Long-distance agreement in the Indo-Aryan languages

Dieter Wunderlich (Berlin), Morphology Network, Amsterdam, 17. Jan. 2009

In Hindi-Urdu, as well as in other Indo-Aryan languages, a matrix verb may agree with the object of an embedded verb, a fact that has been captured under the notion of long-distance agreement (LDA). The first theoretical treatments of this phenomenon were Mahajan (1989), proposing a movement analysis, and Wunderlich (1992/1994), proposing verb complex formation. Since then, many papers were concerned with adjusting the phenomenon more finely, particularly in view of the minimalist program (Bhatt 2005, Boeckx 2004, Chandra 2005, 2007). I am not fully convinced that these efforts are on the right track, and when I try here to revitalize my own former account, I do it also for the reason that it seems to be unknown to the younger authors.

I will start with a few considerations about why agreement and case, although they often seem to be narrowly related, are independent phenomena. Then, I will sketch the fundamentals of the relational theory of case (Kiparsky 1992, Wunderlich 1992/1997) and those of the indexical theory of agreement (Pollard & Sag 1991/1994, Wunderlich 1992/1994). With these two theories in mind, one can characterize a small set of (generally motivated) constraints determining case, and another set of constraints determining agreement. The several Indo-Aryan languages appear to vary in one or the other of these constraints differently ordered, however, it is hard to see how this variation could be captured purely syntactically. Long-distance agreement fells out from the combined forces of AGREE! and *AGR(SUBJECT), which leads to verb complex formation. Those languages in which the verb always agrees with the subject (for one or the other reason) do not show LDA. It is important to see that the account of LDA given here is a local one; even for LDA in Tsez (Polinsky & Potsdam 2001) only a local account is rational under the present assumptions.

1. The alleged complementarity of case and agreement.

Case and agreement are often interconnected in complementary ways: Case tends to mark non-prototypical arguments (high-salient objects, low-salient subjects), while agreement tends to mark prototypical arguments (those that are not marked for case). A NOM-ACC system usually patterns with subject agreement, where the verb agrees with the NOM argument. Likewise, an ERG-ABS system more than often patterns with object agreement, where the verb agrees with the ABS argument; in both instances, the target argument bears the less marked (default) case.

However, case and agreement are not always complementary, in fact, they can be independent of each other. An ERG-ABS system can show subject agreement [Warlpiri, Udi (NE Caucasian), Nepali (4)], regardless of whether the higher argument is marked ergative or not; and, more surprisingly, a NOM-ACC system can show object agreement under certain conditions [Marathi (1), Marwari (Rajasthani) (2), and Kutchi Gujarati (3), all three being narrowly related south-western Central Indo-Aryan languages].

Thus, in general, there is no single set of principles or structures that leads to a simultaneous characterization of case *and* agreement.

(1)   Marathi lacks ERG in the 1sg and 2sg pronouns, but nevertheless shows object agreement in the perfect (Wali 2005:54, cited from Patel 2007).
        mi   dagad   phek-l-a
        1SG.NOM   stone.3SG.M   throw-PERF-3SG.M
        'I threw the stone.'

(2) Marwari (Rajasthani) lacks ERG, but nevertheless shows object agreement in the perfect (Magier 1983, cited from Patel 2007)

a. Raam [ghanii laapsii   jiiml-ii   hii]
    Ram   lots   wheat-gruel.F eat.PERF-F be.PAST.F
    'Ram ate lots of wheat gruel.'

b. aap   [Siita-ne dekh-ii]   ho
    you.pl   Sita-ACC   see.PERF-F be.PRES.2PL
    'You have seen Sita.'
    Note that here the auxiliary agrees with the subject.

(3) Kutchi Gujarati lacks overt ERG, but nevertheless shows optional object agreement in the perfect, even if the object is ACC-marked. (Patel 2007)

a. Reena   chokra-ne mar-ti   thi
    Reena.F.NOM boys-ACC   hit-IPF.F.SG be.PAST.F.SG
    'Reena was hitting the boys'   (*subject agreement*)

b. Reena   chokra-ne mar-ya /   mar-yu
    Reena.F.NOM boys-ACC   hit-PERF.M/N.PL / hit-PERF.N.SG
    'Reena hit the boys'   (*object agreement or default*)

(4) Nepali: The verb agrees with the subject, even if it is ERG-marked in the past. (Bickel & Yādava 2000:348)

a. ma   yas   pasal-mā   patrikā   kin-ch-u.
    1SG.NOM   DEM.OBL store-LOC newspaper.NOM buy-NONPAST-1SG
    'I buy the newspaper in this store.' (*subject agreement*)

b. maile   yas   pasal-mā   patrikā   kin-ē.
    1SG.ERG   DEM.OBL store-LOC newspaper.NOM buy-PAST.1SG
    'I bought the newspaper in this store.' (*subject agreement*)

2. Some fundamentals

a. About case. (I consider here only the four structural cases ACC, ERG, DAT, and NOM, leaving aside genitive, partitive and others.)

(i) The simplest way of distinguishing 4 cases is by means of 2 decompositional features.

(ii) Since case is a means to make the arguments of a predicate distinguishable, case (features) should be regarded as relational in nature. Various decompositional accounts of case have been proposed, e.g., ±governed, ±oblique (Bierwisch 1967), where it is however questionable in what sense these features are relational.

(iii) Case features realized on an argument are matched with the requirements of the corresponding predicate. Is there something like a canonical requirement? E.g., what properties of a predicate give rise to features like ±governed, ±oblique?

(iv) For argument processing, the arguments of a predicate are assumed to be hierarchically ordered (from highest to lowest). (This might be correlated with a hierarchy of semantic roles but doesn't need to.)

Most naturally, case features relate to the positions of argument hierarchy.

Kiparsky (1992): +HR= the highest argument ('subj'), +LR=the lowest argument ('object'). Problem: the least marked positions are associated with +features, and the most marked positions with −features.

Wunderlich (1992, 1997): −hr = the highest argument („there is no higher argument"), −lr = the lowest argument („there is no lower argument"). ACC relates to an argument that is not the highest one (+hr), and ERG relates to an argument that is not the lowest one (+lr).

Consequently: DAT is [+hr,+lr], and NOM is [ ]. The often-used mnemotechnic notions S, A and O/T are instances of [−hr,−lr], [−hr,+lr], and [+hr,−lr], resp.

(5) λx DANCE(x) λy λx HIT(x,y) λz λy λx GIVE(x,y,z)

−hr	+hr	−hr	+hr	+hr	−hr
−lr	−lr	+lr	−lr	+lr	+lr
S	O	A	T	G	A

Lexical marking of case: It is possible that a predicate marks a case lexically. There is still a given hierarchy of arguments, but at least one of the automatic features is overridden. Prominent instances: Dative experiencers: [-hr] of the highest argument is lexically overridden by [+hr] („as if there were a higher argument") – so DAT, or, if DAT is not available, ACC is the required case. Ergative in (agentive) intransitive verbs: [-lr] is overwritten by [+lr] („as if there were a lower argument") – so ERG is the required case.

b. About agreement.

Agreement features are φ−features (person, number, gender), which by themselves are not relational. They are characteristic for nouns and pronouns. By referring to properties of possible referents of N (somewhat grammaticalized), they allow to identify and to keep record of the referents across a piece of discourse (Lehmann 1982).

(i) Nouns and pronouns can agree with each other (which generally is nonlocal).

(ii) The parts of a complex noun phrase can agree with each other (which generally is local).

(iii) A predicate can agree with (one or several of) its (nominal) argument(s).

Agreement per se is, thus, not locally restricted. (Pronouns can agree with their nominal antecedents across sentence-boundaries.) What makes agreement relational is the syntax.

The notion 'long-distance agreement' generally is restricted to instances of (iii), namely predicate-argument agreement, when predicate and argument belong to different domains.

In general, agreement information serves to index referential variables (Wunderlich 1992/ 1994, Pollard & Sag 1991/1994), and is checked in the process of composition. Agreement features are unified on the common binder of the variables used in the representation of lexical items.

Consider the article-noun relationship, where the article, taking a noun, agrees with it:

(6) The German article form *die* is fsg or pl.

die: $\lambda N\ \iota x^{fsg/pl}\ (N(x)\ \&\ \Pi(x))$, with Π=contextually prominent

Frau: λz^{fsg} WOMAN(z)

die Frau: $\iota x^{fsg/pl\ \cup\ fsg}\ ($WOMAN$(x)\ \&\ \Pi(x))\ \Rightarrow \iota x^{fsg}\ (result\ of\ unification)$

Pronouns agree with their nominal antecedent; consider the bound pronoun *she* in:

(7) Every woman believes that she will win.

$\forall x^{fsg}\ [$WOMAN$(x)\ \rightarrow$ BELIEVE$(x,\ [that\ z^{fsg}\ will\ win])]$ If $z = x$: $\Rightarrow \forall x^{fsg\ \cup\ fsg} = \forall x^{fsg}$

Verbs agree with one of their arguments:

(8) betrachten ('look at'): $\lambda y\ \lambda x^{-2pl}$ LOOK_AT$(x,y) \Rightarrow$ *Subject agreement*

die Kinder: $\iota z^{3pl}\ ($CHILD$(z)\ \&\ \Pi(z))$

die Kinder betrachten es: $\exists y$ LOOK_AT$(\iota z^{-2pl\ \cup\ 3pl}\ ($CHILD$(z)\ \&\ \Pi(z)),\ y) \Rightarrow \iota z^{3pl}$

(9) They do not agree with an argument that is lexically case-marked: Lexical marking of the subject forces both NOM on the object and *object agreement* in German.

*gefalle*n: $\lambda y^{-2pl}\ \lambda x^{[+hr]}$ LIKE(x,y)

die Kinder gefallen ihm: LIKE$(he,\ \iota z^{3pl}($CHILD$(z)\ \&\ \Pi(z)))$ ('He likes the children')

3. Some elementary facts about Hindi-Urdu verbs (Mohanan 1994:103, 85)
Hindi has a mixed system with both DSM and DOM.

(10) NOM-NOM or NOM-ACC (with animates or definite inanimates) in the non-perfect; the
      verb agrees with the subject.
  a. ravii   roTii/   roTii-ko   khaa-eg-aa.
      Ravi.M   bread.F/   the bread-ACC   eat-FUT-M.SG
      'Ravi will eat bread/the bread.'
  b. niinaa   kelaa /   kelaa-ko   khaa-eg-ii.
      Nina.F   banana.M/   the banana-ACC   eat-FUT-F.SG
      'Nina will eat a banana/the banana.'

(11) ERG-NOM in the perfect; the verb agrees with the object.
  a. ravii-ne   roTii   khaa-yii.
      Ravi.M-ERG   bread.F   eat-PERF.F.SG
      'Ravi ate bread.'
  b. niinaa-ne   kelaa   khaa-yaa.
      Nina.F-ERG   banana.M   eat-PERF.-M.SG
      'Nina ate a banana.'

(12) ERG-ACC in the perfect; the verb bears default marking M.SG.
  a. ravii-ne   roTii-ko   khaa-yaa.
      Ravi.M-ERG   the bread.F-ACC   eat-PERF.M.SG
      'Ravi ate the bread.'
  b. niinaa-ne   kelaa-ko   khaa-yaa.
      Nina.F-ERG   the banana.M-ACC   eat-PERF.MSG
      'Nina ate the banana.'

(13) Ditransitive verbs never show DOM; the indirect object (the recipient) is always
      marked with ACC, and the direct object (the theme) with NOM.
  a. ilaa   ma̦a̦-ko   rotii/   *rotii-ko   de-tii   hai.
      Ila.NOM   mother-ACC   bread.F.NOM/   *bread-ACC   give-IPF.F   be.PRES
      'Ila gives a/the bread to the mother'

  b. ilaa-ne   ma̦a̦-ko   rotii /   *rotii-ko   di-yii.
      Ila-ERG   mother-ACC   bread.F.NOM /   *bread-ACC   give-PERF.F
      'Ila gave a/the bread to the mother'

4. Determining case.
MAX(F):   Every feature [αF] in the input has a correspondent in the output, with α={+,–}.
*(+F):   Avoid [+F] in the output.

(14) ACC-NOM system: MAX(+hr), *(+lr)   »   MAX(+lr) *(+hr) (Stiebels 2000)
      ERG-ABS system: MAX(+lr), *(+hr)   »   MAX(+hr), *(+lr)
      Mixed system: MAX(+hr), MAX(+lr) »   *(+hr), *(+lr)

Constraints that further restrict possible case patterns (Stiebels 2000, Wdl & Lakämper 2001):
MAX(lexF):   Every lexically assigned feature in the input has a correspondent in the
               output. ("Lexical features have priority.")
DEFAULT: Every linking domain displays the default linker (nominative). ("Linking is
               parsimonious.")
UNIQUENESS: Each linker applies only once in a domain. ("Linkers are distinguishable.")
MAX(+hr,+lr): Every feature combination [+hr,+lr] in the input has a correspondent in the
               output. ("Realizing DAT is more important than realizing ERG or ACC.")

(15) Constraintranking in German and Icelandic (Wunderlich 2003):
  MAX(lexF) » DEF, UNIQ » MAX(+hr,+lr) » MAX(+hr), *(+lr) » MAX(+lr), *(+hr)

In Hindi, the role of default vs. marked case is slightly different. Case = mCase ∪ default, with mCase = {ERG, ACC, INSTR, LOC, …}. Moreover, note that DAT and ACC are indistinguishable.

(16) Hindi adaptation of general constraints
  a. UNIQUENESS(mCase): Each marked case applies only once in a domain.
  b. MAX(+hr)/[+lr]: Every feature combination [+hr,+lr] in the input has a correspondent [+hr] in the output.
  c. DEFAULT/non[+lr]: Every linking domain displays either ERG (=[+lr]) or the default linker (NOM).

(17) Constraintranking in Hindi (Wunderlich 2000):
  MAX(lexF), DEF, UNIQ, MAX(+hr)/+lr » MAX(+hr), MAX(+lr) » *(+hr), *(+lr)

This ranking is somewhat flatter than that in German/Icelandic, but possibly only because the set of data considered wasn't complete enough.

(18) Hindi shows two kinds of split.
*[+lr]/–perf: ERG does not appear in a non-perfect context (unless ERG is lexically required).
*[+hr]/LowS: ACC does not appear on low-salient (indefinite inanimate) items.
Presumably, these constraints are oversimplified.
    *[+lr]/–control: ERG does not appear with verbs referring to non-volitionary (i.e. uncontrollable) events.
Of course, these contextual restrictions must be to the left of " » MAX(+hr), MAX(+lr)".
*Nothing in the determination of the optimal case patterns depends on agreement. Thus, **case is independent of agreement.***

Note, by the way, that the combined effects of Hindi's UNIQ and MAX(+hr)/+lr guarantees that in the ditransitive verb, the recipient is always marked ACC, and the theme argument always NOM, independent of their respective salience values.

5. Determining Agreement
In a Hindi transitive verb, all four case patterns ERG-ACC, ERG-NOM (both in the perfect) and NOM-ACC, NOM-NOM (both in the non-perfect) are possible. Verb agreement is governed by the descriptive rule

(19) The verb agrees with the highest argument associated with NOM case.
    (Gair & Wali 1989, Mohanan 1994:105)
Thus, agreement seems to depend on case.

Hindi verb forms bear information about gender (m,f) and/or number (sg,pl), which is associated with one of the NOM arguments.

(20) AGR(-hr)! Agree with the highest argument.
    AGR(-lr)! Agree with the lowest argument.
    AGR(-hr)! » AGR(-lr)! accounts for the fact that the verb agrees with the *highest* NOM argument.
    *multipleAGR. The Verb agrees at most with one of its arguments.

If, as in the ERG-ACC pattern, no NOM argument is present, the verb bears the default information [-f,-pl] (M.SG) (which by itself is ambiguous: agreement or not).

In order to block agreement with overt case (ERG, ACC, INSTR, LOC), one can hypothesize a general restriction such as (21a). Than, in order to determine agreement, one had first to consider the distribution of case (21b).

(21) a. *AGR(mCase). Do not agree with an entity that bears a marked case (ERG, ...).
      b. *AGR(mCase) » AGR(-hr)! » AGR(-lr)!

Kutchi Gujarati lacks ERG case, but in the perfect, the verb still agrees with the object rather than the subject (Patel 2007). That suggests a constraint such as (22a). Moreover, the verb can agree with an ACC-marked object in that dialect, which suggests that (22b) could be ordered high or low.

(22) a. *AGR(-hr)/+perf: Do not agree with the highest argument in the perfect.
      b. *AGR(-lr)/HighS: Do not agree with the lowest argument if it is high-salient.

Similarly, one might think at blocking agreement with lexically marked arguments separately, namely by a constraint such as *AGR(lexF). These lexical markings, however, override [-hr] or [-lr], so that the relevant agreement features are lacking anyway.

(23) *AGR(-hr)/+perf, *AGR(-lr)/HighS » AGR(-hr)! » AGR(-lr)!

Thus, one arrives at a description of Hindi agreement, which does not recurs on case. Of course, the following restrictions are still complementary to each other

(24) *[+lr]/–perf ≈ *AGR(-hr)/+perf
      *[+hr]/LowS ≈ *AGR(-lr)/HighS

(25) Evaluation of Hindi agreement in a single transitive verb

a. Subject agreement in the non-perfect

	*AGR(-hr)/+perf	AGR(-hr)!	AGR(-lr)!
☞ λy λx$^\phi$ verb			*
λy$^\phi$ λx verb		*!	
λy λx verb		*!	*

b. Object agreement in the perfect

	*AGR(-hr)/+perf	AGR(-hr)!	AGR(-lr)!
λy λx$^\phi$ verb	*!		*
☞ λy$^\phi$ λx verb		*	
λy λx verb		*	*!

c. No agreement in the perfect with a high-salient object

	*AGR(-hr)/+perf	*AGR(-lr)/HighS	AGR(-hr)!	AGR(-lr)!
λy λx$^\phi$ verb	*!			*
λy$^\phi$ λx verb		*	*!	
☞ λy λx verb			*	*

6. Microtypology of Indo-Iryan languages. The variation among the Indo-Arayan languages (s. Masica 1991) can best be described by parameters corresp. to different constraint rankings. (For the sake of communication, I use ERG=[+lr] and ACC=[+hr] for the cases, and S=[-hr] and O=[-lr] for the agreement targets.)

Does the language has ergative?
    Yes: Max(ERG) ≫ *(ERG)   (Hindi, ...)
    No: *(ERG) ≫ Max(ERG)   (Bengali, Oriya, Sinhalese)
Does the language has ergative split?
    Yes: *(ERG)/C ≫ Max(ERG) ≫ *(ERG)   (Hindi, ...)
    No: Max(ERG) ≫ *(ERG)   (Assamese, Bishnapriya)

Is the ergative restricted to perfect?
    Yes: *(ERG)/–perf ≫ Max(ERG) ≫ *(ERG)   (Hindi, ...)
    No: (ergative extends to other aspect-tenses)   (Nepali)
Can ergative and accusative cooccur?
    Yes: (Hindi, ...)
    No: *(ERG & ACC) ≫ Max(ERG), Max(ACC)   (Kashmiri)
Does the verb agree with a ERG-marked subject?
    Yes: Agr(S)   (Assamese, Nepali[#], Shina)
    No: *Agr(mS) ≫ Agr(S)   (Hindi, ...)
      # The Nepali verb agrees with an ERG-subj but not with a (lexically marked) DAT-subj (Bickel & Y. 2000).
Does the verb agree with a marked object (accusative)?
    Yes: Agr(O)   (Gujarati, Marwari)
    No: *Agr(mO) ≫ Agr(O)   (Hindi, ...)
Does the verb agree with NOM-subject in the perfect?
    Yes: Agr(S)   (Bengali, Oriya, Sinhalese)
    No: *Agr(S)/perf ≫ Agr(S)   (Kutchi Gujarati, Marwari, Marathi 1/2sg)
Does the infinitive agree with the object?
    Yes: Agr(O)   (Kashmiri, ...)
    No: *Agr(O)/inf ≫ Agr(O)   (Hindi, ...)
Does the matrix verb agree with the subject, and, simultaneously, the embedded verb with the
    object? Yes: (Kashmiri), No: (Hindi)   etc. etc.

7. LDA in Hindi-Urdu

Let us consider the possible constructions with the control verb *caah-* 'want' (Wunderlich 1994). (The acceptability rates are due to a study by Bornkessel-Schlesewsky (2008); for (26a) they however tested with a masc. subject, so that all verbs ended there in masc –aa.)

(26)   In the perfect, three options are available.
    a. All verbs agree with the embedded object (=**LDA**) 88 %
        Raam-**ne**   roTii   [khaa-nii   caah-ii   thii]
        Ram-ERG   bread.F.SG   eat-INF.F.SG   want-PF.F   be.F.SG
        'Ram had wanted to eat bread.'
    b. The object has contrastive focus, and (therefore?) no verb agrees (**default**). 78%
        Niinaa-ne   roTii   [khaa-naa   caah-aa   thaa]
        Nina-ERG   bread.F.SG   eat-INF.M.SG   want-PF.M   be.M.SG
        'It's bread which Nina had wanted to eat.'
    c. The object is definite (topic), and therefore marked with ACC. Consequently, no verb
        agrees (**default**).
        Niinaa-**ne**   us roTii-**ko**   [khaa-naa   caah-aa   thaa]
        Nina-ERG   that bread.F.SG-ACC   eat-INF.M.SG   want-PF.M be.M.SG
        'Nina had wanted to eat that bread.'

Verb complex (VC) formation is necessary. To be marked for the referential property of an embedded object, the verb 'want' must have access to that object, which is possible by embedding v (rather than vp, ip or cp). By functional composition, the object of the embedded verb is inherited to the verb complex.   This is a 'direct agree' rather than a 'cyclic agree' account (in the sense of Lahne 2008) – verbs do not agree with verbs; they collectively agree with a certain argument. In principle it is possible for LDA that the dependent inf. does not agree at al. (Nepali; Bickel & Yadawa 2000).

(27)   'eat'   *khaa-nii* [+inf]: $\lambda z^{F.SG} \lambda y$ EAT(y,z)
        'want'   *caah-ii* [+perf]: $^{F} \lambda v$[+inf] λx WANT(x, v(x))
        'want to eat'   *khaa-nii caah-ii*: $\lambda z^{F.SG} \lambda x$ WANT(x, EAT(x,z))

The assumption of v-embedding (verb complex formation) is compatible with (26b,c).

(28)   In the imperfect, two options are available.
    a. The matrix verb agrees with the subject, while the dependent verb bears default. 95%
        Niinaa [roTii   khaa-naa] [caah-tii   thii]
        Nina.F bread.F.SG   eat-INF.M   want-IPF.F   be.F.SG
        'Nina habitually wanted to eat bread.'
    b. The matrix verb agrees with the subject, while the dependent verb agrees with its
        object. Not all speakers accept this option, some consider it an incorporation construc-
        tion, which is possible with the expected bread-eating, but not with fish-eating. Bickel
        & Yadawa (2000) find dialectal variation. 67%
        Raam [roTii   khaa-nii]   [caah-taa   thaa]
        Ram   bread.F.SG eat-INF.F.SG   want-IPF.M be.M.SG
        'Ram habitually wanted to bread-eat.'

None of these options is compatible with v-embedding. A non-perfect control verb rather
seems to require vp-embedding. This split could be expressed by

(29)   *VC/non-perfect » VC!   *VC=verb complex*
        where VC is a verb complex triggered by a control verb.

Alternatively, perfect is the triggering element for VC formation: in the perfect, object
agreement is better than default marking (see 26a,b). If the embedded object is chosen as
agreement target, a VC must be formed.

Above, we saw the following alternation in the perfect:
    (A)   LDA: S-ne O VC-agrees-with-O
    (B)   no LDA: S-ne O VC-default
What determines that choice? Different authors propose different things. Possibly dialectal
differences.

According to Wunderlich (1994:24), (B) is used only with contrastive focus.

(30)   *According to* Chandra (2005:12), LDA is possible only with *specific objects*.
a.   Usne   kursii   tor-nii   chaah-ii
     he.ERG   chair.F   break-INF.F   want-PF.F
     'He wanted to break the chair.'   (LDA: specific)
b   Usne   kursii   tor-naa   chaah-aa
     he.ERG   chair.F   break-INF.M   want-PF.M
     'He wanted to break chair.'   (no LDA: generic)

According to Bhatt (2005:799), embedded object quantifiers can get scope over the matrix
verb in the LDA option. That interpretation is not possible with the default option (B).

(31)   Quantifier scope
a.   Naim-ne   har   kitaab   parh-nii   chaah-ii   thii
     Naim-ERG   every   book.F   read-INF.F want-PF.F   be.PAST.F.SG
     'Naim wanted to read every book.'   (LDA: ∀ > *want; want* > ∀)
b.   Naim-ne   har   kitaab   parh-naa   chaah-aa   thaa
     Naim-ERG   every   book.F   read-INF.M want-PF.M   be.PAST.M.SG
     'Naim wanted to read every book.'   (no LDA: *∀ > *want; want* > ∀)

These data suggest that VC formation is even more restricted than assumed before: VC (by
which the matrix verb and the embedded object become clause-mates) only happens together
with LDA; LDA →VC is strengthened to LDA ↔ VC.
    Constraints such as (29) *VC/non-perfect » VC! are not helpful to explain why the
reading *∀ > *want* is blocked in the no-LDA-construction (31b).

Let us instead assume the following constraints, in addition to the preceding ones (20, 22):

(32)　*AGR(-hr)/inf. The infinitive does not agree with the highest argument. (High-ranked!)
　　　*AGR(-lr)/inf. The infinitive does not agree with the lowest argument.
　　　*VC. Avoid a verb complex (in which v rather than vp is embedded).

(33)　　*AGR(-hr)/inf ≫ AGR(-hr)! ≫ AGR(-lr)!, *AGR(-lr)/inf ≫ *VC

The evaluation of the embedded infinitive is given on the left side of the following tableau; both agreeing and non-agreeing with the object are likewise possible options. The evaluation of the matrix verb, embedding either the bare or the vp infinitive, is given on the right side. The optimal combinations are (i) object marking on both the embedded and the matrix verb in the first option of the embedded verb (that is, LDA), while (ii) individual assignment of the default marking in the second option. This result is compatible with the data of quantifier scope. Note particularly that VC formation is only forced to avoid an AGR(-lr)!-violation on the matrix verb. One doesn't need any specific requirement of VC.

(34)　　Stepwise evaluation of agreement marking in the control-verb construction of Hindi

embedded verb	*AGR (-hr)/inf	AGR (-lr)!	*AGR (-lr)/inf	matrix verb	AGR (-lr)!	*AGR (-lr)/inf	*VC
$\lambda y\, \lambda z^\phi$ inf	*!						
☞ $\lambda y^\phi\, \lambda z$ inf			*	☞ $\lambda y^\phi\, \lambda x$ (inf, perf)			*
				$\lambda x\, (\lambda y^\phi$ inf) perf	*!		
☞ $\lambda y\, \lambda z$ inf		*		$\lambda y\, \lambda x$ (inf, perf)	*		*!
				☞ $\lambda x\, (\lambda y$ inf) perf	*		

Thus, under the present constraint-ranking, object-agreement of the embedded infinitive survives only in a LDA-construction. Note that the rank of constraints is subject to probabilistic shifts. If two constraints are co-ranked, it is easily possible that one of them will dominate the other. If *AGR(-lr)/inf becomes dominant, LDA is excluded. (Due to lack of data, I do not know in which of the Indo-Aryan languages this could have happened.) If AGR(-lr)! becomes dominant for infinitives, the second option with default agreement is excluded. This, in fact, has happened in Kashmiri (and presumably also in some Hindi dialects).

(35)　　Kashmiri: the inf. agrees with the object in the imperfect (a), the perfect LDA (b) and non-LDA (c), as well as in the gerund clause with GEN-subject (d) (Chandra 2007:102)
a.　　Raam　　chu　　　　yatsaan　koori　　vuchini.
　　　Ram.NOM　be-PRES.M.SG　want.IPF　girl.F.PL　see.INF.F.PL
　　　'Ram wants to see the girls'
b　　raam-an　　yat-I　　　koori　vuchini.
　　　Ram-ERG　wanted.F.PL　girls　see.INF.F.PL
　　　'Ram wanted to see the girls'
c.　　Raam-an　　chu　　　　hameeSI　yotshImut　　[koori　vuchini].
　　　Ram-ERG　be.PRES.M.SG　always　　wanted.M.SG　girls　see-INF.F.PL
　　　'Ram has always wanted [to see girls].'
d.　　[mohnIn　　　yi philim　　　vuchIn]　　chu　　　　　mumkin.
　　　　Mohan-GEN.SG　this film.F.SG　see.INF.F.SG　be.PRES.M.SG possible.
　　　'[Mohan's seeing this film] is possible.'

(36)　　The LDA constr. of Hindi is recursive (although marginal) (Chandra 2005:10)
　　　John-ne　　　roTi-yaan　khaa-nii　　bhul-nii　　　chaah-ii.
　　　John.M-ERG　bread-PL.F　eat-INF.PL.F　forget-INF.PL.F　want-PF.PL.F
　　　'John wanted to forget to eat bread.'

LDA in Hindi is not restricted to control verbs such as 'want', 'forget'. The raising verb *lag-* 'seem' (which is a raising verb only when it selects perfect) also shows LDA (a). The interaction with quantifier scope shows that ∀ > *seem* is a possible reading, for which VC formation is necessary (b). Moreover, it is also possible (although not usual) to embed a control verb as in (c).

(37) LDA with 'seem' in Hindi
a.   Raam-ne   roTii   khaa-yii   lag-tii   thii.
      Ram.M-ERG bread.FSG eat-PERF.F seem-IPF.F be.PAST.F.SG
      'Ram seemed to have eaten bread.'
b.   Raam-ne   har   kitaab   parh-ii   lag-tii   hai
      Ram.M-ERG every book.F read-PF.F seem-IPF.F be.PRES
      'Ram seems to have read every book.'
c.   Raam ne   har   kitaab   parh-nii   chaah-ii   lag-tii   hai
      Ram.M-ERG every book.F read-INF.F want-PERF.F seem-IPF.F be.PRES
      'Ram seemed to have wanted to read every book.'

(38) LDA is also possible with a raising-to-object verb.
a.   Raam-ne   larki   aa-tii   dekh-ii   thii
      Ram-ERG girl.F come-IPF.F see-PERF.F be.PAST.F.SG
      'Ram had seen a girl coming.'
b.   Raam-ne   larki   aa-tii   dekh-nii   chaah-ii
      Ram-ERG girl.F come-IPF.F see-INF.F want-PERF.F
      'Ram wanted to see a girl coming.'

8. LDA in other Indo-Aryan Languages

My account of LDA in Hindi is compatible with the (few) instances of LDA in other Indo-Aryan languages I have knowledge of. Further work is necessary.

  Kashmiri (Chandra 2005/2007), Kutchi Gujarati (Grosz & Patel 2006)

LDA in KGuj is documented only with one item: the verb *par* 'have to', which lexically assigns 'dative' (ACC?) to its subject. The verb does not agree with a lexically marked subject, so it can only agree with an object – otherwise it bears the default information *neuter* (38c). With an embedded transitive verb, LDA occurs both in the imperfect and in the perfect, and no other options are available.

(39) LDA in Kutchi Gujarati (Grosz & Patel 2006)
a.   Reena-ne   Valji-ne   jo-vo   par-e-ro.
      Reena.F-ACC Valji.M-ACC watch-INF.M have.to-3SG-M
      'Reena has to watch Valji.'
b.   Reena-ne   Valji-ne   jo-vo   par-yo.
      Reena.F-ACC Valji.M-ACC watch-INF.M have.to-PERF.M
      'Reena had to watch Valji.'
c.   Khimji-ne   kha-vu   par-yu.
      Khimji.M-ACC eat-INF.N[D] have.to-PERF.N[D]
      'Khimji had to eat it.'

These facts follow from the constraint ranking:

*AGR(-hr)/perf, *AGR(-hr)/lexCase, *AGR(-hr)/inf ≫ AGR(-hr)! ≫ **AGR(-lr)!** ≫ *AGR(-lr)/inf

which makes for both verbs object agreement to be optimal; in the matrix verb, of course, only if VC is possible.

9. Resumé

The data showing noncomplementary of case and agreement suggest that case and agreement in general are independent of each other.

	subject agreement	object agreement
NOM-ACC	✔	✔
ERG-ABS	✔	✔

▷ Therefore, the description of case should be independent of agreement, and the description of agreement should be independent of case. It is only at the first glance that Hindi agreement depends on case. A more informed investigation shows that both agreement and case depend on aspect (perfect or not), as far as the subject is concerned, and on salience features (such as definiteness) of the object.

▷ The Hindi infinitive does not agree with its object unless it is involved in LDA. This surprising fact (although subject to variation) is explained by the set of constraints proposed for Hindi agreement.

▷ LDA is the outcome of a specific constraint interaction. There is no requirement to achieve this construction.

▷ All examples of Hindi LDA (and so far as I can see, in other Indo-Aryan languages, too) require verb complex (VC) formation, enabled by a small set of (so-called restructuring) matrix verbs. No particular rule for VC formation must be assumed; it is just a way to achieve a construction that optimally fits the specific constraint ranking, thereby establishing LDA.

▷ Agreement, thus, forces in which type of construction it occurs, that is certainly *not* post-syntactic.

References:

Bhatt, Rajesh. 2005. Long-distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23: 757-807.

Bickel, Balthasar & Yogendra P. Yādava. 2000. A fresh look at grammatical relations in Indo-Aryan. *Lingua* 110: 343-373.

Bierwisch, Manfred. 1967. Syntactic features in morphology. General problems of so-called prono-minal inflection in German. In: *To Honor Roman Jacobson*, 239-270. Mouton: The Hague/ Paris.

Boeckx, Cedric. 2004. Long-distance agreement in Hindi: Some theoretical implications. *Studia Linguistica* 58:23-36.

Bornkessel-Schlesewsky, Ina et al. 2008. Bridging the gap between processing preferences and typological distributions: Initial evidence from the online comprehension of control constructions in Hindi. In Marc Richards & Andrej L. Malchukov (eds.) *Scales* (= Ling. Arbeits Berichte 86), 397-436. Institut für Linguistik, Un.iv. of Leipzig.

Chandra, Pritha. 2005. Long-distance agreement in Hindi-Urdu: Agree, AGREE or Spec-Head? Ms., Univ. of Maryland, College Park.

Chandra, Pritha. 2007. (Dis)Agree: Agreement and movement reconsidered. Ph.D dissertation. Univ. of Maryland.

Gair, James W. & Kashi Wali. 1989. Hindi agreement as anaphor. *Linguistics* 27: 45-70.

Grosz, Patrick & Pritty Patel. 2006. Long-distance agreement and restructuring predicates in Kutchi Gujarati. Ms., University of Vienna, University College London.

Kiparsky, Paul. 1992. Structural case. Berlin: Institute for Advanced Study. Unpubl. Ms.

Lahne, Antje. 2008. Specificity-driven syntactic derivation: A new view on long-distance agreement. Ms., Univ. of Leipzig.

Lehmann, Christian. 1982. Universal and typological aspects of agreement. In Hansjakob Seiler & Franz J. Stochowiak (eds.) *Apprehension. Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*, 201-267. Tübingen: Narr.

Magier, David S. 1983. Components of Ergativity in Marwari. *Chicago Linguistic Society (CLS)* 19: 244–255.

Mahajan, Anoop K. 1989. Agreement and Agreement Phrases. *MIT Working Papers in Linguistics* 10, 217-252. Cambridge, Mass.

Mahanan, Tara. 1994. *Argument structure in Hindi*. Stanford: CSLI Publ.

Marantz, Alec (1991): 'Case and Licensing', in: Westphal, German / Ao, Benjamin / Chae, Hee-Rahk (eds.): *Proceedings of ESCOL '91*. Ithaca, New York: Cornell Linguistics Club, 234-253.

Marantz, Alec. 2000. Case and Licensing. In Eric Reuland (ed.) *Arguments and Case: Explaining Burzio's Generalization*. Amsterdam: John Benjamins.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge Univ. Press.

Patel, Pritty. 2007. *Split ergativity and the nature of agreement*. University College London.

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

Pollard, Carl & Ivan Sag. 1991/1994. Information-based syntax and semantics, vol.2. draft 1991. *Head-driven phrase structure grammar*. Univ. of Chicago Press 1994.

Stiebels, Barbara. 2000. Linker inventories, linking splits and lexical economy. In B. Stiebels & D. Wunderlich (eds.) *Lexicon in Focus*, 211- 245. Berlin: Akademie Verlag.

Wali, Kashi. 2005. *Marathi*. Munich: LINCOM Europa.

Wunderlich, Dieter. 1992/1994. Towards a lexicon-based theory of agreement. Working papers SFB 282, No.20, Univ. of Düsseldorf 1992. – *Theoretical Linguistics* 20 (1994) 1- 35.

Wunderlich, Dieter. 1992/1997. CAUSE and the structure of verbs. Working papers SFB 282, No.36, Univ. of Düsseldorf 1992. – *Linguistic Inquiry* 28 (1997) 27-68.

Wunderlich, Dieter. 2000. Optimal case in Hindi. Ms., Univ. of Düsseldorf.

Wunderlich, Dieter. 2001. Grammatical Agreement. In N.J. Smelser and Paul B. Bates (eds.) *2001 International Encyclopedia of Social and Behavioral Sciences*, 6330-6334. Oxford: Pergamon.

Wunderlich, Dieter. 2003. Optimal case patterns: German and Icelandic compared. In Ellen Brandner & Heike Zinsmeister (eds.) *New perspectives on case theory*, 331-367. Stabford: CSLI Publ.

Wunderlich, Dieter & Renate Lakämper. 2001. On the interaction of structural and semantic case. *Lingua* 111: 377-418.