1. The scope of the discussion

This vignette is concerned with agreement in \(\varphi\)-features (person, number, and gender/noun-class)—henceforth, \(\varphi\)-agreement. For our current purposes, \(\varphi\)-agreement should be understood to refer to morpho-phonologically detectable covariance in \(\varphi\)-features between a finite verb and one or more nominal arguments of that verb. An example is given in (1):

(1) a. ha-kelev navax
    art-dog bark.pst(masc)
    ‘The dog barked.’

b. ha-klav-im navx-u
    art-dog-pl bark.pst-3pl
    ‘The dogs barked.’

Finite verb agreement is far from the only case of obligatory covariance between pairs of expressions in natural language, and it is not even the only case where the \(\varphi\)-features of two or more expressions appear to covary. (Anaphoric binding, for example, typically involves matching of \(\varphi\)-features between a binder and a bindee, as well.) However, whether these other phenomena do or do not supervene on the same grammatical relation responsible for \(\varphi\)-agreement in the narrow sense is an empirical question.\(^1\) The current contribution will therefore focus on \(\varphi\)-agreement in the narrow sense: its modular locus (section 2), its computational properties (sections 3–4), its interactions with nominal case (section 5), and the nature of the feature-structures it manipulates (section 6).

2. The modular locus of \(\varphi\)-feature agreement

There has been considerable debate in the recent generative literature regarding the modular locus of \(\varphi\)-agreement. Bobaljik (2008) puts forth an argument that \(\varphi\)-agreement is an exclusively postsyntactic phenomenon. This argument is based on the premise that Marantz’s (1991) “m-case” calculus applies postsyntactically. That premise, in turn, is based on the assumption that m-case never feeds any operations that apply in syntax proper. However, once we have evidence disconfirming this premise (i.e., evidence that m-case does in fact feed certain syntactic operations; Preminger 2014:157–209), Bobaljik’s argument is turned on its head, instead showing that the computation of \(\varphi\)-agreement must be at least partially situated within syntax (idem).\(^2\)

\(^*\)Thanks to the editors Kleanthes Grohmann and Evelina Leivada, as well as an anonymous reviewer, for guidance and feedback on this text. All errors and misrepresentations are my own.

\(^1\)There is reason to believe that the answer to this question is negative—i.e., that a reduction of these other phenomena to the same grammatical relation which underpins \(\varphi\)-agreement is misguided. See Preminger (2013, 2019a), Preminger & Polinsky (2015) and Rudnev (2020), as well as section 3, for discussion.

\(^2\)Landau (2016) has recently attempted to argue in favor of a conclusion similar to Bobaljik’s (2008), based on the behavior of Partial Control. But the argument is based on: (i) a theory of Control that is not uncontroversial (cf. Boeckx & Hornstein 2007, Boeckx, Hornstein & Nunes 2010, Hornstein 1999, but also Grano 2015, McFadden & Sun-
Other authors have advanced the claim that agreement is split into two components, one situated within syntax, and the other situated postsyntactically on the PF branch. Relevant work includes Benmamoun, Bhatia & Polinsky (2009) and Arregi & Nevins (2012), as well as Bhatt & Walkow (2013), who argue that the second component can but does not have to be postsyntactic. Bhatt & Walkow’s arguments, however, are based on agreement with coordinations (as are those of Benmamoun, Bhatia & Polinsky), and there are reasons to believe that agreement with coordinations reflects an extra-grammatical computation in the first place (Lyskawa 2021). This suggests that the position advocated by Arregi & Nevins (as well as Benmamoun, Bhatia & Polinsky) is correct—namely, that what we have descriptively identified as $\varphi$-agreement consists of two components, one syntactic, and the other postsyntactic. Following Arregi & Nevins (2012), I will refer to these as Agree-Link and Agree-Copy, respectively. Agree-Link consists of a probe (a head like $\text{Infl}^0$ or $\nu^0$) searching for and establishing a relation with a goal—typically, an extended nominal projection (i.e., a phrase of category DP)—or with multiple goals (Anagnostopoulou 2003, Hiraiwa 2001, 2004). Agree-Copy is the process by which content from the goal or goals (be it abstract morphosyntactic content, or concrete morpho-phonological content; cf. Dobrin 1995) is shared with, or copied onto, the probe.

3. Minimality and directionality

Agree-Link is subject to minimality:

(2) If $H$ is a probe seeking feature $f$; $H \gg XP \gg YP$ (where ‘$\gg$’ denotes asymmetric c-command); and XP and YP are both $f$-bearers:

\[
\begin{array}{c}
H \ldots \text{XP} \ldots \text{YP} \\
\end{array}
\]

This arises due to the way in which probes search for their goals, as formalized in (3) (see Preminger 2019b:23–24 for further discussion):

(3) Iterative Downward Search
a. Let $P$ be a syntactic probe, and let XP be $P$’s sister
b. query: Is XP a viable goal? If so, halt with “XP” as the search result
c. For every specifier $ZP$ of XP, query: Is $ZP$ a viable goal? If so, halt with “ZP” as the search result
d. query: Is XP a phase? If so, halt with no goal
e. query: Does $X^0$ have a complement? If not, halt with no goal
f. Return to step (b), using the constituent in [Compl,X] as the new “XP”

It is worth noting, in relation to (2) and (3), that it is far from clear that grammatical relations other than $\varphi$-agreement share the same directionality (viz. a probe searching downwards in the structure for a feature-matched goal). This is relevant because recent years have seen attempts to shoehorn daresan 2018, Sundaresan 2014, a.o.); and, more problematically, (ii) the assumption that the transfer of $\varphi$-features from binder to bindee reduces to the same operation that underlies $\varphi$-agreement in the narrow sense (see Preminger 2019a and Rudnev 2020 on why such a reduction is probably incorrect). Since our goal here is to discuss $\varphi$-agreement itself, not putative reductions of other phenomena to the same mechanism, we can safely set aside Landau’s argument here.
Φ-agreement into a structurally inverted schema, where probes search upwards (Bjorkman & Zeijlstra 2019, Wurmbbrand 2011, a.o.). Several notes are in order in this regard. First, the central evidence motivating these proposals comes not from Φ-agreement (in the sense defined here) but from other empirical domains (but see Wurmbbrand & Haddad 2016 for a possible exception). For the most part, this literature puts forth no Φ-agreement data that cannot also be handled in terms of (2–3) (see Polinsky & Preminger 2019:10–11 for discussion).

Second, these proposals are beset by a series of problems when it comes to long-distance agreement (LDA). Bjorkman & Zeijlstra’s (2019) account of LDA in “substandard” Basque, for example, faces at least the following problems: (i) it appeals to uninterpretable Case features (untenable on Basque-internal grounds; Preminger 2011:929–930); (ii) it fails to capture an important distinction between mediated and unmediated LDA (Bhatt & Keine 2017:10–11, Preminger 2009:634); and most importantly, (iii) it requires that Φ-agreement include a checking component, alongside valuation (as argued in Preminger 2014, this assumption is incompatible with an adequate theory of agreement; see also section 4).

Third, such proposals ignore the large body of work showing that (2–3) is absolutely necessary for an adequate account of the Person Case Constraint (PCC; Anagnostopoulou 2003, 2005, Béjar & Rezac 2003, Coon & Keine 2021, Nevins 2007, a.o.).

Lastly, such proposals are incompatible with the crosslinguistic typology of cross-clausal agreement relations (Polinsky & Preminger 2019).

4. Agreement failures and the inadequacy of “interface conditions”

A pervasive property of Φ-agreement is its obligatoriness: sentences in which Φ-agreement has not applied are often ill-formed. The original minimalist account of this obligatoriness attributes it to interface conditions. Probes that have not established an Agree(-Link) relation with an appropriate goal are hypothesized to bear “uninterpretable” features, which are deemed illicit at the interface with LF and/or PF, resulting in ungrammaticality (Chomsky 2000, 2001, a.o.). The motivation behind this type of account is an attempt to comply with Chomsky’s (2004:106) Strong Minimalist Thesis (SMT), which states that the only linguistically-proprietary mechanism of syntax is Merge.

This account fails, however. Consider cases of omnivorous agreement (Nevins 2011), as seen for example in the Agent-Focus construction of the K’ichean (Mayan) languages (Dayley 1978, 1985, Mondloch 1981, Norman & Campbell 1978, Smith-Stark 1978, a.o.):

(4) a. ja rje’ x-e/*∅-tz’et-o rja’
   foc them com-3pl/#3sg.abs-see-AF him
   ‘They saw him.’
   (Kaqchikel)

b. ja rja’ x-e/*∅-tz’et-o rje’
   foc him com-3pl/#3sg.abs-see-AF them
   ‘He saw them.’

The verb in (4a–b) agrees with the plural argument, regardless of whether the argument in question is the subject or the object of the clause.

This pattern arises because 3rd person singular DPs are not viable targets for finite (absolutive) Φ-agreement probes in Kaqchikel. But Agent-Focus clauses in which both the subject and the object are 3rd person singular are nevertheless well-formed:
This shows that derivations in which a $\varphi$-probe has searched for, but not found, an appropriate goal are not ill-formed at the interfaces (i.e., the result is not a “crash”). For a discussion of alternative analyses of these data, and why they are inadequate, see Preminger (2014:18–20, 67–73, 89, 123–128).

Instead, the picture that emerges is one in which Agree-Link is an obligatory operation: the search schematized in (3) must be initiated whenever a $\varphi$-seeking probe is merged into the structure, but what the system enforces is the initiation of this search procedure—not necessarily its successful culmination. This fundamental shift in the theory of agreement, from an approach based on output conditions to one based on derivational conditions, can already be found in the work of Béjar (2003) and López (2007). More recently, Deal (2015, 2020) and Coon & Keine (2021) offer competing proposals on how this approach may be further refined to capture a broader set of empirical domains and typological patterns.

5. The relation between $\varphi$-agreement and case

Early minimalist accounts of $\varphi$-agreement assumed that a side effect of (successful) $\varphi$-agreement between a head H and a DP was the assignment of Abstract Case (Chomsky 1981, Vergnaud 1977) to the DP, where the particular Abstract Case value assigned depended on the identity of H. (Nom-}

ative for H = T, accusative for H = v, and so on.) To the extent that this theory had anything at all to say about morphosyntactic case, it was assumed to be computed based on these Abstract Case values. (See Legate 2008 for what is perhaps the most worked out version of such a theory.) On this view, then, nominal case would depend on (successful) $\varphi$-agreement.

This view is demonstrably incorrect. As shown by Sigurðsson (1993, 2004, 2006) and Bobaljik (2008), a.o., it is $\varphi$-agreement that depends on nominal case, not the other way around. Specifically, whether a given DP can or cannot be targeted for $\varphi$-agreement depends on which case it already has (if any; see below). The crucial data comes from, among other things, quirky-case phenomena, wherein a grammatical subject is marked not with the case normally associated with its grammatical function (e.g. nominative), but with a lexically-specified case associated with the particular choice of verb. Crucially, when this happens, $\varphi$-agreement tracks case, not grammatical function:

(6) a. * [Mörgum studentum ] líka [verkið ]
   [many students ] .DAT like.3pl [job.the ] .NOM
   ‘Many students like the job.’
   (Icelandic) [Harley 1995:208]

   b. [Jóni ] líkuðu [þessir sokkar ].
   [Jon ] .DAT liked.3pl [these socks ] .NOM
   ‘Jon liked these socks.’
   (Jónsson 1996:149)

The examples in (6a–b) involve a verb that idiosyncratically assigns dative to its grammatical subject, and in this case, agreement is controlled not by the subject, which is dative, but by the object,

---

3More broadly, this means that Chomsky’s (2004:106) Strong Minimalist Thesis (SMT), while still a laudable methodological ideal, is incorrect if taken as a substantive hypothesis about the human capacity for language.
which is nominative. (See Andrews 1976, Sigurðsson 1989, Thráinsson 1979, and Zaenen, Maling & Thráinsson 1985, a.o., for arguments that the dative DPs in (6a–b) are indeed grammatical subjects.)

On the basis of this and other evidence, Bobaljik (2008:303) proposes that agreement tracks Marantz’s (1991) “m-case”—along the lines shown in (7)—modifying an earlier proposal by Moravcsik (1974):

(7) unmarked case (e.g. nom, abs) > dependent case (e.g. acc, erg) > lexical case (e.g. dat)

Without going into all the details, the idea is that \( \varphi \)-probes discriminate among potential nominal targets according to (7), with the most common pattern being a probe that can only target bearers of unmarked case. On a view where unmarked case is actually the total absence of valued case features (Bittner & Hale 1996, Kornfilt & Preminger 2015, McFadden 2018, Sigurðsson 2012, a.o.), this amounts to a probe being able to target a DP only if that DP is caseless.

This bears some resemblance to Chomsky’s (2001) Activity Condition approach, whereby DPs that have been assigned Abstract Case are rendered subsequently inaccessible to \( \varphi \)-probing. But there is a significant difference. On Bobaljik’s approach, \( \varphi \)-agreement with a nominative DP (which would perhaps be more accurately characterized as a caseless DP; see above) does not change the DP’s status vis-à-vis case. Therefore, another \( \varphi \)-probe can subsequently agree with the same DP. This is exactly the opposite of what Chomsky’s approach predicts. And, as shown by Polinsky (2003) and Polinsky & Potsdam (2001), a.o., instances where multiple \( \varphi \)-probes agree with the same DP are in fact well attested.

6. The structure of \( \varphi \)-features

Our discussion of \( \varphi \)-agreement has so far abstracted away from the features themselves. An early view, which persisted from the earliest years of generative linguistics (e.g. Chomsky 1965:175) all through the minimalist canon (e.g. Chomsky 1995:37), is that \( \varphi \)-features have a flat attribute-value structure of the kind exemplified in (8):\footnote{This view is by no means restricted to work in the Government & Binding and minimalist traditions. Cf. Pollard & Sag (1994:76), for an example of the treatment of \( \varphi \)-features within HPSG.}

\[
\begin{align*}
\text{a. } \begin{bmatrix}
\text{pers} & 2 \\
\text{num} & \text{pl}
\end{bmatrix} & \quad \text{b. } \begin{bmatrix}
\text{pers} & 3 \\
\text{num} & \text{sg}
\end{bmatrix}
\end{align*}
\]

More recent work has revealed, however, that this is not an adequate representation for \( \varphi \)-features. Instead, \( \varphi \)-features are privative, and arranged in a feature geometry, along the same lines proposed in phonological theory decades earlier (Archangeli 1988, Clements 1985, a.o.).

The feature-geometric representation for \( \varphi \)-features was originally put forth by Harley & Ritter (2002), in order to account for the shape of pronominal inventories across the world’s languages. A simplified geometry is given in (9), abstracting away from many relevant details (including the treatment of 1st person inclusive vs. exclusive, of number systems beyond singular-plural, and of gender/noun-class systems).
To see how such a geometry works, consider person features first. So-called “3rd person” expressions are those where the root node lacks [participant] (as well as [speaker]); “2nd person” expressions are those where the root node has a [participant] but no further dependents; and “1st person” expressions are those where both dependents are present. Similarly, “singular” expressions are those where the [plural] dependent is missing, and “plural” expressions are those in which it is present.

This proposal, despite originally being geared towards a set of morphological desiderata, has proven extremely successful in accounting for a set of syntactic phenomena relating to $\varphi$-agreement. I will briefly mention two such phenomena here. The first, already discussed in section 4, are cases of omnivorous agreement, where a given probe can agree with the subject or the object depending on the features they each bear. Crosslinguistically, cases of omnivorous agreement (across disparate language families like Algonquian, Mayan, and Kartvelian) seem to consistently target 1st/2nd person arguments and plural ones, skipping over 3rd person arguments and singular ones.5 In other words, omnivorous agreement targets $\varphi$-feature values that can be positively specified along a geometry of the kind proposed by Harley & Ritter (see (9)). This geometry therefore allows the typology of omnivorous agreement to be straightforwardly derived, on the simple assumption that probes are unable to search for the absence of a given feature, only for its presence. In contrast, the same typology is not so easily derived on a theory where 1/2/3 and sg/pl are just different possible values in a flat, attribute-value arrangement such as (8). Nor is it so easily derived on a theory where 2nd and 3rd person constitute a natural class on a par with the natural class formed of 1st and 2nd person (as proposed in Ackema & Neeleman 2018, for example).

The second empirical pattern has to do with $\varphi$-neutralization: what happens to $\varphi$-probes when their relation to a goal has been demonstrably disrupted. Consider the pattern of LDA in “substandard” Basque (Etxepare 2006), already mentioned in section 3:

(10) BASELINE: LDA SUCCESSFUL

[ Miren-entzat [harri horiek]$_{\text{ABS}}$ altxa-tze-n ] probatu dituzte
Miren-BEN stone(s) those$_{\text{pl}}$ lift-NMZ-LOC attempted $\text{aux:3pl.abs}$
‘They have attempted to lift those stones for Miren.’
(subject is pro<3pl.erg>)

(11) $\varphi$-NEUTRALIZATION: LDA DISRUPTED $\rightarrow$ 3SG

[ [Lankide-e-i]$_{\text{DAT}}$ [liburu horiek]$_{\text{ABS}}$ irakur-tze-n ] probatu
dute/*dituzte
$\text{aux:3sg.abs}/*$3pl.abs
‘They have attempted to read those books to the colleagues.’
(subject is pro<3pl.erg>)

$\varphi$-neutralization comes from Menominee (Trommer 2008).
(12) \textbf{UNATTES TED:} \( \varphi \)-\textit{NEUTRALIZATION TO, E.G., 2SG}

\[
* \quad [ [ \text{Lankide-e-i}]_{\text{DAT}} \quad [ \text{liburu horiek}]_{\text{ABS}} \quad \text{irakur-\textit{tze-N}} ] \quad \text{probatu} \quad \text{haute} \\
\text{colleague(s)-ART-pl-DAT \quad book(s) those_pl \quad read-NMZ-LOC \quad attempted \text{aux.2sg.ABS}}
\]

‘They have attempted to read those books to the colleagues.’

(subject is \( \text{pro}<3\text{pl.ERG}> \))

In (10), the finite auxiliary in the matrix clause exhibits absolutive agreement morphology controlled by the embedded absolutive argument. In (11), the same kind of agreement relation is disrupted by the presence of a dative intervener in the embedded clause. Crucially, this results in 3rd person singular absolutive morphology on the matrix auxiliary. What never happens, in Basque LDA or in any other LDA pattern I am aware of, is a disruption of this kind that results in some other agreement morphology (e.g. invariant 2nd person plural, instead of invariant 3rd person singular). Why would this be? The geometry in (9) provides an immediate answer: if things like “3rd person” and “singular” are simply the absence of privative dependents like \{participant\} and \{plural\}, it follows that a \( \varphi \)-probe that has not managed to reach an appropriate goal will not have encountered valued dependents of this sort. The agreement morphology that such a probe would display would therefore correspond to the absence of \{participant\} and \{plural\}, i.e., “3rd person singular.”

\begin{center}
\textbf{References}
\end{center}


