

**From German Quirk to Universal Tendency:  
A Speculation on (the Absence of) *Wh*-Infinitives\***  
(Hans-Martin Gärtner)

It is well-known that German disallows embedded infinitival (*wh*-)questions. So far, accounts have tended to attribute this gap to more or less obvious properties of complementizers and the C-system. Also, cross-linguistic comparison appears not to have been undertaken to any satisfactory degree. In this paper, I will try to show that this gap is indicative of a rather wide-ranging, if not universal, correlation linking *wh*-infinitives to the inventory of *wh*-expressions. In particular I will defend the claim that if the pronominal system of a language *L* has a robust interrogative/indefinite ambiguity then *L* does not have embedded infinitival (*wh*-)questions. In addition to a cross-linguistic survey, I will provide a functional account for this gap in German based on aspects of clausal typing as well as the illocutionary force of infinitival (root) questions.

It is a well-known fact about the syntax of English and German that they differ wrt the acceptability of embedded infinitival questions. Thus, consider the contrast in (1). ((1b) is a direct translation of (1a)).

- (1) a. *Mary suddenly remembered [ where to find the keys ]* (En)  
b. \* *Maria erinnerte sich plötzlich [ wo die Schlüssel (zu) finden ]* (Ge)

Although various accounts for the unacceptability of (1b) have been put forward, this phenomenon has rarely been discussed in cross-linguistic perspective. Sabel (1996:chapter 8) is a notable exception.<sup>1</sup> In fact, Sabel (1996:295) provides the rather neat generalization in (2).

- (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).<sup>2</sup>

- (3) *They want* [<sub>CP</sub> *for* [<sub>IP</sub> *Mary to put the soap in the fridge* ] ]

Sabel goes on to implement (2) in terms of the “strength” of the (infinitival) head C° interacting with X°-to-C°-movement and Rizzi’s *Wh*-Criterion (Rizzi 1996). This allows him to derive the contrast in (1). I transfer a sketch of this account to Appendix I.

For the sake of exposition, I will abbreviate Sabel’s generalization as in (4) and I will call WH-INF–languages [+w]-languages and OBGC-INF–languages [+o]-languages

- (4) WH-INF ⇒ OBGC-INF

---

\* For comments and suggestions I’m indebted to Joanna Błaszczak, Michael Cysouw, Bostjan Dvořák, Jürg Fleischer, Beáta Gyuris, Andreas Haida, Helen de Hoop, Shin-Sook Kim, Mark Loudon, Andrej Malchukov, Joachim Sabel, Magdalena Schwager, Szymon Śladowicz, Peter Staudacher, Barbara Stiebels, Peter de Swart, Tonjes Veenstra, and Joost Zwarts as well as audiences at the University of Nijmegen and the 7th “Workshop on the Roots of Pragmasemantics” at Szklarska Poręba. Common disclaimers apply.

<sup>1</sup> See also Sabel (1996) for a survey of analyses of (1b).

<sup>2</sup> I assume without argument that Sabel’s assumptions about particular complementizers in particular languages are correct. Nothing much hinges on this, however. There is a well-known contrast for polar questions in English, as shown in (i) (Kayne 1991:665).

(i) He doesn’t know whether / \* if to go to the movies

For reasons that will become clear momentarily, I won’t discuss polar questions in this article.

(5) represents the (small) typology of languages on which Sabel's study is based.<sup>3</sup>

- (5) a. [+w,+o]: *English, French, Italian, Polish, (Eur.) Portuguese, Spanish*  
 b. [+w,-o]:  $\emptyset$   
 c. [-w,+o]: ?  
 d. [-w,-o]: *Danish, German, Norwegian, Swedish*

Now, the point of this paper is to demonstrate that (5) masks another generalization, which – to my knowledge – has gone unnoticed so far. Thus, German word strings involving *wh*-indefinites and *to*-infinitivals are acceptable as soon as a non-interrogative interpretation is available. This is shown in (6).

- (6) a. *Ich habe was zu tun* ( $\approx$  I have something to do)  
 b. *Ich erinnere mich wohin zu fahren* ( $\approx$  I remember going somewhere (specific))

Crucially, Sabel's [+w,+o]-languages do not allow such (re-)interpretations because their sets of interrogative and indefinite pronouns are strictly disjoint. (7) shows this for (counterparts of) *who*.<sup>4</sup>

- (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 isn't – it warrants the following (preliminary) hypothesis.

- (8) If the pronominal system of a language *L* possesses  
 an interrogative/indefinite ambiguity,  
 then *L* does not possess embedded infinitival (*wh*-)questions

Let me call INT/IND-PRO-languages [+i/i]-languages and abbreviate (8) as in (9).

- (9) INT/IND-PRO  $\Rightarrow$   $\neg$ WH-INF

This turns (5) into a small typology in which German appears to assume an isolated position.

- (10) a. [+w,+i/i]:  $\emptyset$   
 b. [+w,-i/i]: *English, French, Italian, Polish, (Eur.) Portuguese, Spanish*  
 c. [-w,+i/i]: *German*  
 d. [-w,-i/i]: *Danish, Norwegian, Swedish*

However, as we are going to see later, German is indicative of a major typological trend here. Still, somewhat embarrassingly, Dutch, i.e. one of German's closest neighbors, is a language

<sup>3</sup>. Note that Sabel's generalization is not a biconditional. Nevertheless it would be useful to know more about languages that make up class (5c), i.e. [-w,+o]-languages. These languages are predicted to be able to develop into [+w,+o]-languages.

<sup>4</sup> The case of Polish may be slightly more complicated. See footnote 6.

that follows Sabel's generalization in being [+w,+o] (Sabel 1996:294f.), but not (8), the latter shown in (11) (cf. Postma 1994).

- (11) *Ik heb wat te lezen* (≈ I have something to read / I have to read something)

It is therefore important to have a closer look at Dutch. 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).” This lack of generality wrt an [i/i]-ambiguity seems to disqualify Dutch as a “real” counterexample to (8). (12) exemplifies the difference between German and Dutch. ((12b) is due to M. Cysouw, p.c.).

- (12) a. *Ich habe wo gelesen, dass der Sommer schön wird* (Ge)  
 b. *Ik heb \*waar / ergens gelezen, dat de zomer mooi word* (Du)  
 (≈ I read somewhere that summer is going to be nice)

Clearly, Dutch lacks a general, or *robust*, [i/i]-ambiguity. I revise (8) accordingly.

- (13) If the pronominal system of a language *L* possesses  
 a robust interrogative/indefinite ambiguity,  
 then *L* does not possess embedded infinitival (*wh*-)questions

Let me note two further instances of non-robustness, namely, Latin and Russian/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).

- (14) a. *Si quid petieritis me in nomine meo, hoc faciam*  
 if what ask.SUBJ.FUT.2SG me in name my that do.SUBJ.1SG  
 “If ye shall ask anything in my name, I will do it”  
 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?”

This strong distributional restriction instantiates a second type of non-robustness.<sup>5</sup>

Russian (A. Malchukov, p.c.) and Slovene (Herrity 2000:109ff.) restrict their bare interrogatives to non-specific readings.<sup>6</sup> (15a) gives an example from Herrity (ibid.), (15b) is due to B. Dvořák (p.c.).

<sup>5</sup> It has been noted in the literature (e.g. Gärtner and Steinbach 1994:37) that German *w*-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).

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

(ia) can only mean “Who called?” but not “Someone called,” which has to be expressed as in (ib). For some speakers the effect seems to be weaker, especially if it bears an L\*H (“topic”) accent (Reis 1991:45). The restriction in (i) is sometimes mistakenly interpreted as a ban on *w*-indefinites in initial, or procliticized, position (Haspelmath 1997:170). However, as has been shown by Reis (1991:44), such a characterization is incorrect. (ii) illustrates this.

(ii) *Wen Nettes haben wir hier noch nie gesehen*

This can mean, “We have never seen anyone (who is) nice”

<sup>6</sup> According to Cheng (1991:104) this is true of Polish as well. Slovene is a particularly interesting case – meriting further study – because Slovene clearly is a [+w]-language (Marušič 2005:chapter 5). Latin possesses what looks like *wh*-infinitives in a special type of *oratio obliqua* (P. Staudacher, p.c.), a phenomenon that may be analyzed as an embedded root phenomenon rather than a case of proper subordination.

- (15) a. *Ali me je kdó iskál?*  
 Q me AUX who look.for  
 “Has anyone been looking for me?”  
 b. *Včeraj me je nekdo / \* kdó poklical*  
 yesterday me AUX someone / who called  
 “Yesterday someone called me”

The distribution of bare interrogative pronouns seems to approximate the one of negative polarity items. This is the third kind of non-robustness I have encountered so far. Thus, if a language puts (strong) paradigmatic, distributional, or semantic restrictions on bare interrogative pronouns when used as indefinites, it does not count as a robust [+i/i]-language.

Let me turn next to a broader typological perspective. A survey based on Cheng (1991), Haspelmath (1997), and Bhat (2000) yielded the following 57 [+i/i]-languages (out of a sample of roughly 150).

- (16) *Assuriní, Atayal, Bagandji, Belorussian, Burushaski, Chinese, Classical Greek, Diyari, Djaru, Dutch, Dyirbal, Galibi, German, Goajiro, Gooniyandi, Gothic, Guarani, Hmong, Njua, Hopi, Kaingang, Kamaiurá, Khmer, Klamath, Koasati, Korean, Lakhota, Latin, Lithuanian, Mangarayi, Mapuche, Maricopa, Martuthunira, Masalit, Mising, Mundari, Newari, Panare, Panyjima, Pashto, Eastern Pomo, Ancash Quechua, Russian, Vedic Sanskrit, Santali, Shoshone, Siuslaw, Old Church Slavonic, Slovene, Takelma, Thai, Ukrainian, Warndarang, Wintu, Xinh Mul, Yaqui, Yidini, Yup'ik.*

In (16), I have already “crossed-out” languages for which I have evidence that they are [–w] (for Classical Greek I rely on Peter Staudacher, p.c.), not robustly [+i/i], as well as languages of the “Far East type” (e.g. Chinese, Thai, Vietnamese, Cambodian, Hmong) because “they are neutral to the finite/non finite distinction” (Bisang 2001:1408). This leaves us with 47 languages of immediate interest.

However, further hypothesis-testing is hampered by various factors. First of all, many grammars are silent about the phenomenon of *wh*-infinitives. Likewise, there seem to be no monographs about them. Secondly, and clearly more importantly, there is no non-trivial cross-linguistically applicable notion of “infinitive.” Relying in part on Cristofaro (2003), I assume that the following [+i/i]-languages for which (8) has to be still tested possess *bona fide* infinitives.<sup>7</sup>

- (17) *Belorussian, Gothic, Lithuanian, Ancash Quechua, Vedic Sanskrit, Shoshone, Old Church Slavonic, Ukrainian*

A closer look at comprehensive studies of Lithuanian (Vyautas 1997) and Sanskrit (Whitney 1950) makes me reasonably confident that these are [–w]-languages.<sup>8</sup>

<sup>7</sup> For Quechua one can rely on Adelaar (2004:1458), who notes that “Ayacucho Quechua, for instance, has an infinitive in *–y* [ . . . ].”

<sup>8</sup> One has to be careful not to misinterpret certain structures on the basis of their translations. Thus Whitney (1950:352) translates 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. It is equally important not to mix up infinitival questions with the “modal existential construction.” Thus Vyautas (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, is an instance of the “modal existential construction,” which should be analyzed as some kind of free relative (see Grosu 2004 for extensive discussion). Such a view actually helps reducing the number of [+i/i]-languages. Hungarian, for example, allows the modal existential construction (Grosu 2004), which involves *bona fide* interrogative pronouns. The same pronouns can, however, not be used as indefinites

In trying to develop a cross-linguistically valid theory of finiteness one can follow Bisang (2001:section 3), who makes the plausible suggestion that a necessary condition for possession of (counterparts of) infinitives requires a language *L* to be “m-asymmetric” as defined below.<sup>9</sup>

- (18) If a language *L* has partially 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

Now, there are well known cases of m-asymmetry, namely, “converb-languages” and “action nominal languages.”<sup>10</sup> In fact 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 our [+i/i]-languages there appear to be at least 5 converb-, i.e. [+cv]-languages, namely those in (19) (cf. Nuolijärvi 1994:1247).

- (19) *Burushaski, Korean, Mundari, Ancash Quechua, ~~Vedic Sanskrit~~*

However, canonically, converbs are used for modification and coordination, not complementation. Haspelmath (1995b:3) characterizes a converb as “a nonfinite verb whose main function is to mark adverbial subordination.” (20) provides an example from Burushaski (Tikkanen 2001:1117), where the converb *nu-qás* functions as a secondary predicate.<sup>11</sup>

- (20) *In(-e)      nu-qás      ásimi*  
 he(-ERG)    CONV-laugh    he.told.me  
 “He said to me laughing . . .”

Since this functional limitation is a general feature of [+cv]-languages, we can discard them from the list of potential counterexamples to (8).

A look at “action nominal-,” or [+an]-languages, raises additional questions. Comrie & Thompson (1985:350) discuss the following nominalization patterns from Lakhota.

---

otherwise. At the same time, Hungarian is [–w] as (i) shows (B. Gyuris, p.c.). The exposition in É. Kiss (2002:202) is misleading in this respect.

- (i) a. *János elfelejtette keresni a kenyeret*  
 J. forgot look.for.INF the bread.ACC  
 “János forgot to look for the bread”  
 b. \**János elfelejtette, (hogy) hol keresni a kenyeret*  
 J. forgot that where look.for.INF the bread.ACC  
 “János forgot where to look for the bread”

<sup>9</sup> This definition allows languages to be p-asymmetric and m-asymmetric at the same time.

<sup>10</sup> I use these labels for expository purposes denoting languages whose m-asymmetric structures are primarily expressed by means of converbs or action nominals. In fact, both phenomena are wide-spread (cf. Haspelmath and König 1995; Koptjevskaja-Tamm 1993).

<sup>11</sup> It is unclear how to analyze the initial pronoun *he*. Presumably it is a topic rather than part of the constituent headed by the converb. Lezgian (Haspelmath 1995:425) 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*  
 C.-DAT selves-GEN to one youth come-IMC see-AOR  
 “Cükwer saw a young man coming toward them.”

Here *qwe-z* is an imperfective converb. A similar case from Tamil is provided by Bisang (1995:157). However, Haspelmath (ibid.) 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 questions as complements of perception verbs.

- (21) a. *gnayá* >> *wógnaye*      b. *wiyuški* >> *wówiyuški*  
 deceive                  deception                  rejoice                  rejoicing

From an (Indo-)European perspective it seems unlikely that action nominals could function as interrogative complements. A thorough look at Koptjevskaja-Tamm (1993) confirms that. The only exception is the following example from West-Greenlandic going back to Fortescue (Koptjevskaja-Tamm 1993:115).<sup>12</sup>

- (22) *umiarsu-up qassi-nut      tikin-ni-ssa-a      nalunngil-ara*  
 ship-REL    how.many-ALL arrive-AN-FUT-3SG.POSS    know-1SG.3SG:INDIC  
 “I know when [sic!] the ship will arrive”

Of course, the richly inflected AN-form *tikin-ni-ssa-a* does not strike one as a particularly convincing exponent of m-asymmetry. The issue of counterexamples to (8) from [+an]-languages will therefore have to be left unresolved. What is clear, though, is that, among the languages in (16), both Lakota and Maricopa (see Gordon 1986) are [+an]-languages.

Let me update our set of languages in (16), crossing out languages doubly if they can be discarded as counterexamples to (8), and crossing them out singly, if their status as counterexamples to (8) is at least doubtful, for example, because m-asymmetry in these languages is likely to be restricted to [+cv] or [+an]. (See Appendix II for a comprehensive chart of language types).<sup>13</sup>

- (23) *Assurini, Atayal, Bagandji, Belorussian, Burushaski, Chinese, Classical Greek, Diyari, Djaru, Dutch, Dyirbal, Hmong-Njua, Hopi, Galibi, German, Goajiro, Gooniyandi, Gothic, Guarani, Kaingang, Kamaiurá, Khmer, Klamath, Koasati, Korean, Lakota, Latin, Lithuanian, Mangarayi, Mapuche, Maricopa, Martuthunira, Masalit, Mising, Mundari, Newari, Panare, Panyjima, Pashto, Eastern Pomo, Ancash Quechua, Russian, Vedic Sanskrit, Santali, Siuslaw, Old Church Slavonic, Shoshone, Slovene, Takelma, Thai, Ukrainian, Warndarang, Wintu, Xinh Mul, Yaqui, Yidini, Yup'ik.*

This leaves us with 40 languages most urgently in need of further testing.

Let me turn to another interesting field of study. It has been discovered that certain dialects of German do possess embedded infinitival *wh*-questions. These are Reichenau German (RG) (Brandner 2004) and Pennsylvania German (PG) (M. Loudén, p.c.). Examples are given in (24).

- (24) a. *I      zoag dir      schnell [ welle Socke      schtopfe ]* (RG)  
 I.NOM    show you.DAT quickly    which socks.ACC mend.INF  
 “I show you quickly which socks to mend”  
 b. *Ich      hab vergesse [ wo      fer      annegehe ]* (PG)  
 I.nom    have forgotten    where    for    there.go.INF  
 “I forgot where to go”

Brandner (2004) goes on to argue that RG is actually a [+o]-language, the inflected preposition *zum* (‘to.the.DAT’) having turned into an infinitival complementizer. Similar assumptions might be made about *fer* (‘for’) in PG. Thus, both RG and PG would seem to confirm Sabel’s generalization (2).

Interestingly, RG and PG equally confirm (8), i.e. they are [–i/i]. For PG I rely on M. Loudén (p.c.). For RG this can be seen from Bayer & Brandner (2004) where, systematically,

<sup>12</sup> Another example is an AN-based polar question from Mongolian (Koptjevskaja-Tamm 1993:286).

<sup>13</sup> I have crossed out Santali, given that the Santali verb does not seem to be (m-)asymmetric (Neukom 2001:61).

Bavarian indefinite *was* ('what') is rendered as *ebbes* ('something', cf. German *etwas*) in RG.<sup>14</sup>

I thus take (8) to be a generalization worth exploring further. The remainder of this paper will be devoted to a sketch of how to develop a theoretical account of (8) on the basis of German.

The core "functionalist" idea is that the development of embedded infinitival *wh*-questions is blocked in languages where these structures would be "hard to recognize." The German pattern would thus be an instance of the tendency toward the avoidance of structural ambiguities. The task of recognition can be divided in two parts, namely, formal and interpretive. I will have a look at each of them in turn.

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 pronouns into interrogatives.<sup>15</sup> Strategy 1 consists in using specialized pronouns. This is shown for French in (25).

- (25) 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 où?*  
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 (26).<sup>16</sup>

- (26) 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 (27).

- (27) a. *Zheli QUE-LE shenme*  
here is.missing something  
"Something is missing here"  
b. *Zheli que-le SHENME?*  
"What is missing here?"

---

<sup>14</sup> J. Fleischer (p.c.) informs me that this is generally true of the Alemannic (Swiss German) languages, i.e. they are [-i/i]-languages.

<sup>15</sup> I sidestep the use of Q-particles or verbal Q-mood (see Bhat 2000). Among the [+i/i]-languages the following have such marking strategies (cf. Bruening 2004).

(i) *Assurini, 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* [+wh.is]-languages. Among the languages considered in this paper, the ones in (ii) belong to that class (cf. Bruening 2004; Cheng 1991).

(ii) *Cantonese, Hindi, Hopi, Korean, Mandarin, Thai*

<sup>16</sup> According to Cheng (1991:108f) Martuthunira and Panyjima use this strategy for disambiguation, while Diyari fronted *wh*-pronouns are still ambiguous.

The same strategy is used to disambiguate *in-situ w*-indefinites in German multiple questions, as shown in (28).

- (28) a. *Wer hat was gelesen?*  
           who has what read  
           “Who read anything?”  
       b. *Wer hat WAS gelesen?*  
           “Who read what?”

Now, trivially, the safest strategy, i.e. strategy 1, is unavailable in [+i/i]-languages. German uses fronting for one *w*-indefinite and accentuation for further *w*-expressions in multiple questions. Crucially, fronted *w*-indefinites in interrogative clauses are not stressed unless some additional focusing is intended. They are thus not intrinsically identifiable as interrogative pronouns. But strategy 2 alone is not sufficient either, given the fact that German is an OV-language with scrambling. Thus (specific) *w*-indefinites easily end up on the left edge of an infinitive. This has already been shown in (6).

Let us turn to interpretive identification. It seems to be plausible to assume that questions 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 questions don’t seem to allow for such default forces. Let us consider two scenarios for the rise of embedded infinitival questions in German.<sup>17</sup> According to scenario 1 they would arise as complements to verbs allowing both infinitival and interrogative complements.<sup>18</sup> Sabel (1996:284) provides the following list of such verbs.

- (29) *vergessen* (‘forget’), *entscheiden* (‘decide’), *erklären* (‘state’),  
       *(sich) überlegen* (‘ponder’)

None of these denote default acts of seeking information. Thus their complements would not easily be recognizable as questions from interpretation alone.<sup>19</sup>

According to scenario 2, embedded infinitival questions would start out as embedded root phenomena, i.e. as root clauses that undergo some form of pseudo-embedding. 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 (2003:189f.) 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 (30a) but not by an infinitival one (30b).<sup>20</sup>

<sup>17</sup> The following sketch will have to be squared with the available historical evidence, something I leave for further research.

<sup>18</sup> Clearly, English *wh*-infinitives are allowed as complements to verbs that are semantically compatible with questions although formally incompatible with infinitives. *Wonder* is an obvious example here.

<sup>19</sup> English *wh*-infinitives have been shown to possess a peculiar modal component which reduces the environments they can be embedded into (Bhatt 1999). The following examples are taken from Reis (2003:173).

(i) a. \* John predicted who to invite  
       b. \* It is important who to talk to at the party  
       c. \* What to do depends on where to be

<sup>20</sup> Note that German only allows bare infinitives in this function. This distinguishes it from Dutch where similar deliberative questions are expressed by *to*-infinitives, as shown in (i) (Zwart 1993:102).

(i) *Wat \*(te) doen?*  
       what to do



- (30) a. *An wen soll ich mich wenden?*  
 at who shall I me turn  
 “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 identifying a structure as a question from interpretation alone.

To summarize, formal as well as interpretive aspects conspire against the use of *wh*-infinitives as embedded infinitival questions in German. Neither available strategies of clausal typing nor pragmatic assignment of illocutionary force allows a straightforward identification of the relevant structures as questions. Thus, if a language is a robust [+i/i]-language, then the likelihood that it lacks embedded infinitival (*wh*-)questions is high. How high exactly, will have to be established. For the time being I claim that  $\text{INT/IND-PRO} \Rightarrow \neg \text{WH-INF}$ .

## Appendix I:

Sabel (1996:296) assumes the following parametrization of infinitival  $\text{C}^\circ$ .

- (i)  $\text{C}^\circ_{\text{INF}}$  is [-/+ strong], if the infinitival C-system can / cannot be overtly filled with a base-generated element

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

- (ii) *Wh*-Criterion (Rizzi 1996:64)  
 a. Each  $\text{X}^\circ_{+\text{wh}}$  must be in a Spec-Head relation with a *Wh*-phrase  
 b. Each *Wh*-phrase must be in a Spec-Head relation with an  $\text{X}^\circ_{+\text{wh}}$

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

The initial stages of deriving (1a) in English would look like in (iii).

- (iii) a. [ $\text{C}^\circ$   $\text{C}^\circ_{\text{inf,wh,agr}}$  [ $\text{AGRP}$  PRO [ $\text{AGR}^\circ$   $\text{AGR}^\circ_{\text{agr}}$  [ $\text{TP}$  *to put the soap where* ] ] ] ]  
 b. [ $\text{C}^\circ$   $\text{C}^\circ_{\text{inf,wh,agr}}$  [ $\text{AGRP}$  PRO [ $\text{AGR}^\circ$   $\text{AGR}^\circ_{\text{agr}}$  [ $\text{TP}$  *to put the soap where* ] ] ] ]

Following Chomsky (2000) one can interpret “weak”  $\text{C}^\circ$  as involving a weak [*agr*]-feature. Thus, the transition from (iiia) to (iiib) would be brought about by pure “Agree.” [*wh*] in  $\text{C}^\circ$  is left unaffected and *wh*-movement can still apply.

In contrast, (iv) shows the German situation.

- (iv) a. [ $\text{C}^\circ$   $\text{C}^\circ_{\text{inf,wh,AGR}}$  [ $\text{AGRP}$  PRO [ $\text{AGR}^\circ$   $\text{AGR}^\circ_{\text{agr}}$  [ $\text{TP}$  *die Seife wohin zu legen* ] ] ] ]  
 b. [ $\text{C}^\circ$  [ $\text{C}^\circ$  [ $\text{AGR}^\circ_{\text{agr}}$  ]<sub>i</sub>  $\text{C}^\circ_{\text{inf,AGR}}$  ] [ $\text{AGRP}$  PRO [ $\text{AGR}^\circ$   $t_i$  [ $\text{TP}$  *die Seife wohin zu legen* ] ] ] ]

Strong [*AGR*] in  $\text{C}^\circ$  has to be eliminated by “Move,” which – by stipulation – overwrites [*wh*]. Thus, no subsequent *wh*-movement to  $\text{C}^\circ$  of an infinitive is possible and (1b) is ungrammatical.

## Appendix II<sup>21</sup>

	m-as	i	cv	an	w	o	i/i	r-i/i	qp/qm	wh.is
<i>Assurini</i>							+		+	
<i>Atayal</i>							+			
<i>Bagandji</i>							+			
<i>Belorussian</i>	+	+					+			
<i>Burushaski</i>	+		+				+			
<i>Chinese</i>	—						+			+
<i>Danish</i>	+	+			—	—	—	—		
<i>Diyari</i>							+			
<i>Djaru</i>							+			
<i>Dutch</i>	+	+			+	+	+	—		
<i>Dyirbal</i>							+			
<i>English</i>	+	+			+	+	—	—		
<i>French</i>	+	+			+	+	—	—		+
<i>Galibi</i>							+		+	
<i>German</i>	+	+			—	—	+	+		
<i>Reichenau German</i>	+	+			+	+	—	—		
<i>Pennsylvania German</i>	+	+			+	+	—	—		
<i>Goajiro</i>							+			
<i>Gooniyandi</i>							+			
<i>Gothic</i>	+	+					+			
<i>Classical Greek</i>	+	+			—		+			
<i>Guarani</i>							+		+	
<i>Hindi</i>	+	+			+		—	—		+
<i>Hmong Njua</i>	—						+			
<i>Hopi</i>							+		+	+
<i>Hungarian</i>	+	+			+	—	—	—		
<i>Italian</i>	+				+	+	—	—		
<i>Kaingang</i>							+			
<i>Kamaiurá</i>							+			
<i>Khmer</i>	—						+			
<i>Klamath</i>							+		+	
<i>Koasati</i>							+		+	
<i>Korean</i>	+		+				+		+	+
<i>Lakhota</i>	+			+			+		+	
<i>Latin</i>	+	+					+	—		
<i>Lithuanian</i>	+	+			—?		+		+	
<i>Mangarayi</i>							+			
<i>Mapuche</i>	+						+			
<i>Maricopa</i>	+			+			+		+	

<sup>21</sup> I have often preferred leaving blanks to filling in “trivial” minuses. Other blanks, of course, signal my current ignorance.

<i>Martuthunira</i>							+		–	
<i>Masalit</i>							+			
<i>Mising</i>							+		+	
<i>Mundari</i>			+				+			
<i>Newari</i>							+			
<i>Norwegian</i>	+	+			–	–	–			
<i>Panare</i>							+			
<i>Panyjima</i>							+		–	
<i>Pashto</i>							+			
<i>Polish</i>	+	+			+	+	–?	–		
<i>Eastern Pomo</i>							+			
<i>Portuguese</i>	+	+			+	+	–	–		
<i>Ancash Quechua</i>	+	+	+				+			
<i>Russian</i>	+	+					+	–		
<i>Vedic Sanskrit</i>	+	+	+		–?		+			
<i>Santali</i>	–?						+			
<i>Shoshone</i>	+	+					+			
<i>Siuslaw</i>							+			
<i>Old Church Slavonic</i>	+	+					+			
<i>Slovene</i>	+	+			+		+	–		
<i>Spanish</i>	+	+			+	+	–	–		
<i>Swedish</i>	+	+			–	–	–	–		
<i>Takelma</i>							+		+	
<i>Thai</i>	–						+		+	+
<i>Ukrainian</i>	+	+					+			
<i>Warndarang</i>							+			
<i>Wintu</i>							+		+	
<i>Xinh Mul</i>							+			
<i>Yaqui</i>							+			
<i>Yidini</i>							+			
<i>Yup'ik</i>							+		+	

**m-as:** *L* is minus-asymmetric in the sense of Bisang (2001)

**i:** *L* uses infinitives in subordinate clauses

**cv:** *L* uses converbs in subordinate clauses

**an:** *L* uses action nominals in subordinate clauses

**w:** *L* has embedded infinitival questions

**o:** *L* has base-generated elements in its infinitival C-system

**i/i:** *L* has an interrogative/indefinite ambiguity in its pronominal system

**r-i/i:** *L* has a robust interrogative/indefinite ambiguity in its pronominal system

**qp/qm:** *L* has question particles or verbal question mood

**wh.is:** *L* has (unmarked) *wh-in-situ* questions

## References

- Adelaar, Willem. 2004. "Quechua." Pp. 1453-1464 in *Morphology*, edited by Geert Booij. Berlin: de Gruyter.
- Bayer, Josef, and Ellen Brandner. 2004. "Klitisiertes *zu* im Bairischen und Alemannischen." Pp. 160-188 in *Morphologie und Syntax deutscher Dialekte und Historische Dialektologie des Deutschen*, edited by Franz Patocka and Peter Wiesinger. Wien: Edition Praesens.
- Bhat, D.N.S. 2000. "The Indefinite-Interrogative Puzzle." *Linguistic Typology* 4:365-400.
- Bhatt, Rajesh. 1999. "Covert Modality in Non-Finite Contexts." Ph.D. Dissertation, University of Pennsylvania.
- Bisang, Walter. 1995. "Verb Serialization and Converbs - Differences and Similarities." Pp. 137-188 in *Converbs in Cross-Linguistic Perspective*, edited by Martin Haspelmath and Ekkehard König. Berlin: Mouton de Gruyter.
- . 2001. "Finite vs. Non-finite Languages." Pp. 1400-1413 in *Language Typology and Language Universals*, edited by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible. Berlin: de Gruyter.
- Brandner, Ellen. 2000. "Scope Marking and Clausal Typing." Pp. 45-75 in *Wh-Scope Markin*, edited by Uli Lutz, Gereon Müller, and Arnim von Stechow. Amsterdam: John Benjamins.
- . 2004. "Microvariation within German *to*-Infinitives." Unpublished manuscript, University of Konstanz.
- Bruening, Benjamin. 2004. "Wh-Indefinites, Question Particles, and Wh-in-Situ: There is No Relation." Unpublished Manuscript, University of Delaware.
- Cheng, Lisa L.S. 1991. "On the Typology of *WH*-Questions." Ph.D. Dissertation, MIT.
- Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework." Pp. 89-155 in *Step By Step*, edited by Roger Martin, David Michaels, and Juan Uriagereka. Cambridge MA: MIT Press.
- Comrie, Bernard, and Sandra A. Thompson. 1985. "Lexical Nominalization." Pp. 349-398 in *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*, edited by Timothy Shopen. Cambridge: CUP.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.
- É. Kiss, Katalin. 2002. *The Syntax of Hungarian*. Cambridge: CUP.
- Gärtner, Hans-Martin, and Markus Steinbach. 1994. "Economy, Verb Second, and the SVO-SOV Distinction." *Working Papers in Scandinavian Syntax* 53:1-59.
- Gordon, Lynn. 1986. *Maricopa Morphology and Syntax*. Berkeley: University of California Press.
- Grosu, Alexander. 2004. "The Syntax and Semantics of Modal Existential *Wh* Constructions." Pp. 405-438 in *Balkan Syntax and Semantics*, edited by Olga Tomic. Amsterdam: John Benjamins.
- Haspelmath, Martin. 1995a. "Contextual and Specialized Converbs in Lezgian." Pp. 415-440 in *Converbs in Cross-Linguistic Perspective*, edited by Martin Haspelmath and Ekkehard König. Berlin: Mouton de Gruyter.
- . 1995b. "The Converb as a Cross-Linguistically Valid Category." Pp. 1-55 in *Converbs in Cross-Linguistic Perspective*, edited by Martin Haspelmath and Ekkehard König. Berlin: Mouton de Gruyter.
- . 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.
- Haspelmath, Martin, and Ekkehard König (Eds.). 1995. *Converbs in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.
- Herrity, Peter. 2000. *Slovene*. London: Routledge.
- Kayne, Richard. 1991. "Romance Clitics, Verb Movement, and PRO." *Linguistic Inquiry* 22.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.
- Marušič, Franc. 2005. "On Non-Simultaneous Phases." Ph.D. Dissertation, Stony Brook.

- Neukom, Lukas. 2001. *Santali*. Müncher: Lincom.
- Nuolijärvi, Pirkko. 1994. "Finiteness." Pp. 1245-1248 in *The Encyclopedia of Language and Linguistics*, edited by Ronald E. Asher. Oxford: Pergamon Press.
- Postma, Gertjan. 1994. "The Indefinite Reading of WH." Pp. 187-198 in *Linguistics in the Netherlands 1994*, edited by Reineke Bok-Bennema and Crit Cremers. Amsterdam: John Benjamins.
- Reis, Marga. 1991. "Was konstituiert w-Interrogativsätze." *Arbeitsberichte des SFB 340* 6:35-68.
- . 2003. "On the Form and Interpretation of German *Wh*-Infinitives." *Journal of Germanic Linguistics* 15:155-201.
- Rizzi, Luigi. 1996. "Residual Verb Second and the Wh-Criterion." Pp. 63-90 in *Parameters and Functional Heads*, edited by Adriana Belletti and Luigi Rizzi. Oxford: OUP.
- Sabel, Joachim. 1996. *Restrukturierung und Lokalität*. Berlin: Akademie-Verlag.
- Tikkanen, Bertil. 2001. "Converbs." Pp. 1112-1123 in *Language Typology and Language Universals*, edited by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible. Berlin: de Gruyter.
- Truckenbrodt, Hubert. 2004. "Zur Strukturbedeutung von Interrogativsätzen." *Linguistische Berichte* 199:313-350.
- Vyautas, Ambrazas (Ed.). 1997. *Lithuanian Grammar*: Baltos Lankos.
- Whitney, William D. 1950. *Sanskrit Grammar*. Cambridge MA: Harvard University Press.
- Zwart, Jan-Wouter. 1993. "Dutch Syntax. A Minimalist Approach." Ph.D. Dissertation, University of Groningen.