the transition from (A3a) to (A3b) would be brought about by pure "Agree". [*wh*] in C° is left unaffected and *wh*-movement can still apply.

In contrast, (A4) shows the German situation, where C°_{inf} is [+strong].

(A4)   a. $[_{C'} C°_{inf,wh,AGR} [_{AGRP} \text{PRO} [_{AGR'} AGR°_{agr} [_{TP} \textit{die Schlüssel wo zu finden}]]]]$

b. $[_{C'} [_{C°} [AGR°_{\text{agr}}]_i C°_{inf,\text{AGR}}] [_{AGRP} \text{PRO} [_{AGR'} t_i [_{TP} \textit{die Schlüssel wo zu finden}]]]]$

Strong [*AGR*] in C° has to be eliminated by "Move", which – by stipulation – overwrites [*wh*]. Thus, no subsequent *wh*-movement to C° of an infinitive is possible and (1b) is ungrammatical.

Appendix B: Crosslinguistic survey summarized

"nf" gives non-finite forms: an (action nominal), cv (converb), inf (infinitive), vn (verbal noun), ±nf = presence/absence of another m-asymmetric form. "Example" provides numbers of example sentences or forms. "Section" mentions the section in which a language is discussed. "Footnote" mentions a footnote in which a language is discussed.

Language	[i=i]	[enfi]	[nf]	Example	Section	Footnote
Aghem	+/%	−			3	20, 37
Assuriní	+/%				3	39
Atayal	%	+	+	(41)	3	
Bagandji	−				3	20
Basque	−	+		(9a)	2	18
Belorussian	%				2, 3	
Burushaski	+/%		cv, inf	(32)	3	26, 28
Chinese	%	−	−	(48)	3, 4	21, 39
Danish	−	−	inf		2	
Diyari	+/%				3	35, 40
Djaru	+/%				3	
Dutch	%	+	inf	(12), (13b), Footnote 44	2–4	44
Dyirbal	+/%				3	
English	−	+	inf		1–4	
Middle English	−	+	inf	(53)	4	47
Old English	+	−	inf	(51), (52), (54)	3,4	20, 26, 46
Finnish Sign Language	+/%				3	20
French	−	+	inf	(7), (46)	2, 4	27
Scottish Gaelic	−	−	vn	(33)	3	30

Language	[i=i]	[enfi]	[nf]	Example	Section	Footnote
Galibi	+/%				3	39
German	+	−	inf		1–4	
Old High German	+/%		inf		2,3	
Pennsylvania German	−	+	inf	(25b)	2	
Reichenau German	%	+	inf	(25a), Foot-note 15	2	16, 50
Goajiro	+/%				3	
Gooniyandi	+/%		+		3	35
Gothic	+	−	inf	(21)	2–4	22, 26
Classical Greek	+/%	−	inf		3	26, 37
West Greenlandic			an	(35)		
Guaraní	+/%				3	39
Modern Hebrew	−	+	inf	(9b)	2	
Hindi/Urdu	−	−	inf		3	23
Hmong Njua	+/%	−	−		3	
Hopi	+/%				3	39
Hungarian	−	−	inf	Footnote 27	3	27
Icelandic	−	−	inf		2	4
Irish	−	+	vn		3	
Italian	−	+	inf	(7)	2	
Kaingang	+/%				3	
Kamaiurá	+/%				3	33
Khmer	+/%	−	−		3	
Klamath	+/%				3	39
Koasati	+/%				3	39
Korean	+	±	cv, ?	(55)	3, 4	39, 48
Lakhota	+/%		an		3	39
Latin	%	±	inf	(15), (16)	2, 3	8, 49
Lezgian			cv		3	29
Lithuanian	%	+	inf	(24a)	2, 3	27, 39
Mangarayi	+/%				3	
Mapuche	%	+	an	(36)	3	
Maricopa	+	−	an	(37), (38)	3	33, 39
Martuthunira	+/%		+		3	35, 40
Masalit	+/%				3	
Mising	+/%				3	39
Mongolian					3	31
Mundari	+/%		cv		3	
Newari	%				3	
Norwegian	−	−	inf		2	
Panare	+/%				3	
Panyjima	+	−	−	(39)	3	34, 40
Pashto	+/%		inf		3	26, 28
Passamaquoddy	+				3	22, 27
Polish	−	+	inf	(7)	2, 3	11
Eastern Pomo	%				3	
(European) Portuguese	−	+	inf	(7)	2	
Ancash Quechua	+		cv, inf		3	26, 28
Huallaga Quechua	−	+	inf	Footnote 28	3	28
Russian	%	+	inf	(24a)	2, 3	14, 26
Vedic Sanskrit	+/%	−	cv, inf		3	27

Language	[i=i]	[enfi]	[nf]	Example	Section	Footnote
Santali	+/%	−	−		3	
Shoshone	+	−	inf		3	26
Siuslaw	+/%				3	
Old Church Slavonic	+/%		inf		2, 3	25
Slovene	%	+	inf	(17), (24a)	2, 3	
Spanish	−	+	inf	(7)	2	
Swedish	−	−	inf		2	4
Takelma	+/%				3	39
Tamil					3	29
Thai	+/%	−	−		3	25, 39
Ukrainian	+/%		inf		2, 3	26
Warndarang	+/%				3	
Wintu	+/%				3	39
Xinh Mul	+/%				3	
Yaqui	+/%	−	−	(40), Footnote 36	3	36
Yidini	+/%				3	
Yindjibarndi	+/%				3	
Yup'ik	+				3	22, 39

References

Adger, David. 2007. Three domains of finiteness: A minimalist perspective. In Nikolaeva (ed.) 2007, 23–58.

Ambrazas, Vytautas (ed.). 1997. *Lithuanian grammar*. Vilnius: Baltos Lankos.

Amir Coffin, Edna & Shmuel Bolozky. 2005. *A reference grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Amir-Babenko, Svetlana & Franz Pfiegl. 2005. *Praktische Kurzgrammatik der ukrainischen Sprache*. Hamburg: Buske.

Anderson, Stephen & Edward Keenan. 1985. Deixis. In Shopen (ed.) 1985, 259–308.

Austin, Peter (ed.). 1988. *Complex sentence constructions in Australian languages*. Amsterdam: Benjamins.

Banfield, Ann. 1973. Narrative style and the grammar of direct and indirect speech. *Foundations of Language* 10. 1–39.

Barrie, Michael. 2007. Control and WH-infinitivals. In William Davies & Stanley Dubinsky (eds.), *New horizons in the analysis of control and raising*, 263–279. Dordrecht: Springer.

Barwise, Jon & John Perry. 1983. *Situations and attitudes*. Cambridge, MA: MIT Press.

Bayer, Josef & Ellen Brandner. 2004. Klitisiertes *zu* im Bairischen und Alemannischen. In Franz Patocka & Peter Wiesinger (eds.), *Morphologie und Syntax deutscher Dialekte und historische Dialektologie des Deutschen*, 160–188. Wien: Praesens.

Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager*. Wiesbaden: Harrassowitz.

Beukema, Frits & Marcel den Dikken. 1989. The position of the infinitival marker in the Germanic languages. In Dany Jaspers, Willem Klooster, Yvan Putseys & Pieter Seuren (eds.), *Sentential complementation and the lexicon*, 57–75. Dordrecht: Foris.

Bhat, D. N. S. 2000. The indefinite-interrogative puzzle. *Linguistic Typology* 4. 365–400.

Bhat, D. N. S. 2004. *Pronouns*. Oxford: Oxford University Press.

Bhatt, Rajesh. 2006. *Covert modality in non-finite contexts*. Berlin: Mouton de Gruyter.

Bisang, Walter. 1995. Verb serialization and converbs: Differences and similarities. In Haspelmath & König (eds.) 1995, 137–188.

Bisang, Walter. 2001. Finite vs. non finite languages. In Haspelmath et al. (eds.) 2001, 1400–1413.

Bisang, Walter. 2007. Categories that make finiteness: Discreteness from a functional perspective and some of its repercussions. In Nikolaeva (ed.) 2007, 115–137.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17. 189–216.

Blutner, Reinhard & Henk Zeevat (eds.). 2004. *Optimality theory and pragmatics*. Houndmills: Palgrave Macmillan.

Bolkestein, A. Machtelt. 1996. Reported speech in Latin. In Theo A. J. M. Janssen & Wim van der Wurff (eds.), *Reported speech*, 121–140. Amsterdam: Benjamins.

Brandner, Ellen. 2000. Scope marking and clausal typing. In Uli Lutz, Gereon Müller & Arnim von Stechow (eds.), *Wh-scope marking*, 45–75. Amsterdam: Benjamins.

Brandner, Ellen. 2004. Microvariation within German *to*-infinitives. Unpublished manuscript, Universität Konstanz.

Brandon, Frank Roberts & Lucy Ferreira Seki. 1984. Moving interrogatives without an initial +WH node in Tupí. In Cook & Gerdts (eds.) 1984, 77–103.

Bruening, Benjamin. 2004. Wh-indefinites, question particles, and Wh-in-situ: There is no relation. Unpublished manuscript, University of Delaware.

Bruening, Benjamin. 2007. *Wh*-in-situ does not correlate with *Wh*-indefinites or question particles. *Linguistic Inquiry* 38. 139–166.

Cheng, Lisa L. S. 1991. *On the typology of wh-questions*. Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step*, 89–155. Cambridge MA: MIT Press.

Cole, Peter. 1984. Clause reduction in Ancash Quechua. In Cook & Gerdts (eds.) 1984, 105–121.

Comrie, Bernard. 1990. Russian. In Comrie (ed.) 1990, 329–347.

Comrie, Bernard (ed.). 1990. *The world's major languages*. Oxford: Oxford University Press.

Comrie, Bernard & Sandra A. Thompson. 1985. Lexical nominalization. In Shopen (ed.) 1985, 349–398.

Cook, Eung-Do & Donna B. Gerdts (eds.). 1984. *The syntax of Native American languages* (Syntax and Semantics 16). Orlando, FL: Academic Press.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.

Dahl, Östen. 1999. Review of Haspelmath 1997. *Linguistics and Philosophy* 22. 663–678.

Dayley, Jon. 1989. *Tümpisa (Panamint) Shoshone grammar*. Berkeley, CA: University of California Press.

Dedrick, John & Eugene Casad. 1999. *Sonora Yaqui language structures*. Tucson, AZ: University of Arizona Press.

Dench, Alan. 1988. Complex sentences in Martuthunira. In Austin (ed.) 1988, 98–139.

Dench, Alan. 1991. Panyjima. In Robert Dixon & Barry Blake (eds.), *The handbook of Australian languages, Vol. 4*, 125–243. Oxford: Oxford University Press.

Duffley, Patrick & Peter Enns. 1996. Wh-words and the infinitive in English. *Lingua* 98. 221–242.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Evans, Nicholas. 2007. Insubordination and its uses. In Nikolaeva (ed.) 2007, 366–431.

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. *The syntax of Early English*. Cambridge: Cambridge University Press.

Fodor, Janet D. & Ivan A. Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5. 355–398.

Franks, Steven & James Lavine. 2006. Case and word order in Lithuanian. *Journal of Linguistics* 42. 239–288.

Gärtner, Hans-Martin & Markus Steinbach. 1994. Economy, verb second, and the SVO-SOV distinction. *Working Papers in Scandinavian Syntax* 53. 1–59.

Giusti, Giuliana. 1986. On the lack of Wh-infinitives with 'zu' and the projection of COMP in German. *Groninger Arbeiten zur Germanistischen Linguistik* 28. 115–169.

Gordon, Lynn. 1986. *Maricopa morphology and syntax*. Berkeley, CA: University of California Press.

Grosu, Alexander. 2004. The syntax and semantics of modal existential *Wh* constructions. In Olga Tomić (ed.), *Balkan syntax and semantics*, 405–438. Amsterdam: Benjamins.

Haspelmath, Martin. 1989. From purposive to infinitive: A universal path of grammaticalization. *Folia Linguistica Historica* 10. 287–310.

Haspelmath, Martin. 1995a. The converb as a cross-linguistically valid category. In Haspelmath & König (eds.) 1995, 1–55.

Haspelmath, Martin. 1995b. Contextual and specialized converbs in Lezgian. In Haspelmath & König (eds.) 1995, 415–440.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Clarendon.

Haspelmath, Martin. 2005. Indefinite pronouns. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 190–193. Oxford: Oxford University Press.

Haspelmath, Martin & Ekkehard König (eds.). 1995. *Converbs in cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.). 2001. *Language typology and language universals: An international handbook*. Berlin: Walter de Gruyter.

Hercus, Luise. 1982. *The Bagandji language* (Pacific Linguistics B-67). Canberra: Australian National University

Herrity, Peter. 2000. *Slovene*. London: Routledge.

Heycock, Caroline. 2006. Embedded root phenomena. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, Vol. 2, 174–209. Oxford: Blackwell.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4. 465–497.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Hudak, Thomas John. 1990. Thai. In Comrie (ed.) 1990, 757–775.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Partee (eds.), *Quantification in natural languages*, 451–486. Dordrecht: Kluwer.

Jacobson, Steven. 1995. *A practical grammar of the Central Alaskan Yup'ik Eskimo language*. Fairbanks, AK: Alaska Native Language Center.

Jónsson, Jóhannes Gísli. 1996. Clausal architecture and case in Icelandic. Amherst, MA: University of Massachusetts doctoral dissertation.

Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive*. Cambridge: Cambridge University Presss.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3–44.

König, Ekkehard & Johan van der Auwera. 1988. Clause integration in German and Dutch conditionals, concessive conditionals, and concessives. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 101–133. Amsterdam: Benjamins.

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.

Koptjevskaja-Tamm, Maria. 1994. Finiteness. In Ronald E. Asher (ed.), *The encyclopedia of language and linguistics*, 1245–1248. Oxford: Pergamon.

Lamb, William. 2001. *Scottish Gaelic*. München: Lincom Europa.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory* 22. 811–877.

Lehmann, Winfred. 1994. Gothic and the reconstruction of Proto-Germanic. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages*, 19–37. London: Routledge.

Lorenz, Manfred. 1979. *Lehrbuch des Pashto*. Leipzig: Enzyklopädie.

Los, Bettelou. 2005. *The rise of the to-infinitive*. Oxford: Oxford University Press.

Lunt, Horace. 1959. *Old Church Slavonic grammar*. Den Haag: Mouton.

MacKenzie, David N. 1990. Pashto. In Comrie (ed.) 1990, 547–565.

Marušic, Franc. 2005. On non-simultaneous phases. Stony Brook, NY: Stony Brook University doctoral dissertation.

McGregor, William. 1988. Mood and subordination in Kuniyanti. In Austin (ed.) 1988, 37–67.

McLendon, Sally. 1975. *A grammar of Eastern Pomo*. Berkeley, CA: University of California Press.

Mustanoja, Tauno. 1960. *A Middle English syntax*. Helsinki: Société Néophilologique.

Neukom, Lukas. 2001. *Santali*. München: Lincom Europa.

Nikolaeva, Irina. 2007a. Introduction. In Nikolaeva (ed.) 2007, 1–19.

Nikolaeva, Irina. 2007b. Constructional economy. In Nikolaeva (ed.) 2007, 138–180.

Nikolaeva, Irina (ed.). 2007. *Finiteness*. Oxford: Oxford University Press.

Perlmutter, David. 2007. In what ways can finite and nonfinite clauses differ? Evidence from Russian. In Nikolaeva (ed.) 2007, 250–304.

Platzack, Christer. 1986. Comp, Infl, and Germanic word order. In Lars Hellan & Kirsti Koch-Christensen (eds.), *Topics in Scandinavian syntax*, 185–234. Dordrecht: Kluwer.

Postma, Gertjan. 1994. The indefinite reading of *WH*. In Reineke Bok-Bennema & Crit Cremers (eds.), *Linguistics in the Netherlands 1994*, 187–198. Amsterdam: Benjamins.

Quirk, Randolph & Charles L. Wrenn. 1989. *An Old English grammar*. London: Routledge.

Rau, Der-Hwa Victoria. 1992. A grammar of Atayal. Ithaca, NY: Cornell University doctoral dissertation.

Reis, Marga. 1991. Was konstituiert *w*-Interrogativsätze. *Arbeitsberichte des SFB 340* 6. 35–68.

Reis, Marga. 2003. On the form and interpretation of German *Wh*-infinitives. *Journal of Germanic Linguistics* 15. 155–201.

Rissanen, Matti. 1987. Old English indefinite pronouns meaning 'some' and 'any', with special reference to *HW*-forms. In Leena Kahlas-Tarkka (ed.), *Neophilologica Fennica*, 411–428. Helsinki: Société Néophilologique.

Rizzi, Luigi. 1996. Residual verb second and the Wh-criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and functional heads*, 63–90. Oxford: Oxford University Press.

Ruys, Eddy. 1993. *The scope of indefinites*. Utrecht: Universiteit Utrecht doctoral dissertation.

Sabel, Joachim. 1996. *Restrukturierung und Lokalität*. Berlin: Akademie-Verlag.

Sabel, Joachim. 2006. Impossible infinitival interrogatives and relatives. In Patrick Brandt & Eric Fuß (eds.), *Form, structure, and grammar*, 243–254. Berlin: Akademie-Verlag.

Shopen, Timothy (ed.). 1985. *Language typology and syntactic description*, Vol. 3*: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.

Smeets, Ineke. 2008. *A grammar of Mapuche*. Berlin: Mouton de Gruyter.

Sohn, Ho-Min. 1994. *Korean*. London: Routledge.

Streitberg, Wilhelm. 1981. *Gotische Syntax*. Heidelberg: Winter.

Szabolcsi, Anna & Frans Zwarts. 1997. Weak islands and an algebraic semantics for scope taking. In Anna Szabolcsi (ed.), *Ways of scope taking*, 217–262. Dordrecht: Kluwer.

Tappe, Hans Thilo. 1984. On infinitival clauses without COMP. In Wim de Geest & Yvan Putseys (eds.), *Sentential complementation*, 227–237. Dordrecht: Foris.

Thráinsson, Höskuldur. 1998. Infinitival complements in some Old and Modern Germanic languages. In John Ole Askedal (ed.), *Historische germanische und deutsche Syntax*, 335–363. Frankfurt am Main: Lang.

Tikkanen, Bertil. 2001. Converbs. In Haspelmath et al. (eds.) 2001, 1112–1123.

Truckenbrodt, Hubert. 2004. Zur Strukturbedeutung von Interrogativsätzen. *Linguistische Berichte* 199. 313–350.

Ultan, Russell. 1978. Some general characteristics of interrogative systems. In Joseph Greenberg (ed.), *Universals of human language*, Vol. 4*: Syntax*, 211–248. Stanford, CA: Stanford University Press.

Weber, David John. 1983. *Relativization and nominalized clauses in Huallaga (Huanuco) Quechua*. Berkeley, CA: University of California Press.

Weber, David John. 1996. *Una gramática del quechua del Huallaga (Huánuco)*. Lima: Instituto Lingüístico de Verano.

Whitney, William D. 1950. *Sanskrit grammar*. Cambridge, MA: Harvard University Press.

Ylikoski, Jussi. 2003. Defining non-finites: Action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16. 185–237.

Zaefferer, Dietmar. 1991. Weiß wer was? Wer weiß was? Wer was weiß . . . : W-Interrogative und andere w-Konstruktionen im Deutschen. In Marga Reis & Inger Rosengren (eds.), *Fragesätze und Fragen*, 77–93. Tübingen: Niemeyer.

Zeshan, Ulrike. 2004. Interrogative constructions in signed languages: Crosslinguistic perspectives. *Language* 80. 7–39.

Zúñiga, Fernando. 2000. *Mapudungun*. München: Lincom Europa.

Zwart, Jan-Wouter. 1993. *Dutch syntax: A minimalist approach*. Groningen: Rijksuniversiteit Groningen doctoral dissertation.