Vogel, R. (2001). Towards An Optimal Typology Of The Free Relative Construction. In A. Grosu (Ed.), Papers From The Sixteenth Annual Conference And From The Research Workshop Of The Israel Science Foundation "The Syntax And Semantics Of Relative Clause Constructions" (pp. 107-119). Tel Aviv: University Of Tel Aviv.

(please note that pagination follows a 2-on-1 format, i.e. the first two pages are page 1, etc.)

Towards an Optimal Typology of the Free Relative Construction*

Ralf Vogel
Institute of English Linguistics, University of Stuttgart
ralf.vogel@ims.uni-stuttgart.de

This paper is part of a larger research project on OT Syntax and the typology of the free relative (FR) construction. For a discussion of issues that are not covered in this paper see (Vogel 2001), (Vogel to appear*a*) and (Vogel to appear*b*). The present paper concentrates on a discussion of the typology to be accounted for and how it can be modeled within optimality theory. The first part of the paper presents the data in some detail and the second part discusses the structure of the OT approach that I propose. The typological discussion focuses on the variable treatment of *case conflicts* in FRs standing for a verbal argument.

An example of an English free relative (FR) clause is the subordinate *wh*-clause in (1), taken from Bresnan & Grimshaw (1978):

(1) [$_{CP}$ I drank [$_{FR}$ *whatever* there was]]

I assume that FR clauses have the structure of other ordinary subordinate clauses and that the label 'FR' is to be replaced by 'CP' in (1). Rooryck (1994) argues for this proposal in detail.[1]Many earlier accounts claim that there must be an NP node heading the FR clause. Bresnan & Grimshaw (1978) assume that this NP node hosts the FR pronoun. However, none of

*I would like to thank the following colleagues for helpful comments and fruitful discussions: Artemis Alexiadou, Gisbert Fanselow, Hans-Martin Gärtner, Anastasia Giannakidou, Jane Grimshaw, Alex Grosu, Fabian Heck, Gunnar Hrafn Hrafnbjargarson, Shin-Sook Kim, Jaklin Kornfilt, Geraldine Legendre, Gereon Müller, Peter Öhl, Henk van Riemsdijk, Doug Saddy, Tanja Schmid, Melita Stavrou, Ruth Seibert, Halldor Sigurðsson, Markus Steinbach, Sten Vikner. I'm further thankful to the audiences of presentations of parts of this paper at the Stuttgart 1999 OT Syntax Workshop, the Workshop on Conflicting Rules in Potsdam, December 1999, and IATL 16, June 2000, Tel Aviv. All remaining errors are mine. The work on this paper was fully supported by a grant for the DFG (German research society) research project "Optimalitätstheoretische Syntax des Deutschen" (MU 1444/2-1), University of Stuttgart.

[1]See also (Åfarli 1994) for a similar proposal. In fact, this is a revival of the first generative account of this construction by Kuroda (1968).

these proposals could convincingly prove the need for the stipulated exceptional structure of FR clauses by showing exceptional syntactic behaviour.

The only exceptional property of FRs is the so-called 'matching effect': the FR pronoun seems to be sensitive to the requirements of both the matrix verb and the relative clause internal verb.[2] As an effect of this, English only has well-formed FRs if the pronoun is able to 'fulfil' both requirements simultaneously. This is the case in (1). But note that the *requirements* of the verbs do *not* match here literally: the matrix verb requires a direct object, while the FR pronoun is the subject of the FR clause. But the *form* of the pronoun is the same for subject and direct object, so its form *matches* both requirements.

Rooryck (1994) notes that a more complicated structure is not necessary to allow for such effects. The accessibility of the [Spec,CP] position of the subordinate clause for the matrix verb is also assumed by a standard explanation for complementation by subordinate *wh*-clauses:

(2) [$_{CP}$ Mary asked [$_{CP}$ *what* Peter said]]

The verb 'ask' requires a *wh*-complement, but it is the [Spec,CP] position of the subordinate clause, and not the CP node itself, that fulfils the *wh*-requirement.[3]In FRs, the situation is not much different. Here it is a case requirement which is fulfilled by the element in the [Spec,CP] position of the subordinate clause.

In languages with a more elaborated case system a *case conflict* can occur on the FR pronoun between the case required by the matrix verb (henceforth `m-case`) and the case required by the FR-internal verb (henceforth `r-case`). The data to be accounted for are as in (3):

(3) German:

a. Wer einmal lügt, (der) lügt auch ein zweites Mal
who-NOM once lies, (he-NOM) lies also a second time
'People who lie(d) once, will lie a second time'

b. Wer einmal lügt, *(dem) glaubt man nicht
who-NOM once lies, *(he-DAT) believes one-NOM not
'One doesn't believe people who lie(d) once any more'

[2]It is a bit surprising to me that most researchers seem to agree without any doubt that the matching effect, first of all a morphological phenomenon, requires a treatment in terms of syntactic structure.

[3]The crucial data are possible extractions out of a complex *wh*-DP in [Spec,CP]. See (Chomsky 1986) for a discussion of these issues.

Both examples contain clause-initial FRs. In (3–b), the resumptive pronoun that picks up the referent of the FR within the main clause is obligatory, in (3–a) it is optional. I assume that the obligatoriness of the resumptive element in (3–b) is an effect of the case requirements of the matrix verb: the FR stands for a dative object, but the FR pronoun carries nominative case. Without the resumptive, dative would not be 'realised' morpho-phonologically. With Alexiadou & Varlokosta (1995) and Grosu (1994) I assume that such constructions with an overt resumptive pronoun in addition to the FR are instances of left dislocation of the FR. Correlative constructions often appear in the same configuration.

In (3–a) the FR pronoun matches the case required by the matrix verb – the resumptive can be omitted. The typology I am exploring here compares the situations under which it is possible to omit such a resumptive, i.e., those situations where the FR pronoun seems to serve two case assigners simultaneously. Only those kinds of structures are called 'FR constructions' in this paper.

Many languages allow for different varieties of non-matching FRs. German, for example, allows for the omission of nominative and accusative resumptives in non-matching configurations (*vertrauen* requires a dative and *einladen* an accusative object):

(4) a. Ich lade ein *wen/ wem ich vertraue
I-NOM invite *who-ACC/ who-DAT I trust

b. Mich läd ein *wer/ wem ich vertraue
Me-ACC invites *who-ACC/ who-DAT I trust

The details of the typology of this phenomenon will be presented in the next section.

1 The Typology of FRs

The syntax of FRs varies in 2 dimensions: first, languages differ in whether they have FRs at all, only matching FRs or also some or all kinds of non-matching FRs; second, languages differ in their resolution strategies for case conflicts. We observe three different ways of realising FRs: we find FRs with the FR pronoun realising `r-case` and `m-case` remaining unrealised; we find FRs with the FR pronoun realising `m-case` and `r-case` remaining unrealised; and we find FRs with the FR pronoun realising `m-case` and an additional FR-internal resumptive

pronominal element realising `r-case`. The fourth type of 'resolution' that also has to be considered in an OT account is the shift to another construction like, for instance, a correlative construction. This section introduces the language types that have to be considered.

1.1 Languages without FRs

There are languages that do not allow for free relatives. One example in case is Hindi (Dayal 1996). The usual way to translate a clause like 'I didn't like whatever Anu ordered' is by using a correlative construction:

(5) jo ciizeN anu-ne mangaayiiN **ve** mujh-ko nahiiN pasand aayiiN
which things Anu-Erg ordered **them** I-Dat not like come-P
'Which things Anu ordered, I didn't like them' (Dayal 1996, 213)

Another language that might belong to this class is Tok Pisin:

(6) Wanem ol kaikai ol i givim yu, yu no ken kaikai
what Pl. food they give you, you Neg can eat
'Whatever food they give you you must not eat' (Woolford 1978, 484)

Although Tok Pisin is classified as a language having FRs in the literature (cf. Bresnan & Grimshaw 1978 and Woolford 1978), the data can be interpreted in a different way. The reason is that here the FR looks exactly like a headed relative construction, cf.:

(7) Ol samting mipela salim i go long yu i kamap pinis long yu
Pl. thing we sent go to you come Aspect to you
'The things that we sent you arrived' (Woolford 1978, 485)

According to Woolford (1978), restrictive relative clauses in Tok Pisin look like ordinary clauses. They are not introduced by a complementiser or a relative pronoun. So we only have to replace the 'FR' pronoun, *wanem ol kaikai*, in (6) with an ordinary NP, *ol samting*, to yield a headed relative construction as in (7).

Bresnan & Grimshaw argue on the basis of the Tok Pisin data that FRs are structurally different from interrogative *wh*-clauses. Tok Pisin does not have *wh*-movement, i.e., *wh*-pronouns in interrogative clauses remain in situ:

(8) Yutupela sutim husat tru?
you shot who really
'Who did you really shoot?' (Woolford 1979, 43)

This is the same in subordinate *wh*-clauses. From Bresnan & Grimshaw's (1978) point of view, the interrogative pronoun remains in situ, while the FR pronoun occupies the clause-initial position – we have two different syntactic patterns. If Tok Pisin has no FRs, as I suggest above, then this argument breaks down, and the parallelism of FRs and other subordinate *wh*-clauses is re-established.

Korean, another *wh*-in situ language, uses a sort of correlative FR construction, very much like Hindi. The FR pronoun, an ordinary *wh*-element, remains in situ and a resumptive pronoun inside the matrix clause is obligatory:

(9) a. Nuku-ka na-lul ch'otaeha-tunchi *ku-nun* Hans-to ch'otaeha-n-ta
who-Nom ich-Acc invite-ever he-Top Hans-also invite-Pres-Dec
'Whoever invites me, also invites Hans.'

b. Maria-ka nuku-lul ch'otaeha-tunchi na-to *ku-lul* ch'otaeha-n-ta
Maria-Nom who-Acc invite-ever I-also he-Acc invite-Pres-Dec
'Whoever Maria invites, I will also invite him.'

c. Maria-ka nuku-eke ka-tunchi na-to *ku-eke* ka-n-ta
Maria-Nom who-Dat go-ever I-also he-Dat go-Pres-Dec
'To whoever Maria goes, I will go to him, too.'

If the above analysis of Tok Pisin is correct, then languages use two different syntactic alternatives to FRs: a correlative or left dislocation construction, as in Hindi and Korean, and/or a headed relative construction, as in Tok Pisin.

1.2 Languages with only matching FRs

In languages which only have matching FRs the surface form of the FR pronoun has to 'match' the correct forms required for the realisation of both `m-case` and `r-case`. English is such a language:

(10) (Bresnan & Grimshaw 1978):
a. I drank whatever there was
b. I'll reread whatever paper John has worked *(on)
c. *I'll reread on whatever paper John has worked

d. I'll live wherever you live
e. I'll live in whatever town you live (in)

If the matrix verb requires a NP/DP, then the FR pronoun has to be of that category, as we see in (10-a-c). The same holds for a PP requirement (10–e). However, English has preposition stranding. Although there is a conflict in (10–b) with respect to the forms required by the verbs – the matrix verb requires a direct object, i.e. a NP/DP, and the embedded verb a PP –, a FR is possible, if the pronoun moves on its own and strands the preposition (10–b). Pied-piping as in (10–c) yields ungrammaticality. This shows again that it is not the requirements of the two verbs that have to match, but it is the *element in the* [Spec,CP] *position* that has to match the matrix requirement, and on the other hand fulfil its requirements inside the embedded clause. One might argue that English only has this matching effect, because it has preposition stranding: (10–c) is odd because of the possibility of (10–b). Groos & van Riemsdijk (1981, 173) show that Dutch is also a matching language, but Dutch does not have English type preposition stranding.[4]

(11) a. *Ken jij met wie zij flirt?
know you with who she flirts?
'Do you know (the person) with who she is flirting?'

b. Ken jij wie zij net kuste?
know you who she just kissed?
'Do you know (the person) who she just kissed?'

Norwegian also seems to be a matching language, as reported by Åfarli (1994). German is classified as a matching language by Groos & van Riemsdijk (1981). Pittner (1991) and Vogel (2001) show that many German speakers do accept non-matching FRs. I also show in the latter paper that German speakers vary in which kinds of non-matching FRs they accept. I propose two variants, German A and German B. It cannot be excluded that there are speakers of German who only accept matching FRs as proposed by Groos & van Riemsdijk (1981). These speakers would then constitute a third variant, German C, which is, in this respect, like Dutch and English.

[4]Dutch only has preposition stranding in a very restricted way:

(i) Waar heb je ob gerekend?
where have you on counted?
'What have you counted on?'

This is the same in most variants of German, where preposition stranding is also possible only with the *r*-pronouns *wo* ('where') and *da* ('there'). See (Herslund 1984) and (Müller 2000) for further discussion.

1.3 Icelandic

Icelandic has an interesting and somewhat surprising pattern. The FR pronoun always bears `m-case`.[5]In addition, there do not seem to be any restrictions on the suppression of cases and Icelandic has preposition stranding. So it is hard to find a configuration that does not yield a well-formed FR.

In the following examples, two headed restrictive relative constructions (12–a,c) are paired with two FR constructions (12–b,d). The chosen verbs are *hjálpa* ('help'), which requires a dative object, and *elska* ('like'), which requires an accusative object. In German the same configuration would yield ungrammaticality for (12–b). This is not the case here. Icelandic FR pronouns always take `m-case` and `r-case` is simply suppressed. This is, however, not very surprising, if we look at restrictive relative clauses. They are uniformly introduced by the complementiser *sem*, and the relativised argument (which is represented by a relative pronoun in many other languages) remains unrealised, no matter what case it should have:[6]

(12) a. ég hjálpa þeim/ *þann sem ég elska
I help those-DAT/ those-ACC that I like

b. ?ég hjálpa hverjum/ *hvern (sem) ég elska
I help who-DAT/ who-ACC (that) I like

c. ég elska *þeim/ þann sem ég hjálpa
I like those-DAT/ those-ACC that I help

d. ?ég elska *hverjum/ hvern (sem) ég hjálpa
I like who-DAT/ who-ACC (that) I help

The classification of (12–b,d) as FRs and (12–a,c) as headed relatives is based on two observations: the complementiser can be omitted in the FRs with the *wh*-pronoun, and in (12–a,c) the d-pronoun can be separated from the relative clause:[7]

(13) a. þeim hjálpa ég sem ég elska
those-DAT help I that I like

[5]This phenomenon is called *case attraction* in the literature.

[6]The FRs in (12–b,d) are judged as 'archaic' or 'a bit strange' by my informants. But they agree that they are possible. The complementiser *sem* is optional here, contrary to restrictive relative clauses.

[7]To be honest, (13–b) is not ill-formed because the FR is disrupted. The pronoun 'hverjum' can be interpreted as interrogative. But now it is hard, if not impossible, to make sense of the clause.

b. *hverjum hjálpa ég (sem) ég elska
who-DAT help I (that) I like

1.4 German A

German has matching FRs and, in addition, non-matching FRs, if `m-case` is one of the structural cases nominative and accusative.

As already noted, we can identify at least two, if not three different variants of German with respect to FRs. In Vogel (2001) I discuss two variants I call German A and German B. These differ only in the treatment of one particular case conflict configuration, namely, if `m-case` is accusative and `r-case` is nominative:[8]

(14) a. *Er zerstörte, wer ihm begegnete
He destroyed who-NOM him-DAT met
'He destroyed who he met'

b. Er zerstörte was ihm begegnete
he destroyed what-NOM him-DAT met
'He destroyed what he met'

The given judgement is for German B. German A differs from German B in that here (14–a) is fine. One possible interpretation of these facts could be that in German B FRs are sensitive to a case hierarchy: only FRs that suppress the lower marked case are acceptable. German A could then be seen as a kind of mirror image of Icelandic in that it does not care about the case hierarchy and always realises the FR pronoun with `r-case`, suppressing `m-case`.

The difference to Icelandic is, however, that oblique forms, i.e., dative, genitive and PPs, cannot be suppressed at all in German. But the fact that in Icelandic relative clauses any case form can be suppressed is quite exceptional and surprising anyway.[9]

[8]The well-formedness of (14–b) is due to the fact that the FR pronoun *was* has the same form for accusative and nominative – because of this we find a matching configuration here, although the required/assigned cases are in conflict. Cf. the analogous case in English discussed above.

[9]An interesting difference between German and Icelandic is that Icelandic does not have what is called 'free dative' in German:

(i) a. Ich backte meiner Mutter einen Kuchen
I baked my mother-DAT a cake-ACC
'I baked my mother a cake'
b. *Èg bakaði mömmu minni köku
I baked mother my (a) cake

1.5 German B

As shown above, German B is a language that has matching FRs. It also has non-matching FRs, but only if the suppressed case is lower than the case realised on the FR pronoun on a hierarchical scale that Pittner (1991) proposes to look like this:

(15) nominative $\prec$ accusative $\prec$ dative, genitive, PP

The FR pronoun has to bear `r-case` in German. This means that FRs are impossible in German B, if `m-case` is higher than `r-case`, but possible in the opposite situation:

(16) a. `m-case`=ACC;`r-case`=NOM:

*Er zerstörte, wer ihm begegnete
He destroyed who-NOM him-DAT met
'He destroyed who met him'

b. `m-case`=NOM;`r-case`=ACC:

Ihm begegnete, wen er zerstören wollte
Him-DAT met who-ACC he destroy wanted
'Him met who he wanted to destroy'

Conflicts between accusative and nominative can only be resolved by a FR construction, if accusative is `r-case`. The same holds for a conflict between accusative and dative, now dative has to be `r-case`:

(17) a. `m-case`=DAT;`r-case`=ACC:

*Er begegnete, wen er zerstören wollte
He-NOM met who-ACC he destroy wanted
'He met who he wanted to destroy'

Likewise, free dative FRs are possible in German, but not in Icelandic.

(ii) a. Ich backe einen Kuchen wem ich vertraue
I bake a cake-ACC who-DAT I trust
'I bake a cake for whom I trust' (German)

b. *Ég bakaði köku hverjum/hvern ég elska
I bake a cake who-DAT/who-ACC I like
'I bake a cake for whom I like' (Icelandic)

I attribute this difference, as well as the one addressed above to the different status of case in the grammars of Icelandic and German, respectively. See (Vogel to appear*b*) for further discussion. See (Holmberg & Platzack 1995) for a related proposal about the status of case in Icelandic.

b. `m-case`=ACC;`r-case`=DAT:

Er zerstörte, wem er begegnete
He-NOM destroyed who-DAT met
'He destroyed who he met'

1.6 Gothic and Romanian

German and Icelandic are languages that uniformly realise either `m-case` (Icelandic) or `r-case` (German) on the FR pronoun, but cannot shift between the two. This is possible in Gothic and Romanian. In these languages, it is the 'higher' case that is realised on the FR pronoun.[10]

(18) Romanian, nominative vs. dative:

a. Cui i se dă de mîncare trebuie să muncească
who-DAT him self give of food must SUBJ work
'(He) who gets food must work' (Grosu 1994, 116)

b. Mă voi adresa cui mă poate înțelege
me will-I address who-DAT
'I shall turn to who can understand me' (Grosu 1994, 120)

In (18–a) the embedded verb requires dative on the pronoun, while the FR is the subject of the clause. In (18–b) the FR pronoun is subject of the FR clause, while the FR itself serves as dative object of the matrix clause. In both instances, the FR pronoun must bear dative case. The same behaviour can be observed in Gothic, as reported by Harbert (1983):

(19) Gothic, nominative vs. accusative (Harbert 1983, 248f):

a. jah *þo*-ei ist us Laudeikaion jus ussiggwaid
and Acc-Compl is from Laodicea you read
'and read (the one) which is from Laodicea' (Col 4:16)

b. *þan-ei* frijos siuks ist
Acc-Compl you-love sick is
'(The one) whom you love is sick' (Joh. 11:3)

In (19–a) the `m-case` is accusative and in (19–b) it is the `r-case`. Nevertheless, the FR pronoun bears accusative morphology in both instances. Accusative is, however, always suppressed, if it conflicts with higher marked dative or genitive:

[10]Romanian is a more complicated case than Gothic, because animate accusatives are usually realised with a preposition and this blocks suppression. See (Grosu 1994) for detailed discussion.

(20) Gothic, accusative vs. dative/genitive: (Harbert 1983, 248f)

a. hva nu wileiþ ei taujau *þamm-ei* qiþiþ þiudan Iudaie?
what now you-want that I-do DAT-Compl you-say king of-Jews
'What now do you want that I do to him (whom) you call the king of Jews?' (Mk 15:12)

b. bugei *þiz-ei* þaurbeima
buy GEN-Compl we-might-have-need-of
'Buy (that) of which we might have need' (Joh 13:29)

If one tried to reduce the hierarchy at work in Gothic to a two-element hierarchy of, say, 'marked' and 'unmarked' (which actually could not really be called a 'hierarchy'), one would have to decide whether accusative counts as marked or as unmarked. Depending on what this decision would be, it would either be wrongly predicted that accusative cannot lose against dative or genitive (because it is marked) or that it cannot win against nominative (because it is unmarked). Gothic and German B are languages that really use a scale, not only, e.g., a distinctive feature.

1.7 Modern Greek

Modern Greek shares with Icelandic that the FR pronoun bears `m-case`:

(21) Agapo opjon/*opjos me agapa
love-1Sg whoever-ACC/*NOM me loves
'I love whoever loves me' (Alexiadou & Varlokosta 1995, 12)

The FR pronoun shows obligatory *case attraction*. If the otherwise suppressed `r-case` is an oblique dative/genitive, then there has to occur a resumptive clitic realising `r-case`:

(22) Tha voithiso opjon tu dosis to onoma mu
FUT help-1S whoever-ACC cl-GEN give-2S the name my
*opjou 'whoever-GEN'
*s'opjon 'to whoever'
*opjou tu 'whoever-GEN him-GEN'
'I will help whoever you give him my name'
(Alexiadou & Varlokosta 1995, 13)

The conflict is resolved by realising both cases without giving up the FR structure. This only happens, when `m-case` is nominative or accusative and `r-case` is dative/genitive.

1.8 Summary

Table 1 gives a summary of the typology to be accounted for. There are languages without FRs and languages with only matching FRs. And then there is a language with an overall strategy, Icelandic. The other languages are obviously sensitive to the case hierarchy. German A and Modern Greek seem to make use of the case hierarchy in a different way than the others. They change their strategy, when their standard mode of conflict resolution would yield suppression of oblique case. German A has no FRs in this situation, while Modern Greek uses the resumptive pronoun strategy. Gothic and German A also take care of accusative-nominative conflicts. As shown above, many of these typological patterns are observed in more than one language.

Conflict[11]	Hindi	Engl.	Icel.	Ger.A	Ger.B	Gothic	M. Greek
m=NOM;r=ACC	–		M	R	R	R	M
m=NOM;r=OBL	–	–	M	R	R	R	RES
m=ACC;r=OBL	–	–	M	R	R	R	RES
m=ACC;r=NOM	–		M	R	–	M	M
m=OBL;r=NOM	–	–	M	–	–	M	M
m=OBL;r=ACC	–	–	M	–	–	M	M
m=r	–	FR	FR	FR	FR	FR	FR

Table 1: Typology of case conflict resolution in FRs

There is no need to assume that table 1 is complete. On the other hand, the data suggest a certain systematicity. The conflicts are sorted into two groups: in the first three conflict types `r-case` is higher than `m-case`, the next three types have the opposite pattern. Only two languages do not seem to have a uniform strategy for the same conflict type. But this might be an artefact of the mode of presentation. If German A and Modern Greek only distinguish between structural and oblique case and judge the two structural cases nominative and accusative as equivalent, although they are morphologically distinct, then the pattern is quite uniform again: nominative and accusative would not conflict, the fields in the first and fourth lines of

[11]Only those forms of nominative and accusative are taken into account that differ, so English has no conflicts between nominative and accusative forms, because these forms match. The abbreviations M, R and RES stand for the three different types of FRs: those with the pronoun realising `r-case` (R) and `m-case` (M), and those that use the resumptive pronoun strategy (RES). These abbreviations will be used throughout the paper.

the columns of German A and Modern Greek should be left blank.

Four of the seven language types are sensitive to the case hierarchy, while they differ in whether they use a twofold hierarchy (structural vs. oblique, as in German A and Modern Greek), or a threefold hierarchy (nominative, accusative, oblique).

2 An OT account

The preference of a given language for one or more of the four case conflict resolution strategies discussed in the preceding section result from evaluations of their advantages and disadvantages. Neither of the strategies is a 'perfect' solution. Which strategy is chosen by a language depends on the relative importance of the strategies' advantages and disadvantages. This is the optimality theoretic view of the problem. Conflicting requirements and constraint violations are assumed to occur everywhere.

FRs have often been proposed to have a syntactic structure that is more complex than that of a usual subordinate clause. An important motivation for this treatment seems to me to lie in the need to avoid a configuration that would yield constraint violation. The most popular analysis is the one developed by Groos & van Riemsdijk (1981). They claim that FRs are relative clauses with a phonetically empty NP head, usually represented as `pro`:

(23) $[_{NP}\ [_{N}$ `pro` $]\ [_{CP} \ldots$ FR $\ldots\ ]]$

This configuration avoids the case conflict by assuming that `m-case` and `r-case` are assigned to two different NPs; `m-case` is assgned to `pro` and `r-case` is assigned to the FR pronoun. Problematic cases are instances of case attraction, i.e., those cases where the FR pronoun has `m-case` morphology, although it is only assigned `r-case`, according to the structure in (23)– a very obvious constraint violation, without even having a conflict. Harbert (1983) tries to attack this problem by assuming that case assignment can apply at the level of Phonetic Form (PF), which is very hard to believe to be the case. My impression is that most attempts that try to avoid a conflicting or ambiguous syntactic representation end up in making strange additional assumptions. I conclude that the "mistake" that these proposals made is trying to avoid to represent what is obviously the case, namely, that the FR pronoun somehow is 'assigned' case twice.

This treatment raises at least as many problems as it solves.

I thus assume, with Rooryck (1994), that FRs are ordinary CPs with the FR pronoun in [Spec,CP]. Furthermore, I assume that abstract case is assigned to XPs in case positions, hence the FR-CP is assigned abstract `m-case` and the FR pronoun is assigned abstract `r-case`. I finally assume, again with Rooryck (1994), that the C^0 head of the FR-CP has the propoerties of an agreement head. Thus, the C-AGR head and the FR pronoun are in a Spec-head relation that is also a suitable configuration for case 'assignment'.

I share Harbert's (1983) idea that case attraction is a surface phenomenon. My proposal is that the case conflict arises not with abstract case, but with 'surface' case morphology – as we saw in the discussion in the first section, the matching effect is basically about surface forms rather than about abstract case. While the abstract cases are assigned to two different XPs, there is only one element that can morpho-phonologically realise the two cases, the FR pronoun. It can realise `r-case`, because it is assigned this case via ordinary case assignment, and it can realise `m-case`, because it stands in a Spec-head agreement relation with the C-AGR head of the FR. Table 2 gives a summary of this idea.

	`m-case`	`r-case`
abstract case assigned to	FR-CP	FR pronoun
surface case to realise on	FR pronoun (Spec-head-agreement with C-AGR)	FR pronoun (case assignment)

Table 2: Case Conflict Configuration

One way of interpreting this might be that the whole problem occurs, because the languages under consideration do not have the possibility to realise case morphology on subordinate clauses, and that the FR pronoun is a kind of a last resort for the realisation of `m-case`.

With this general picture in mind, I will now introduce the constraints I am assuming.

2.1 The Constraints

The first constraint is responsible for the ban on FR constructions as such in some languages, e.g., Hindi, Korean, Tok Pisin.[12] FR-CPs without overt case markers violate the constraint in (24) that requires a one-to-one correspondence between abstract and surface case markers:[13]

(24) **Case Uniqueness (CU)** If XP has abstract case at LF, then the PF-correspondent of XP has case morphology.

If CU is ranked high enough, then FRs are banned. In such a case an alternative structure 'wins' the OT competition. This is a correlative or a headed relative construction, as discussed in section 1.2.These 'deviant' constructions are included in the candidate set of the OT competition.[14] I assume that the input for our competition is a full-fledged syntactic structure, i.e., that of a FR construction. Correlatives differ from the input and violate a constraint on input-output faithfulness:[15]

(25) **Input-output faithfulness (FAITH-IO)** The input is preserved in the output.

The relative ranking of these two constraints determines the possibility of FRs:

(26) CU $\gg$ FAITH-IO = No FRs, only CORR
FAITH-IO $\gg$ CU = FRs possible

Those languages that allow for FRs can use strategy M (`m-case` on the FR pronoun), R (`r-case` on the FR pronoun) and RES (`m-case` on the FR pronoun and an additional resumptive element with `r-case`). Let us first compare M and R. The advantage of strategy M is that it is made explicit how the FR is 'integrated' into its superordinate clause – it is clear which thematic role it is assigned, for instance. The constraint that I am thinking of favors features of the matrix clause over features of a subordinate clause. An observation that can be made over and over again is that it is much easier to have 'gaps', e.g., omitted or deleted constituents in subordinate clauses than in matrix clauses. The constraint in (27) expresses this tendency. It might in itself

[12]The description that I am giving here is very informal and intuitive. A more precise and formal discussion of the proposed model is given in (Vogel to appear*a*).

[13]This constraint is a simplified version of the one that is used in (Vogel to appear*a*). It is neutral about whether the case morphology expresses the correct abstract case. It is only violated by categories that carry abstract case and cannot have case morpholgy, like subordinate clauses in the languages under discussion.

[14]I use the abbreviation 'CORR' for this candidate in the discussion below.

[15]See McCarthy & Prince (1995) for a detailed discussion of faithfulness in OT.

be an effect of the collaboration of other constraints, but I do not want to study this here. So this constraint can be read as an abbreviation for a yet to be determined system of constraints.

(27) **Matrix integration (MI)** This constraint is violated by constituents that contain no indication about how they are integrated into their clause.

MI is only violated by candidate R in non-matching competitions. Candidates M and RES fulfil MI by realising `m-case` on the FR pronoun. But this in turn leads to another 'imperfection': the FR pronoun is assigned abstract `r-case`. So the abstract case feature of the FR pronoun and its surface case morphology do not match. This is evaluated by the following constraint:[16]

(28) **One chain – one Case (1C1C)** The surface case morphology of an XP has to match its abstract case features.

This constraint is violated by strategy M, and also by strategy RES, but not by strategy R. The ranking of MI and 1C1C with respect to each other mirrors the preference for one or the other strategy.

(29) MI $\gg$ 1C1C = M, RES preferred
1C1C $\gg$ MI = R preferred

In additiion to 1C1C, candidate RES has another imperfection, namely, the occurrence of a resumptive pronoun. I assume, with Pesetsky (1998), that this resumptive pronoun inside the FR spells out the trace of the FR pronoun. It violates the following constraint against spelling out traces that has been proposed by Pesetsky (1998):

(30) **Silent trace (ST)** Don't pronounce the traces of a moved constituent

If ST is ranked high enough, in particular, higher than MI and at leats as high as 1C1C, then candidate RES is banned.

(31) ST MI $\gg$ 1C1C = M preferred

The possible rankings of this system of constraints give us three different grammar types, with CORR, M and R as an 'overall strategy'. It is not yet possible for strategy RES to win over M. Its advantage has not yet been integrated into the system of constraints. But what we have up

[16]The constraint that I am assuming in (Vogel to appear*a*) is more general and speaks of correspondence between abstract and surface morphology in a broader sense.

to now cannot be the full picture anyway. We saw that there are languages that allow only for matching FRs and some that allow for non-matching FRs that are in accordance with a case hierarchy. To capture these, and the advantage of RES, I assume two additional constraints that are weaker versions of UC. The first constraint requires abstract cases to be realised overtly, but allows for one element to realise two cases simultaneously, as is the case in matching FRs.

(32) **Realise Case (RC)** Each abstract case feature X at LF corresponds with a surface case marker realising X at PF.

This constraint is fulfilled in all matching FRs, and by RES FRs even in a case conflict configuration. But it is not fulfilled by the strategies M and R in non-matching FRs, because here only one of the two cases is overtly realised.

A language with only matching FRs (and without FR-internal resumptive pronoun), like English, has the following partial ranking:

(33) RC ST $\gg$ FAITH-IO $\gg$ UC

The RES candidate is preferred under a partial ranking that ranks FAITH-IO high, in order to ban CORR, and RC, in order to ban M and R in non-matching competitions. The constraints violated by RES (ST, 1C1C, UC) have to be ranked low:

(34) FAITH-IO RC $\gg$ ST 1C1C UC

To include the case hierarchy, I use a more liberal version of RC that assumes an abstract case to also be realised by the surface form of a hierarchically higher case:

(35) **Realise Case (relativised) (RCr)** Each abstract case feature X at LF corresponds with a surface case marker realising X at PF or a surface case marker of a more marked case form, as determined by the language particular case hierarchy.

The case hierarchies that are at issue are hierarchies of *surface case forms*.[17] Languages are only sensitive to the case hierarchy, if RCr is ranked high. We saw that not all languages we examined are sensitive to a case hierarchy, and that some languages seem to have only a two-membered hierarchy, others a three-membered one. Abstract case might be universal, but case forms are

[17]This can be shown in a comparison of those languages that are sensitive to case hierarchies. For a detailed discussion of this issue see (Vogel to appear*b*, Vogel to appear*a*).

language particular, and these are crucial here, as we saw above. Thus, the case hierarchies we are talking about are language particular, and for this reason I make no attempt to encode them directly into a set of universal markedness constraints with a fixed ranking (like, e.g., 'Realise Dative' ≫ 'Realise accusative' etc.), as done, e.g., by Woolford (2000) to account for the typology of case systems. The constraint RCr compares the case-hierarchical ranking of the given form of a FR pronoun in a candidate with the ranking of the form that the suppressed case would yield – if the given form is not ranked higher than or identical to the suppressed form, then RCr is violated.

German B allows for non-matching FRs only if they obey the case hierarchy. This is only minimally different from matching languages like English. In the partial ranking for German B we would only have to take the partial ranking of English, as given in (33) and replace RC with RCr. In addition, we have to block the M candidate and the RES candidate, so 1C1C is ranked high instead of or in addition to ST:

(36) 1C1C RCr (ST) ≫ FAITH-IO ≫ UC (RC)

The first 'stratum' of constraints can only be survived by CORR and R (if it obeys the case hierarchy in a case conflict configuration) or R/M (in a matching configuration). FAITH-IO, the constraint in the second stratum, blocks CORR, but only if it still has a competitor. Hence, German B has no FRs except for matching ones and non-matching ones of the R type that obey the case hierarchy.

An overview of the constraint violations and competitions is given in table 3. The next section discusses some more details of the predicted typology.

3 The Typology

The typology of the system developed in the last section is much more restricted than one might imagine. Three different situations have to be taken into account: `m-case` and `r-case` are identical, `m-case` is 'higher' than `r-case` and `m-case` is 'lower' than `r-case`. The latter two situations have four possible outcomes, M, R, RES and CORR, while the first situation has only three, because the outcomes M and R are indistinguishable. This means that there are 3 ×

	FAITH-IO	ST	UC	1C1C	RC	RCr	MI
m-case $\prec$ r-case							
R			1		1		1
M			1	1	1	1	
RES		1	1	1			
CORR	1						
m-case $\succ$ r-case							
R			1		1	1	1
M			1	1	1		
RES		1	1	1			
CORR	1						
m-case $=$ r-case							
R/M			1				
RES		1	1				
CORR	1						

Table 3: Summary of constraints, competitions and violations

$4 \times 4 = 48$ logically possible different patterns that an individual grammar can produce.

With the seven constraints that we are using we get a total number of $7! = 5040$ possible rankings. However, these different rankings only yield 11 different patterns. 37 logically possible patterns are predicted never to occur.[18]

Eight of these eleven patterns can be related to one of the languages discussed in section 1. Each of these languages is accounted for by one of the predicted grammars. The rankings are given in (37). Table 4 shows the winners for each of the three competition types.

(37) Hindi: UC (ST) (1C1C) (RC) (RCr) (MI) $\gg$ FAITH-IO
Icelandic: FAITH-IO ST MI $\gg$ 1C1C RC RCr UC
German A: FAITH-IO ST 1C1C $\gg$ UC RC RCr MI
German B: ST 1C1C RCr $\gg$ FAITH-IO $\gg$ RC UC MI
Gothic: FAITH-IO ST RCr $\gg$ 1C1C RC UC MI
Modern Greek: FAITH-IO RCr MI $\gg$ ST (1C1C) (UC) $\gg$ RC
English: ST RC (1C1C) (RCr) (MI) $\gg$ FAITH-IO $\gg$ UC

Only the crucial rankings are indicated. Constraints that occur in brackets are ranked as high as

[18]The typology has been calculated with the assistance of Bruce Hayes's constraint ranking software 'OTSOFT' (Hayes 1998).

Conflict	Hindi	Engl.	Icel.	Ger.A	Ger.B	Gothic	M. Greek
`m-case` $\prec$ `r-case`	CORR	CORR	M	R	R	R	RES
`m-case` $\succ$ `r-case`	CORR	CORR	M	R	CORR	M	M
`m-case=r-case`	CORR	R/M	R/M	R/M	R/M	R/M	R/M

Table 4: Predicted typology of case conflict resolution in FRs

possible, but could also be ranked lower. Their ranking is not crucial.

The pattern for Hindi and other non-FR languages is straightforward. FAITH-IO is the only constraint that is violated by candidate CORR. It must be ranked lower than UC or another combination of constraints such that each of the FR candidates violates a constraint that is ranked higher than FAITH-IO.

English allows for FRs, so FAITH-IO is ranked above UC. But RC is ranked above FAITH-IO, which blocks M and R in non-matching FRs. ST is also ranked high, which blocks candidate RES. So only the M/R candidate of a matching competition is a winning, i.e., well-formed FR.

Icelandic has an overall preference for M. This results from ranking constraints high that are violated by the other candidates, but not by M. FAITH-IO is violated by CORR, ST by RES and MI by R, none of them is violated by M.

The ranking attributed to German A also results in an overall strategy. Here, R is the winner. The difference to Icelandic is that 1C1C, which is violated by M, is ranked high instead of MI. This ranking correctly predicts that a FR is well-formed, if `m-case` is accusative and `r-case` is nominative – i.e., German A does not obey the case hierarchy. But it is now also predicted that dative can be suppressed in favour of a less marked nominative or accusative. This is a configuration that does not yield a well-formed FR. The explanation that I gave in (Vogel 2001) was that such a candidate is indeed syntactically well-formed, but that it is semantically uninterpretable and thus ‘crashes’ in the semantics component of the grammar. Oblique case may not be suppressed in German, because it makes a semantic contribution. Nevertheless, such a clause might be syntactically optimal.

The ranking for German B is slightly different. FAITH-IO is in an intermediate position. ST and 1C1C are ranked high to exclude M and RES, and RCr is ranked high to exclude those R candidates that do not obey the case hierarchy. The ranking for German B could also be used

for German A, if we assumed that the German A case hierarchy makes no difference between nominative and accusative. The difference between the two dialects would then not be attributed to different rankings, but to different case hierarchies.

The Gothic pattern is again quite straightforward. FAITH-IO and ST are ranked high, so RES and CORR are blocked and M and R remain. We know that both of them can win in Gothic, the crucial factor is the case hierarchy. So RCr is also ranked high. Depending on the type of the case conflict, RCr is either violated by M and fulfilled by R or vice versa.

The ranking proposed for Modern Greek has a problem that is similar to the one that ocurred with German A. Strategy RES is proposed to win, if `m-case` is lower than `r-case`. But we only observe RES, if `r-case` is dative/genitive, not if it is accusative. As for German A, we can claim that the Modern Greek case hierarchy treats nominative and accusative as equivalent. This would mean that strategy M is optimal when `m-case` is nominative and `r-case` is accusative. Candidate R is blocked because of the high ranking of MI. RCr is not violated by M, if `r-case` is accusative. So the high ranking of RCr is no longer crucial for this conflict. FAITH-IO is also ranked high, so CORR is blocked.

The three predicted, but not attested languages have the rankings in (38). The outcomes are summarised in table 5.

(38) Unattested # 1: FAITH-IO RCr $\gg$ 1C1C (UC) $\gg$ RC (MI) $\gg$ ST
Unattested # 2: FAITH-IO RC (RCr) (MI) $\gg$ ST UC 1C1C
Unattested # 3: ST RCr MI $\gg$ FAITH-IO $\gg$ RC UC 1C1C

Conflict	Unattested # 1	Unattested # 2	Unattested # 3
`m-case` $\prec$ `r-case`	R	RES	CORR
`m-case` $\succ$ `r-case`	RES	RES	M
`m-case`=`r-case`	R/M	R/M	R/M

Table 5: Predicted, but not yet attested patterns of case conflict resolution in FRs

None of these grammars is unreasonable. The first language is a mirror image of Modern Greek in that its default strategy is realising `r-case` on the pronoun. It shifts to the resumptive pronoun strategy, when `m-case` is higher than `r-case`. A mirror image of German B is the

third language. Its default is realising the FR pronoun with `m-case`. If `r-case` is the higher case, a FR is impossible.

The second language uses resumptives for both non-matching FR types. This is also a reasonable strategy. Future research will show, whether the three predicted, but unattested languages exist, and whether there exist other languages that are wrongly predicted not to exist by this account.

An interesting gap in the typology is worth mentioning: the system predicts that strategy RES is unavailable if `m-case` and `r-case` are identical, i.e., in a matching FR. We can see that from the constraint violations for this competition in table 3. The candidates R/M and RES both violate UC, but RES violates ST in addition. There is no ranking under which RES can win against R/M in this competition. RES is 'harmonically bounded'. And indeed, a language with matching FRs of the RES type has not been attested yet.

4 Summary

The central hypothesis underlying the present study is that the typology of the FR construction is crucially determined by the role that case plays in the grammar of a language. I assume that in the languages at issue FRs are intrinsically 'imperfect' insofar as they cannot realise the case they are assigned by themselves. There are several strategies that a language can choose to 'repair' this: shift to a more complex construction like a correlative construction, spell out a trace, use case attraction, suppress the 'less important case' etc. Each of these strategies has different imperfections. The relative importance of these imperfections determines the choice of the strategy in a language.

The picture drawn here is nevertheless incomplete. It has not been studied how prepositional FRs and other FRs that involve pied-piping fit into it. FRs are not restricted to NPs and PPs. We also find adjectival and perhaps verbal FRs, adjunct and predicate FRs etc. Their typology has not been studied within the present framework.

Although the analysis focuses on a single construction, it entails some more general claims. If a language disfavours resumptive pronouns in FRs, it might do so as well in other environments. If it allows for case attraction here, it will do so at other occasions as well. If the case

hierarchy plays a role in FRs, it might do so elsewhere. The used constraint set in itself is not construction specific.

OT syntax analyses are sometimes suspected to be more complex than the problem they try to solve. One hears comments like "… no surprise that you predict eight languages with seven constraints." I want to emphasise once more that it is not at all that simple. We are not dealing with a 'one-clause' typology here. Instead, we have to predict and evaluate different *patterns* of behaviour. I considered three different cases for each language, and this already increases the potential typology drastically. The predictive power of the proposed system can easily be calculated: a total of 48 logically possible patterns is reduced to eleven – 37 patterns are predicted to be impossible. This is done by the assumed constraint set. This would not be possible, if the constraints were language particular or arbitrarily chosen. All languages that I examined are covered by the given typology, and there is only a very small set of predicted, but not yet attested languages.

The value of a new proposal is determined in relation to its predecessors. Previous accounts of the typology of FRs distinguish only, e.g., between matching and non-matching languages (Groos & van Riemsdijk 1981, Bresnan & Grimshaw 1978), some assume in addition a group of 'partially matching languages' (Grosu 1994). The diversity of possible FR construction types and their different distribution in different languages have not been studied in such a detail in earlier work. Case attraction has always been a notorious problem. One reason for the difficulties that these accounts face is that FR constructions in themselves are somewhat special, 'marked' or 'imperfect'. Traditional grammar models assume well-formed constructions to be perfect constructions. OT offers a different point of view of which I hope to have shown that it is useful for the analysis of FRs, precisely because it presupposes imperfection.

References

Åfarli, Tor A. (1994): 'A promotion analysis of restrictive relative clauses', *The Linguistic Review* **11**, 81–100.

Alexiadou, Artemis & Spyridoula Varlokosta (1995): 'The Syntactic and Semantic properties of Free Relatives in Modern Greek', *ZAS Working Papers in Linguistics* **5**, 1–30.

Bresnan, Joan & Jane Grimshaw (1978): 'The Syntax of Free Relatives in English', *Linguistic Inquiry* **9**, 331–391.

Chomsky, Noam (1986): *Barriers*. MIT Press, Cambridge/Massachussetts.

Dayal, Veneeta (1996): *Locality in WH Quantification*. Kluwer, Dordrecht.

Groos, Anneke & Henk van Riemsdijk (1981): Matching Effects with Free Relatives: a Parameter of Core Grammar. *In:* A. Belletti, L. Brandi & L. Rizzi, eds, *Theories of Markedness in Generative Grammar*. Scuola Normale Superiore di Pisa, Pisa, pp. 171–216.

Grosu, Alexander (1994): *Three Studies in Locality and Case*. Routledge, London/New York.

Harbert, Wayne (1983): 'On the Nature of the Matching Parameter', *The Linguistic Review* **2**, 237–284.

Hayes, Bruce (1998): 'OTSOFT.EXE. Optimality Theory Software', http://www.humnet.ucla.edu/humnet/linguistics/people/hayes/otsoft/otsoft.htm.

Herslund, Michael (1984): 'Particles, Prefixes, and Preposition Stranding', *Topics in Danish Syntax* **14**, 34–71.

Holmberg, Anders & Christer Platzack (1995): *The Role of Inflection in Scandinavian Syntax*. Oxford Studies in Comparative Syntax, Oxford University Press, New York/Oxford.

Kuroda (1968): 'English Relativization and Certain Related Problems', *Language* **44**, 244–266.

McCarthy, John & Alan Prince (1995): Faithfulness and Reduplicative Identity. *In:* J. Beckman, L. Walsh-Dickie & S. Urbanczyk, eds, *Papers in Optimality Theory*. Vol. 18, UMass Occasional Papers in Linguistics, Amherst, Massachussetts, pp. 249–384.

Müller, Gereon (2000): 'Das Pronominaladverb als Reparaturphänomen', *Linguistische Berichte* **182**, 139–178.

Pesetsky, David (1998): Some Optimality Principles of Sentence Pronunciation. *In:* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds, *Is the Best Good Enough?*. MIT Press, Cambridge, Massachusetts, pp. 337–384.

Pittner, Karin (1991): Freie Relativsätze und die Kasushierarchie. *In:* E. Feldbusch, ed., *Neue Fragen der Linguistik*. pp. 341–347.

Rooryck, Johan (1994): 'Generalized Transformations and the WH-Cycle: Free Relatives and bare Wh-CPs', *GAGL* **37**, 195–208.

Vogel, Ralf (2001): Case Conflict in German Free Relative Constructions. An Optimality Theoretic Treatment. *In:* G. Müller & W. Sternefeld, eds, *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 341–375.

Vogel, Ralf (to appear*a*): Free Relative Constructions in OT Syntax. *In:* G. Fanselow & C. Féry, eds, *Sonderheft Optimality Theory*. Linguistische Berichte.

Vogel, Ralf (to appear*b*): Surface Matters. Case Conflict in Free Relative Constructions and Case Theory. *In:* E. Brandner & H. Zinsmeister, eds, *New Perspectives on Case Theory*. CSLI Publications.

Woolford, Ellen (1978): Free Relatives and Other Base Generated WH Constructions. In: *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*. University of Chicago, Chicago, Illinois, pp. 482–490.

Woolford, Ellen (2000): Case Patterns. *In:* J. Grimshaw, G. Legendre & S. Vikner, eds, *Optimality Theoretic Syntax*. MIT Press, Cambridge/Mass.

Woolford, Ellen B. (1979): *Aspects of Tok Pisin Grammar*. Number 66 *in* 'Pacific Linguistics, Series B', Department of Linguistics, Research School of Pacific Studies, The Australian National University, Canberra.

To appear in: Proceedings of IATL 16/2000