Agreement: Interactions with Case and Raising

Alex Alsina
Universitat Pompeu Fabra

Eugenio M. Vigo
Universitat Pompeu Fabra

Proceedings of the LFG'17 Conference

University of Konstanz

Miriam Butt, Tracy Holloway King (Editors)

2017

CSLI Publications

pages 3–23

http://csli-publications.stanford.edu/LFG/2017

Keywords: case, raising, agreement, OT

Alsina, Alex, & Vigo, Eugenio M. (2017). Agreement: Interactions with Case and Raising. In Butt, Miriam, & King, Tracy Holloway (Eds.): *Proceedings of the LFG'17 Conference, University of Konstanz* (pp. 3–23). Stanford, CA: CSLI Publications.

Abstract

The goal of this paper is to describe verbal agreement in languages like Icelandic in which the finite verb agrees with the nominative SUBJ, if there is one; otherwise, it agrees with a nominative OBJ; otherwise, it shows 3rd person singular features in the default gender (neuter in Icelandic). Special attention is paid to agreement in raising constructions, the raising verb may agree with the nominative OBJ of its infinitival complement. Similar facts occur in English locative inversion (Bresnan, 1994). These facts support the claim that verbs do not specify the Person-Number-Gender (PNG) features of any particular GF in their lexical entries. Instead, they specify the clausal PNG features as the feature structure AGR(EEMENT), which is unified with the AGR of the appropriate GF satisfying a set of OT constraints (as in Alsina & Vigo, 2014).

1 Introduction

Whereas verb agreement in many languages can be described as an agreement relation in which the agreement trigger is always the subject, in other languages the trigger of verb agreement cannot be defined as the subject or any other specific grammatical function (GF), as the agreement trigger varies from one GF to another, if any, depending on different properties of the clause. An example of such a language is Icelandic, whose behavior with respect to verb agreement can be illustrated as follows (the agreement trigger is shown here in boldface):

(1) a. **Við** hjálpuðum stelpunum
we.nom *helped*.1.pl *girl*.dat.f.pl
'We helped the girls' (Sigurðsson, 2004)

b. Henni líkuðu **hestarnir**
she.dat *liked*.3.pl *horse*.nom.m.pl
'She liked the horses' (Sigurðsson, 2004)

c. Mér býður við setningafræði
I.dat *nauseated*.med.3.sg *by syntax*
'I am nauseated by syntax' (Zaenen et al., 1990)

The generalization that covers the agreement facts illustrated here is as follows (see Andrews, 1990; Sigurðsson, 1996, 2004; Zaenen et al., 1990, among others):

(2) **The Icelandic agreement facts**: The finite verb agrees with the nominative GF that is highest in the subject > non-subject hierarchy; if there is no nominative GF, the verb is in the third person singular form.

[†]The research presented in this paper is supported by research project *Highest Argument Agreement* (HAA), FFI2014-56735-P (Spanish Ministry of Economy and Competitiveness).

In (1a), the verb agrees with the subject, which is the only nominative GF in the clause; in (1b), the verb cannot agree with the subject, as it is in the dative case, but agrees instead with the object, which is nominative; in (1c), the verb does not agree with any GF, as the clause includes no nominative GF, and so is in its third person singular form.

These facts pose a problem for the standard LFG approach to agreement, according to which the agreement target, such as the verb in verb agreement, lexically specifies the GF that it agrees with, along with the features of person, number, gender, etc. of this GF. So, for example, the Latin verb form *amamus* 'we love' is claimed to lexically specify that its subject has the features of $\left[\text{PERS} \quad 1\right]$ and $\left[\text{NUM} \quad \text{PL}\right]$. And it is not only this verb form, but all finite verb forms in Latin that impose featural requirements on their subjects, and on no other GF. But what we see in Icelandic is that finite verb forms cannot impose featural requirements on a particular GF, because the morphological form of a verb may depend on the subject, on the object, or on neither the subject nor the object.

The goal of this paper is to propose an analysis of finite verb agreement in Icelandic. An essential element of this analysis is the idea that the features involved in agreement—typically, person, number, and gender—are grouped in a feature structure, referred to as AGR (for "agreement features"), present in the f-structure representation of nominal constituents, but also, crucially, in the f-structure representation of the clause. These features of the clause are normally overtly expressed on the finite verb, as well as on the agreement trigger, if there is one. In this we follow Alsina & Vigo (2014, 2017); Vigo (2016) and other work. In addition, we assume that there are well-formedness constraints on the f-structure that apply according to the principles of Optimality Theory (OT).

In section 2, we present the theory of agreement that we propose as an alternative to the standard LFG approach to agreement to account for the basic facts of agreement in Icelandic, arguing for the AGR feature structure and laying out the OT constraints adopted. We also bring out similarities with Hindi. In section 3, we show how this analysis provides an immediate account of the phenomenon of "long-distance agreement" in Icelandic, where the object of an infinitival complement can agree with the verb under which the infinitival complement is embedded. In section 4, we show how this analysis can be adapted to English with minimal changes so as to explain the agreement facts in locative inversion, as well as in non-inverted constructions. Finally, in section 5, we draw the main conclusions and make some comparisons with other theories.

2 An AGR-based theory of agreement

The two central elements of the theory of agreement to be presented are the AGR feature structure and the set of OT constraints that refer to it. Unless otherwise indicated, the data presented in this section is from Icelandic.

2.1 Arguments for AGR

The AGR feature bundle, as used in this theory, fulfils two functions: 1) it groups together the features involved in agreement and 2) it provides agreement targets (e.g. verbs and adjectives) with their own agreement features, separate from those of potential agreement triggers (typically, nouns and noun phrases). The former function is found in the INDEX feature of much work in HPSG (Pollard & Sag, 1994, among others); in this work, the representation of nouns and NPs is assumed to include this feature, which specifies the person, number, and gender of the noun or NP. The way agreement is handled in this line of work is by having the verb or other agreement target specify in its lexical entry one or more of the INDEX features of one of its GFs, but the verb does not have its own set of agreement features. The second function of AGR—that of giving the verb and the clause its own set of agreement features, separate from those of the agreement trigger—is proposed in Kathol (1999), within HPSG, and is used in LFG in Haug & Nikitina (2012) for the analysis of participial clauses in Latin. It is applied for the first time in an LFG work to the analysis of finite verb agreement in Alsina & Vigo (2014); see also Haug & Nikitina (2016); Vigo (2016); Alsina & Vigo (2017).[1] In these works, the f-structure of the verb and the clause includes an AGR feature structure that is structure-shared with the AGR of one of the GFs of that verb or clause.[2]

AGR is not linked to any particular GF at the lexical level, but may be linked to one at the f-structure level by means of OT constraints. For example, the lexical entries of the Icelandic verb forms *hjálpuðum* 'helped.1.pl' and *líkuðu* 'liked.3.pl' (used in (1a) and (1b), respectively) provide this information to the f-structure of the sentence:

(3) a. *hjálpudum* : $\begin{bmatrix} \text{PRED} & \text{'}\textit{help}\ \langle\ \text{arg arg}\ \rangle\text{'} \\ \text{TENSE} & \text{PAST} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{CASE} & \text{DAT} \end{bmatrix} \end{bmatrix}$

b. *líkuðu* : $\begin{bmatrix} \text{PRED} & \text{'}\textit{like}\ \langle\ \text{arg arg}\ \rangle\text{'} \\ \text{TENSE} & \text{PAST} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{DAT} \end{bmatrix} \end{bmatrix}$

[1] The present theory does not deal with NP-internal agreement and could be enriched, if necessary, with an additional set of features such as CONCORD as in King & Dalrymple (2004). In their work, in the representation of NPs, INDEX is used as the equivalent of our AGR.

[2] For languages where verb forms show agreement with more than one GF in the sentence, we also assume that there is only one AGR bundle. For example, morphs that specifically target the object are linked to the AGR of the OBJ, not to the AGR of the clause.

One of the advantages of having AGR is that it allows us to avoid stipulating which GF each verb form must agree with. Which GF the verb agrees with follows from general principles.[3]

A second advantage afforded by AGR is that it also allows for a unique lexical representation of verb forms that alternate between agreement with the subject, agreement with the object, and agreement with no GF, e.g. auxiliary verbs like *vera* 'to be' and *hafa* 'to have'. This alternation is illustrated for the 3rd person plural *voru* in (4).

(4) a. **Drengirnir** voru/*var sýndir honum
the.boys.nom.m.pl *were*.3.pl/*sg *shown*.nom.m.pl *him*.dat
'The boys were shown to him' (based on Andrews, 1990)

b. Henni voru/*var gefnir/*gefin/*gefið **hestarnir**
she.dat *were*.3.pl/*sg *given*.nom.m.pl/*f.sg/*n.sg *horses*.nom.m.pl
'She was given horses' (based on Sigurðsson, 2004)

The form *voru* agrees with the subject, in (4a), or with the object, in (4b). The standard LFG approach would require us to have two diferent lexical entries for *voru* or a lexical entry with a disjunction of different sets of functional annotations, as in Butt & Sadler (2003). Using AGR we only need one lexical entry for each form, namely:

(5) *voru* : $\begin{bmatrix} \text{PRED} & \text{'}be\ \langle \text{arg} \rangle\text{'} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \end{bmatrix}$

A third advantage of having AGR in the representation of verbs and clauses is that it also allows us to explain cases of long-distance agreement, in which the verb agrees with the object of its complement (as we shall see in section 3).

2.2 OT constraints on AGR

We assume an OT-LFG approach where general constraints are applied to candidate f-structures. Candidates from the same input share the same meaning and for our purposes only differ with respect to agreement. We assume that all candidates comply with Consistency, Completeness, and Coherence (Bresnan, 2000; Kuhn, 2003; Prince & Smolensky, 2004).

[3]We are not concerned here with the morphological aspects of the verb forms involved. Current LFG approaches to morphology that deal with agreement include Bögel & Butt (2013); Butt & Sadler (2003); Dalrymple (2015); Kaplan & Butt (2002). Our approach to agreement is compatible with alternative approaches to morphology, such as the ones mentioned.

All clauses whose verb form agrees with some GF satisfy AGRSHARE, i.e. the requirement that the AGR features of the clause unify with those of a dependent GF.[4]

(6) AGRSHARE : $\begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{GF} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}_f$

For f-structure f that maps to a constituent of category V

The choice of the agreeing GF (GF_{AGR}) is determined by a set of constraints. Languages differ in the ranking of the constraints that refer to GF_{AGR}.

In order to explain the facts of agreement in Icelandic, we need to block agreement with non-nominative GFs, even in the case of subjects. This task is performed by constraint *AGRCASE, which bars unifying the AGR of the verb with the AGR of a dependent GF that is not nominative:

(7) *AGRCASE : $*\begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{GF} & \begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{CASE} & \neg\text{NOM} \end{bmatrix} \end{bmatrix}_f$

For f-structure f that maps to a constituent of category V

As will be shown later, this formulation of the constraint is preferable to a formulation that requires GF_{AGR} to be nominative (i.e. caseless scenarios).

Given that in Icelandic there is only one nominative per sentence (leaving aside copular sentences, to be analyzed in §2.3), the application of both *AGRCASE and AGRSHARE guarantees that the verb agrees with a nominative GF if there is one. Usually this results in subject agreement, as nominative is assigned by default to the subject. Therefore, subject agreement in Icelandic is just a consequence of the interaction of case assignment rules and our constraints. But if the subject is non-nominative (a "quirky case subject"), the two constraints mentioned are satisfied by agreement with a nominative object.

When the clause lacks a nominative GF, we find the verb in the 3rd person singular neuter. Although finite verb forms in Icelandic do not show differences in terms of gender, participles do show gender agreement and the form they adopt when there is no nominative GF for them to agree with is the neuter singular form. We can account for the 3rd person singular neuter forms in such cases as a result of satisfying AGRDEF(AULT):

(8) AGRDEF : $\begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{defgen} & \end{bmatrix} \end{bmatrix}_f$

For f-structure f that maps to a constituent of category V

[4] Identity of structure, or structure-sharing, in f-structures and in constraints on f-structures is signalled by means of the tag notation commonly used in HPSG.

AGRDEF constrains the features of the verb's AGR to be 3rd person singular and in the default gender (defgen) of the language. Given that there is no single gender value that can be universally considered to be the default gender, we assume that *defgen* is a placeholder that is replaced at the definition of the constraint by the appropriate attribute-value pair specific for every language: GEND = NEUT in Icelandic, GEND = MASC in Hindi, or no pair in English (as it lacks the GEND attribute), etc.

We provisionally propose the following ranking of the three constraints assumed so far:

(9) CONSTRAINT RANKING (PROVISIONAL):
*AGRCASE ≫ AGRSHARE ≫ AGRDEF

We shall now see how these constraints operate in selecting the grammatical verb form with a few examples. For each sentence, we need to consider the various competing candidates, which, as stated earlier, are f-structures. For ease of exposition we will use sentences in place of the corresponding f-structures, we only consider the more harmonic candidates and will start by seeing how the choice between the two verb forms in (4a) is decided. The following three sentences correspond to the three competing f-structures that we will consider.

(10) a. **Drengirnir** voru sýndir honum
the.boys.nom.m.pl *were*.3.pl *shown*.nom.m.pl *him*.dat
'The boys were shown to him'

b. * Drengirnir var sýnt honum
the.boys.nom.m.pl *was*.3.sg *shown*.nom.n.sg *him*.dat

c. * Drengirnir var sýndum **honum**
the.boys.nom.m.pl *was*.3.sg *shown*.dat.m.sg *him*.dat

The agreeing expression is shown in boldface: the nominative subject in (10a), the dative object in (10c), and there is no agreeing expression in (10b).[5] The optimization tableau is given in Tableau 1.

	*AGRCASE	AGRSHARE	AGRDEF
☞ (10a)			*
(10b)		*!	
(10c)	*!		*

Tableau 1: optimization for (10)

(10a) is selected as the grammatical structure, because the constraints violated by the alternative candidates are more highly ranked than the one it violates. (10c)

[5]The passive participle agrees in gender and number with its GF_{AGR} if there is one. In addition, it has a case specification which has to match that of the GF_{AGR}, as in (10) above, or be in the default nominative case if there is no GF_{AGR}, as in (12a).

is discarded because it violates the most highly ranked constraint of the three under consideration—*AGRCASE—as the verb agrees with a dative expression. (10b) is discarded because it violates the second constraint in ranking—AGRSHARE—given that the clausal AGR is not shared with that of any GF in the clause.[6] This leaves (10a) as the optimal candidate, even though it violates AGRDEF, the lowest ranking of the three. The f-structure of the optimal candidate is given in Figure 1.[7] The AGR of the clause is shared with that of a GF of the clause, satisfying AGRSHARE, and with that of a GF that is not non-nominative, satisfying *AGRCASE. The fact that its features are not 3rd person singular neuter causes a violation of the low ranking AGRDEF, which does not make the structure ungrammatical.

$$\begin{bmatrix} \text{PRED} & \text{`}\textit{be.shown}\ \langle\ \cancel{\text{arg}}\ \text{arg}_1\ \text{arg}_2\ \rangle\text{'} \\ \text{AGR} & \boxed{1} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`}\textit{boy}\text{'} \\ \text{AGR} & \boxed{1}\begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{MASC} \end{bmatrix} \\ \text{CASE} & \text{NOM} \end{bmatrix}_2 \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \end{bmatrix} \\ \text{CASE} & \text{DAT} \end{bmatrix}_1 \end{bmatrix}$$

Figure 1: f-structure of (10a)

Let us consider now a structure in which the verb agrees with a nominative object, as in (1a), repeated here as (11a). The three competing candidates are: agreement with the nominative object, as in (11a), no agreement, in (11b), and agreement with the dative subject, in (11c).

(11) a. Henni líkuðu **hestarnir**
she.dat *liked*.3.pl *horse*.nom.m.pl
'She liked the horses.'

b. * Henni líkaði hestarnir
she.dat *liked*.3.sg *horse*.nom.m.pl

[6]The difference between (10b) and (10c) with respect to agreement is revealed by the form of the participle. On the assumption that the finite verb form and the participle share the same AGR features, the neuter singular form of the participle in (10b) indicates that there is no agreeing GF, whereas the dative masculine singular form in (10c) indicates that the agreeing GF is *honum*, dative masculine singular.

[7]In the f-structures represented in this paper we are using the following two notational conventions: Crossing out of the most promiment argument in the PRED value signals the suppression of the logical subject of the passive; the subscripted number in the PRED value show the correspondence of each argument with a GF.

c. * **Henni** líkaði hestarnir
she.dat *liked*.3.sg *horse*.nom.m.pl

As shown in the optimization tableau for (11), in Tableau 2, the violations of AGR-SHARE and *AGRCASE that (11b) and (11c), respectively, incur leave (16a) as the grammatical structure, in which neither of these constraints is violated. The f-structure of the grammatical (11a) is shown in Figure 2.

	*AGRCASE	AGRSHARE	AGRDEF
☞ (11a)			*
(11b)		*!	
(11c)	*!		*

Tableau 2: optimization for (11)

$$\begin{bmatrix} \text{PRED} & \text{`}\textit{like}\ \langle\ \text{arg}_1\ \text{arg}_2\ \rangle\text{'} \\ \text{AGR} & \boxed{1} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{FEM} \end{bmatrix}_1 \\ \text{CASE} & \text{DAT} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`}\textit{horse}\text{'} \\ \text{AGR} & \boxed{1} \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{MASC} \end{bmatrix}_2 \\ \text{CASE} & \text{NOM} \end{bmatrix} \end{bmatrix}$$

Figure 2: f-structure of (11a)

Finally, we need to consider the type of sentence where there is no trigger of agreement for the verb, i.e., the verb agrees with no GF. The two candidates to take into account are (12a), where the clause contains no GF whose AGR is structure-shared with that of the clause, and (12b), where the agreeing GF is the subject.

(12) a. Þeim var hjálpað
them.dat *was*.3.sg *helped*.nom.n.sg
'They were helped'

b. * **Þeim** voru hjálpaðum
them.dat *were*.3.pl *helped*.dat.m.pl

The corresponding tableau 3 indicates that (12a), despite not having any agreeing GF, is the optimal candidate. It shows that a grammatical structure can violate

AGRSHARE. In such a situation, the effects of AGRDEF are revealed, requiring the verb form to be in the 3rd person singular—other candidates, not shown here, with different AGR values (plural, 1st person, feminine, etc.), are ruled out because they violate AgrDef. The f-structure of the grammatical (12a) is given in Figure 3.

	*AGRCASE	AGRSHARE	AGRDEF
☞ (12a)		*	
(12b)	*!		*

Tableau 3: optimization for (12)

$$\begin{bmatrix} \text{PRED} & \text{`}\textit{be.helped}\ \langle\ \sout{\text{arg}}\ \text{arg}_1\ \rangle\text{'} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{NEUT} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix}_1 \\ \text{CASE} & \text{DAT} \end{bmatrix} \end{bmatrix}$$

Figure 3: f-structure of (12a)

2.3 Subject agreement

Up to this point, we have not introduced a principle accounting for the observation that the verb agrees preferentially with the subject. So far, this fact follows from the combined effect of *AGRCASE, which excludes any non-nominative as the agreeing GF, and the principles of case assignment in Icelandic, by which the subject is assigned nominative case by default and co-occurring GFs are in other cases. The issue that we haven't yet addressed is what happens when there are two nominative GFs in the sentence. We find this in Icelandic copular sentences with two nominative GFs, i.e. subject and complement. In this situation the verb always agrees with its subject (leaving aside copular constructions in which the subject is *þetta* or *það*, to which we will return). See for example the following data from Sigurðsson, 1996:

(13) a. **Bítlarnir** **hafa**/*hefur lengi verið frægasta
the.Beatles.nom.m.pl *have*.3.pl/*sg *long been most.famous*
hljómsveitin
the.band.nom.m.sg
'The Beatles have long been the most famous band'

b. **Frægasta hljómsveitin hefur**/*hafa lengi verið
most.famous the.band.nom.m.sg *have*.3.sg/*pl *long been*
Bítlarnir
the.Beatles.nom.m.pl
'The most famous band has long been The Beatles'

The claim that in the presence of two nominative NPs finds support in the facts from Hindi. Some Hindi transitive sentences allow the subject to alternate between ergative case and nominative case, depending on the aspect of the verb (see Butt & King, 2004 for the full compexity of Hindi/Urdu case). Following the observations in Mohanan (1994, 2016), if the subject is ergative and there is a nominative object, the verb agrees with the object, as expected: (14). When the subject is nominative, the verb agrees with it, despite the presence of a nominative object: (15) (examples from Mohanan, 2016).

(14) ravii-ne / niinaa-ne **santraa** khaayaa/*khaaii
Ravi-erg.m / *Nina*-erg.f *orange*-nom.m *eat*.perf.m.sg/*f.sg
'Ravi/Nina ate orange(s)'

(15) **niinaa** santraa / roTii khaaegii/*khaaegaa
Nina-nom.f *orange*-nom.m / *bread*-nom.f *eat*.fut.f.sg/*m.sg
'Nina will eat orange/bread'

The preference for the subject as an agreement trigger is explained by positing a new principle named AGRSUBJ, defined as in (16) (informally, $\text{GF}_{\text{AGR}} = \text{SUBJ}$), and placing it in the provisional ranking of constraints in (17).

(16) AGRSUBJ : $\begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}_f$
For f-structure f that maps to a constituent of category V

(17) CONSTRAINT RANKING (PROVISIONAL VERSION 2):
*AGRCASE ≫ AGRSUBJ ≫ AGRSHARE ≫ AGRDEF

In sentences with two nominatives, the choice between the two is settled in favor of the subject.[8]

The position of AGRSUBJ in the hierarchy in (17) is decided as follows. That constraint must rank below *AGRCASE because there is never agreement with a

[8]The person restriction reported in Sigurðsson (2004) can be interpreted as a prominence matching constraint: the most prominent case feature—nominative—may be aligned with the more prominent person features—first and second—only if they correspond to the most prominent argument at argument structure. By this constraint, the nominative object of *líka* 'like' cannot be first or second person. But the nominative complement of the copula, being the single and, therefore, most prominent argument of this verb is not prevented from being first or second person. Inverse agreement in copular clauses arises only when the object of the copula is *það* or *þetta*. We can assume that these forms are lexically specified for the features of AGR, as indicated in Sigurðsson (2004); and that they acquire these features through identity of AGR between a predicative element and its subject.

non-nominative subject. AGRSUBJ must rank higher than AGRDEF because agreement with a nominative subject is obligatory. In order to determine the relative ranking between AGRSUBJ and AGRSHARE, we can use evidence from Jónsson (2016): In some variants of Icelandic, lack of agreement is preferred over nominative object agreement, suggesting AGRDEF ≫ AGRSHARE in these variants, a reordering of constraints that shows an advantage of using an OT approach. Nominative subject agreement remains obligatory in these variants. Given that AGRSUBJ ≫ AGRDEF, we can deduce that AGRSUBJ ≫ AGRSHARE.

3 Raising and long-distance agreement

In raising constructions, the raising verb has two dependent GFs: the subject and the complement (the embedded clause). If the subject is nominative, we expect that the verb should agree with it. If it is not, the expectation would be that the verb agrees in the 3rd singular with its clausal complement. Facts such as (18) show that the verb does not necessarily have 3rd person singular features, but may have the features of the nominative object of the embedded clause, either in the active or the passive (Sigurðsson, 2004). The examples below correspond to the three most harmonic candidates: showing agreement with the object of the complement clause, (18a), showing agreement with no nominal GF, (18b), and showing agreement with the dative subject of the raising verb, (18c). The corresponding optimization is shown in Tableau 4.

(18) a. Henni eru taldir hafa verið sýndir
she.dat *are*.3.pl *believed*.nom.m.pl *have*.inf *been* *shown*.nom.m.pl
bilarnir
car.nom.m.pl
'She is believed to have been shown the cars'

b. * Henni er talið hafa verið sýndir
she.dat *is*.3.sg *believed*.nom.n.sg *have*.inf *been* *shown*.nom.m.pl
bilarnir
car.nom.m.pl

c. * **Henni** er talinni hafa verið sýndir
she.dat *is*.3.sg *believed*.dat.f.sg *have*.inf *been* *shown*.nom.m.pl
bilarnir
car.nom.m.pl

The f-structure of the grammatical (18a) is shown in Figure 4. The main points to highlight are the following: the AGR of the embedded clause is structure-shared with that of its nominative object, bearing in mind that the subject is dative and, therefore, prevented from agreeing; this is a type of covert agreement, as the verb of the embedded clause is an infinitive and, as such, does not express any agreement feature; the subject of the embedded clause undergoes raising, i.e. is structure-shared with the subject of the embedding clause; the AGR of the embedding clause

	*AGRCASE	AGRSUBJ	AGRSHARE	AGRDEF
☞ (18a)		**		**
(18b)		**	*!	
(18c)	*!*			**

Tableau 4: optimization for (18)

cannot be shared with that of its subject, which is dative, but instead is shared with that of its object—the complement clause. Nothing said so far prevents the sharing of the AGR of a clause with that of an embedded clause, and that is what happens here. The apparent long-distance agreement seen in (18a) is, in fact, a combination of two local agreement relations: the sharing of the AGR of the raising clause with the AGR of its infinitival complement and the sharing of this AGR with that of the object of the infinitive.

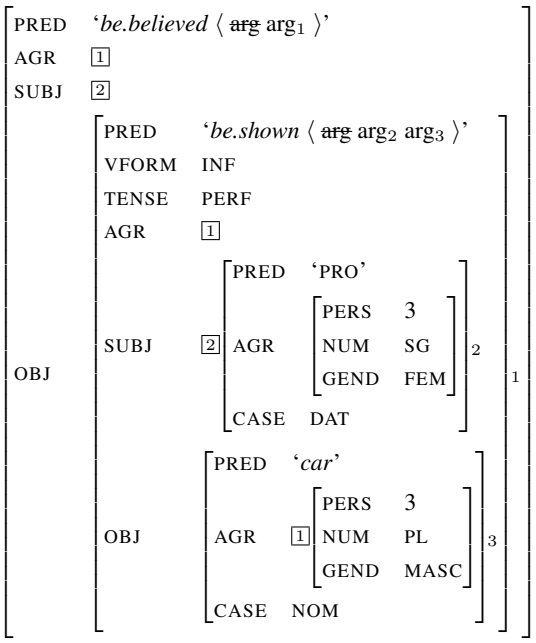

Figure 4: f-structure of (18a)

The facts of agreement in raising constructions such as these argue for formulating *AGRCASE as stated in (7), that is, as a negative constraint (the AGR of a dependent GF whose case is not nominative cannot be involved in AGR-sharing), rather than as a positive constraint (the AGR of a clause must be shared with that of a nominative dependent). There is no evidence that a complement clause is specified for case. So, although we cannot say that a complement clause is nominative (or any other case specification), we can say that it lacks case and therefore is not non-nominative (but see Butt (2014), where the infinitive complement clauses in a

similar construction in Hindi/Urdu bear nominative case).

When the clause embedded under a raising verb does not include a nominative argument, the raising verb shows the default 3rd person singular form, as in (19a). (19b) is the competing candidate in which both the raising verb and the infinitive agree with the dative subject. The optimization is given in Tableau 5.

(19) a. Þeim virðist hafa verið hjálpað
they.dat *seems*.3.sg *have*.inf *been helped*.nom.n.sg
'They seem to have been helped'

b. * **Þeim** virðast hafa verið hjálpaðum
they.dat *seems*.3.pl *have*.inf *been helped*.dat.m.pl

	*AGRCASE	AGRSUBJ	AGRSHARE	AGRDEF
☞ (19a)		**	*	
(19b)	*!*			**

Tableau 5: optimization for (19)

The facts of raising sentences are explained by the same set of constraints that we proposed for monoclausal structures. The agreement of the raising verb with the nominative object of the embedded clause is possible because the raising verb shares its AGR with that of the embedded clause, regardless whether there is a nominative object. However, not all embedded clauses allow their AGR to be shared with that of the embedding clause. The evidence indicates that the AGR of an embedded clause is only available for sharing with the AGR of the higher clause if there is raising-to-subject (RTS) from the embedded clause. Examples, like (20) and (21) do not allow agreement of the main clause verb with the nominative object of the embedded clause.

(20) Mér hefur/*hafa alltaf virst honum líka **bækur**
I.dat *has*.3.sg/*pl *often seemed he*.dat *like*.inf *book*.nom.m.pl
'It has often seemed to me that he likes books' (Schütze, 1997)

(21) Mér virðist/*?virðast stráknum líka **þessir bílar**
I.dat *seems*.3.sg/*?pl *the.boy*.dat.m.sg *like*.inf *these*.nom.pl *cars*.nom.pl
'It seems to me that the boy likes these cars' (Watanabe, 1993)

In examples (20) and (21) there is no RTS as the subject of the matrix clause is occupied by an argument of this verb. This is the crucial difference between these examples and the cases in which structure sharing is found between the matrix and the embedded AGRs.

In order to block the possibility of AGR-sharing across clauses without RTS, we posit the constraint Clausal Transparency (CLTRANS):[9]

[9]One could state this constraint by saying that cross-clausal sharing of AGR only occurs when

(22) CLTRANS : $\begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{GF} & \left[\text{AGR} \quad \boxed{1}\right]_g \end{bmatrix}_f \rightarrow \begin{bmatrix} \text{SUBJ} & \boxed{2} \\ \text{GF} & \left[\text{GF} \quad \boxed{2}\right]_g \end{bmatrix}_f$

For f-structures f, g that map to constituents of category V.

The final ranking of OT constraints is assumed to be as follows:

(23) CONSTRAINT RANKING (FINAL):
CLTRANS, *AGRCASE $\gg$ AGRSUBJ $\gg$ AGRSHARE $\gg$ AGRDEF

The relative order of CLTRANS and *AGRCASE is impossible to determine as we have not found any instance in which either one is violated by an optimal candidate. We therefore assume both rank equally.

In order to see the effect of CLTRANS on a sentence like (20), let us consider the three candidates in (24): (24a), without sharing of the two clausal AGRs and with default agreement on the matrix AGR; (24b), without sharing of the clausal AGR and with agreement of the matrix verb with its dative subject; and (24c), with sharing of the clausal AGRs and with long-distance agreement. Tableau 6 shows that (24a) is the optimal candidate, as the other two candidates violate one of the two highest ranking constraints in (23).

(24) a. Mér hefur alltaf virst honum líka bækur
I.dat *has* *often seemed he*.dat *like*.inf *book*.nom.m.pl
'It has often seemed to me that he likes books'

b. * MÉR hef alltaf virst honum líka bækur
I.dat *have*.1.sg *often seemed he*.dat *like*.inf *book*.nom.m.pl

c. * Mér hafa alltaf virst honum líka **bækur**
I.dat *have*.3.pl *often seemed he*.dat *like*.inf *book*.nom.m.pl

	CLTRANS	*AGRCASE	AGRSUBJ	AGRSHARE	AGRDEF
☞ (24a)			**	*	*
(24b)		*!	*		**
(24c)	*!		**		**

Tableau 6: optimization for (24)

The analysis of long-distance agreement proposed here lends itself to accounting for other instances of long-distance agreement discussed in the literature. For example, there is a construction with long-distance agreement in Hindi/Urdu according to Bhatt (2005); Butt (2014) in which the infinitive may agree with its

the lower AGR is that of an XCOMP. However, we are using a reduced inventory of GFs that does not include COMP or XCOMP, as in Alsina (1996); Alsina et al. (2005); Patejuk & Przepiórkowski (2016). See Patejuk & Przepiórkowski (2016) for evidence against distinguishing XCOMP from OBJ as GFs.

nominative object and in turn with the subordinating verb. In this construction the infinitive reflects its AGR features by means of a suffix. Another phenomenon of long-distance agreement is discussed in Polinsky & Potsdam (2001) and Haug & Nikitina (2016) involving (optional) agreement of a verb with the absolutive argument of the complement clause in a structure without RTS in Tsez. A possible way to analyze these facts within the present approach is to assume AGR-sharing of the embedded clause and the embedding clause, as a result of a different position of CLTRANS in the hierarchy of constraints in Tsez. Developing the analysis of Hindi/Urdu and Tsez long-distance agreement is beyond the scope of this paper.

4 Agreement and locative inversion in English

Although English has traditionally been analyzed as a language with subject agreement, the facts of locative inversion (LocInv) indicate that English is not that different from Icelandic. When an oblique case locative appears in subject position, the verb agrees with the postverbal NP, which is analyzed as a complement, following Bresnan (1994).[10]

(25) In the swamp was/*were found **a child**

(26) In the swamp were/*was found **two children**

We assume as in Bresnan (1994) that the PP locatives are oblique case and the postverbal NPs in these constructions are direct case. We cannot analyze these constructions exactly like structures with quirky case subjects in Icelandic because the postverbal NP in English LocInv is not nominative. Although distinctions in terms of case on NPs in English are only found in pronouns and pronouns are pragmatically hard to use in the postverbal NP position of LocInv because of the presentational function of this position, which is inconsistent with the anaphoric function of pronouns, pronouns can be used deictically in this position and can only be used in their accusative (not in their nominative) form (Bresnan, 1994):

(27) Among the guests of honor was sitting HER [pointing]

It seems that the relevant distinction in terms of case for agreement in English is between direct and oblique case (not between nominative and non-nominative). In order to analyze the facts of English, we can assume that *AGRCASE is relegated to a low position in the hierarchy of constraints, so that it has no effect in English, and in its place in the ranking of constraints we can assume a *AGROBL constraint prohibiting agreement (AGR sharing) with an oblique case argument.

[10] Bresnan (1994, p. 95, footnote 31) notes that in rare instances, in this construction, the verb is found in the 3rd singular not agreeing with the postverbal NP, a possibility reported by a reviewer. The analysis within the present theory would imply that for some speakers the two lower-ranking constraints in (23), AGRSHARE and AGRDEF, rank equally, giving rise to two outcomes in free variation: an agreeing and a non-agreeing form.

With this difference with Icelandic, we can assume the rest of the theory unchanged for English. In order to explain the grammatical agreement form in (26), we can consider the competing candidates in (28): with agreement of the verb with the direct case object, (28a); without agreement, with the default morphology on the verb, (28b); with agreement with the oblique case subject, (28c). The competition is resolved as in Tableau 7 and the f-structure of the well-formed (28a) is given in Figure 5.

(28) a. In the swamp were found **two children**

b. * In the swamp was found two children

c. * **In the swamp** was found two children

	CLTRANS	*AGROBL	AGRSUBJ	AGRSHARE	AGRDEF
☞ (28a)			*		*
(28b)			*	*!	
(28c)		*!			

Tableau 7: optimization for (28)

$$\begin{bmatrix}
\text{PRED} & \text{`}\textit{be.found}\ \langle\ \sout{\text{arg}}\ \text{arg}_1\ \text{arg}_2\ \rangle\text{'} \\
\text{TENSE} & \text{PAST} \\
\text{AGR} & \boxed{1} \\
\text{SUBJ} & \begin{bmatrix}
\text{PRED} & \text{`}\textit{in}\ \langle\ \text{arg}_3\ \rangle\text{'} \\
\text{OBJ} & \begin{bmatrix}
\text{PRED} & \text{`}\textit{swamp}\text{'} \\
\text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix}
\end{bmatrix}_3 \\
\text{CASE} & \text{OBL}
\end{bmatrix}_2 \\
\text{OBJ} & \begin{bmatrix}
\text{PRED} & \text{`}\textit{child}\text{'} \\
\text{QUANT} & \text{NUMERAL} \\
\text{AGR} & \boxed{1}\begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix} \\
\text{CASE} & \text{DIR}
\end{bmatrix}_1
\end{bmatrix}$$

Figure 5: f-structure of (28a)

The facts of LocInv in raising structures are also covered by our theory; the verb is correctly predicted to agree with the complement of the embedded clause if the raised subject is a locative PP: (29) and (30) (adapted from Bresnan, 1994).

(29) On the hill appears/*appear to be located **a cathedral**

(30) On the hill appear/*appears to be located **two towers**

As in Icelandic, sharing of AGRs across clauses is only possible if RTS is involved (as required by CLTRANS), as in (29) and (30). The candidates considered for (29) are given in (51) and the corresponding optimization in (52).

(31) a. On the hill appear to be located **two towers**
b. * On the hill appears to be located two towers
c. * **On the hill** appears to be located two towers

	CLTRANS	*AGROBL	AGRSUBJ	AGRSHARE	AGRDEF
☞ (31a)			**		**
(31b)			**	*!	
(31c)		*!*			

Tableau 8: optimization for (31)

5 Conclusion

The theory of agreement proposed in this paper rests on two fundamental ideas: a) the f-structure representation of both the clause and nominal expressions includes the feature matrix AGR containing the features involved in agreement (PERS, NUM, GEND, and others); and b) the value of the verbal AGR is determined by a set of OT constraints. The lexical entry of a finite verb form in languages like English, Icelandic, or Hindi specifies the features of AGR, but does not specify what GF this AGR corresponds to. Verbal agreement is the sharing of the verb's AGR with that of one of its dependent GFs. The job of the OT constraints is to ensure that the right GF is chosen to share its AGR with that of the clause it belongs to.

Some of the consequences of the theory are the following:

1. Agreement with nominative expressions in Icelandic:
   - High-ranking *AGRCASE excludes expressions with case values other than nominative for AGR sharing: this restricts eligible agreement triggers to nominative GFs and clauses (lacking in case values).[11]
   - AGRSUBJ breaks the tie in favor of the subject when two expressions are nominative.
2. Long-distance agreement:
   - Apparent long-distance agreement (a verb agreeing with the complement of an embedded clause) is a set of local agreement relations: the AGR of the main clause is shared with that of the embedded clause, which, in turn, is shared with the AGR of a complement.

[11] In English, *AGROBL, instead of *AGRCASE, leaves GFs with direct case (as well as those lacking in case) as potential triggers of agreement.

- The impression of long-distance agreement is enhanced by the fact that the AGR-sharing of an infinitival clause is covert agreement, i.e., not morphologically encoded.
- Cross-clausal AGR-sharing is restricted to RTS structures (CLTRANS).

3. Verbal features are 3rd sg when there is no agreement:

   - Agreement fails to arise when AGRSHARE cannot be satisfied: No available GF is nominative or a clause that undergoes RTS.
   - In the absence of AGR-sharing, AGRDEF requires the clausal AGR to be 3rd sg (neuter in Icelandic).
   - This 3rd sg form is exactly the same that is used when agreement with a 3rd sg constituent is required (not a homophonous form).

Proposals within LFG to account for the Icelandic agreement facts discussed here can be found in Andrews (1990) and Otoguro (2005). Space limitations prevent us from making an in-depth comparison of those proposals with the present one. One clear advantage of the present proposal is that, whereas both Andrews (1990) and Otoguro (2005) treat the verb form that is used in the absence of agreement as homophonous with the form that agrees in the 3rd person singular, the present proposal assumes that there is only one 3rd singular form. A verb form whose AGR includes the features of 3rd person and number singular can be used both when this AGR is shared with another AGR with the same features or when it is not shared with any AGR, as the result of AGRDEF.

References

Alsina, Alex. 1996. *The role of argument structure in grammar: evidence from Romance*. Stanford: CSLI Publications.

Alsina, Alex, Karuvannur Puthanveettil Mohanan & Tara Mohanan. 2005. How to get rid of the COMP. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG05 conference*, 21–41. Stanford: CSLI Publications.

Alsina, Alex & Eugenio M. Vigo. 2014. Copular inversion and non-subject agreement. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG14 conference*, 5–25. Stanford: CSLI Publications.

Alsina, Alex & Eugenio M. Vigo. 2017. Fixing LFG to account for direct-inverse agreement: the case of Plains Cree. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG17 conference*, Stanford: CSLI Publications.

Andrews, Avery D. 1990. The VP-Complement analysis in Modern Icelandic. In Joan Maling & Annie Zaenen (eds.), *Syntax and semantics: Modern Icelandic syntax*, 165–185. San Diego: Academic Press.

Bhatt, Rajesh. 2005. Long distance agreement in hindi-urdu. *Natural Language and Linguistic Theory* 23(4). 757–807.

Bresnan, Joan. 1994. Locative inversion and the architecture of Universal Grammar. *Language* 70(1). 72–131.

Bresnan, Joan. 2000. Optimal syntax. In Joost Dekkers, Frank van der Leeuw & Jerome van der Weijer (eds.), *Optimality theory: phonology, syntax and acquisition*, 334–385. Oxford: Oxford University Press.

Butt, Miriam. 2014. Control vs. complex predication. *Natural Language and Linguistic Theory* 32(1). 155–190.

Butt, Miriam & Tracy Holloway King. 2004. The status of case. In Veneeta Dayal & Anoop Mahajan (eds.), *Clause structure in South Asian languages*, 153–198. Dordrecht: Kluwer Academic Publishers.

Butt, Miriam & Louise Sadler. 2003. Verbal morphology and agreement in Urdu. In Uwe Junghangs & Luka Szucsich (eds.), *Syntactic structures and morphological information*, 57–100. Berlin: Mouton de Gruyter.

Bögel, Tina & Miriam Butt. 2013. Possessive clitics and *ezafe* in Urdu. In Kersti Börjars, David Denison & Alan Scott (eds.), *Morpho-syntactic categories and the expression of possession*, 291–322. Amsterdam: John Benjamins.

Dalrymple, Mary. 2015. Morphology in the LFG architecture. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG15 conference*, 64–83. Stanford: CSLI Publications.

Haug, Dag Trygve Truslew & Tanya Nikitina. 2012. The many cases of non-finite subjects: the challenge of "dominant" participles. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG12 conference*, 292–311. Stanford: CSLI Publications.

Haug, Dag Trygve Truslew & Tanya Nikitina. 2016. Feature sharing in agreement. *Natural Language and Linguistic Theory* 34. 865–910.

Jónsson, Jóhannes Gísli. 2016. Testing agreement with nominative objects. *Working papers in Scandinavian syntax* 97. 57–75.

Kaplan, Robert M. & Miriam Butt. 2002. The morphology-syntax interface in LFG. Abstract in *Proceedings of the LFG15 Conference*.

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In *Studies in contemporary Phrase Structure Grammar*, 223–274. Cambridge: Cambridge University Press.

King, Tracy Holloway & Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40(1). 69–104.

Kuhn, Jonas. 2003. *Optimality-Theoretic syntax: a declarative approach*. Stanford: CSLI Publications.

Mohanan, Tara. 1994. *Argument structure in Hindi*. Stanford: CSLI Publications.

Mohanan, Tara. 2016. Issues in verb agreement in Hindi. Manuscript.

Otoguro, Ryo. 2005. Agreement and path specification in Icelandic. In *Proceedings from the annual meeting of the Chicago Linguistic Society*, vol. 41 1, 385–399. Chicago: Chicago Linguistic Society.

Patejuk, Agnieszka & Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 541–559. Stanford: CSLI Publications.

Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19. 583–646.

Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: The University of Chicago.

Prince, Alan & Paul Smolensky. 2004. *Optimality theory: constraint interaction in generative grammar*. Malden, MA: Blackwell Publishing.

Schütze, Carson T. 1997. *INFL in child and adult language: agreement, case and licensing*: Massachusetts Institute of Technology dissertation.

Sigurðsson, Halldór Ármann. 1996. Icelandic finite verb agreement. *Working papers in Scandinavian syntax* 57. 1–46.

Sigurðsson, Halldór Ármann. 2004. Icelandic non-nominative subjects: Facts and implications. In Peri Bhaskararao & Karumuri V. Subbarao (eds.), *Non-nominative subjects*, vol. 2, 137–159. Amsterdam: John Benjamins.

Vigo, Eugenio M. 2016. *Copular inversion and non-subject agreement*: Universitat Pompeu Fabra dissertation.

Watanabe, Akira. 1993. *Agr-based case theory and its interaction with the A-bar system*: Massachussets Institute of Technology dissertation.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1990. Case and grammatical functions: The Icelandic passive. In Joan Maling & Annie Zaenen (eds.), *Syntax and semantics: Modern Icelandic syntax*, 95–136. San Diego: Academic Press.