Agreement, grammatical

Dieter Wunderlich (University of Düsseldorf)

Grammatical agreement is a phenomenon in which word forms cooccurring in a clause are sensitive to each other. Inflected forms often agree in their values of number, gender, or person, as can be seen from the contrasting examples in (1a,b) to (4a,b):

(1) French   a. La^{fsg} fille^{fsg} est^{3sg} belle^{fsg}.   b. Les^{pl} garcons^{mpl} sont^{3pl} beaux^{mpl}.
                'The girl is beautiful'   'The boys are handsome'

(2) French   a. Je^{1sg} les^{3pl} ai^{1sg} vues^{fpl}.   b. Tu^{2sg} l'^{3sg} a^{2sg} vu^{msg}.
                'I have seen them'   'You have seen him'

(3) Finnish   a. Taina löysi minun$^{1sg}_{GEN}$ kirja-ni^{1sgPoss}.   b. Taina löysi heidän$^{3pl}_{GEN}$ kirja-nsa^{3Poss}.
                'Taina found my book'   'Taina found their book'

(4) German   a. Ein kleines^{nsg} Mädchen^{n} saß^{3sg} auf der Treppe, und es^{3nsg} lachte.
                'A little girl sat on the stairs, and she laughed'
           b. Ein kleiner^{msg} Junge^{msg} stand^{3sg} daneben, und er^{3msg} heulte.
                'A little boy stood beside, and he cried'

In (1), the definite article agrees with the following noun, and both the copula verb and the predicative adjective agree with the whole subject phrase. If any item of (1a) is exchanged with the corresponding item of (1b), the sentence becomes ungrammatical (e.g., *la fille est beaux*). (2) shows two agreement relations interwined: the auxiliary agrees with the subject pronoun, while the participle agrees with the object clitic. (3) is an example of a possessed noun agreeing with the possessor in the genitive. In (4) it is the pronoun in the second clause that agrees with the antecedent noun phrase in the first clause. Notice that *Mädchen* 'girl' is a neuter noun, triggering a neuter pronoun even though it refers to a female being; in contrast, the English pronouns in the translation are selected solely on semantic grounds.

Grammatical agreement can be seen as an interplay of morphological, syntactic and semantic aspects (Anderson 1988, Lapointe 1988, Corbett 1994, 1995, 1998, Pollard and Sag 1994). Morphosyntactically, agreement is a relation between syntactic constituents in virtue of the fact that the word forms they consist of bear similar information, either inherently or by means of morphological affixation (inflection). Semantically, agreement serves to keep record of discourse referents: only constituents that relate to the same referent may overtly agree with each other (Lehmann 1982). The sentences in (1) are about one referent (a single person or a group of persons), while those in (2) are about two referents in a seeing-relation.

Generally, agreement relations are found among the following elements:

(a)   NP- (or DP-) internally: determiner and noun, attribute and noun, possessor and possessed noun;

(b)   a predicate and its arguments: verb-subject, verb-object, preposition-object, predicative noun/adjective-argument;

(c)   a pronoun or anapher and its antecedent.

The categories that may appear in these relations are gender (e.g., feminine f, masculine m, animate, or neuter n), number (e.g., dual, plural pl, and singular sg), and person (1st, 2nd and 3rd, or inclusive 1+2).

Nouns and pronouns (or pronominal affixes/clitics) function as the controller (or source) of an agreement relation in virtue of their features. Nouns are inherently classified for gender (or noun class), and sometimes also for number (cf. *scissors, trousers, brethren*), whereas pronouns are inherently classified for person and number, and sometimes also for gender. In some instances, this inherent classification is semantically based: as a default, female beings are represented by feminine nouns, but male beings by masculine nouns; aggregates are represented by plural nouns; adressees are represented by 2nd person pronouns, but speakers by 1st person pronouns. In most other cases feature assignment is purely grammatical (consider the distinct genders in French *lune*^{f}, *soleil*^{m} vs. German *Mond*^{m}, *Sonne*^{f} 'moon', 'sun') and may even override the default (cf. German *Mädchen*^{n} 'girl') (Corbett 1991).

The grammatical elements that agree with the noun or the pronoun are said to be the agreement targets (or controllees); in (1), the article *la* is the target of *fille*, which is f, and both the copula *est* and the adjective *belle* are targets of the DP *la fille*, which is fsg. The agreement features of controller and target do not have to be identical, but they must be compatible with each other, as in *la fille*^{fsg} and *est*^{3sg}. In order for a DP to be the controller of an agreement relation, the features of the noun and the other DP-internal elements have to be unified in the formal sense of feature structures (see below).

For short, agreement features (person, number, or gender) are often dubbed as ϕ-features; they index the referential elements of a discourse (pronouns and DPs) as well as the categories that predicate of these elements.

The notion of 'agreement' is often used in a broader sense, including other types of correspondences as well, such as case concord in (5), and sortal correspondence in (6).

(5) a. Latin: DP-internal case concord
        Illa^{fsg}-m$_{ACC}$ femina^{fsg}-m$_{ACC}$ bella^{fsg}-m$_{ACC}$ videbam.
        'I saw that beautiful lady'
    b. Finnish: DP-internal case concord
        tuo-n$_{GEN}$ nariseva-n$_{GEN}$ tuoli-n$_{GEN}$ alla
        that   squeaky   chair   under
        'under that squeaky chair'
    c. German: Case concord between the predicative noun and its argument
        Man nannte ihn$_{ACC}$ einen Idioten$_{ACC}$; er$_{NOM}$ wurde ein Idiot$_{NOM}$ genannt.
        'He was called an idiot'

(6)   The horse neighs; the dog barks; the cock crows.

These relations, however, differ in many respects from agreement as described above and therefore are not included here. Case reflects the grammatical role of a constituent (subject or object) rather than its referential status, whereas sortal correspondence is based on much finer semantic distinctions than those provided by ϕ-features.

Several theories have been proposed to account for agreement phenomena. At a first systematic account, the constituents introduced by a phrase structure rule, such as S → DP + VP, are indexed with the same set of ϕ-feature values:

(7) a. S → DP$^{\phi i}$ + VP$^{\phi i}$ (subject+verb phrase)
  b. DP$^{\phi j}$ → D$^{\phi j}$ + NP$^{\phi j}$ (determiner+noun phrase)

This account on the one hand is too rigid since it does not allow the instantiation of different (still compatible) feature values, as in the French example (1) in which one constituent is fsg and another one is 3sg. Moreover, it does not account for anapher-antecedent agreement, where the elements involved do not originate from a common syntactic node. On the other hand, this account is not specific enough because it would describe case concord with the same means.

Notice that (7) is symmetrical between controller and target. In contrast, Generalized Phrase Structure Grammar (Gazdar et al. 1985) proposed an asymmetric analysis: controller elements can inherit a feature structure AGR lexically, whereas target elements (mostly the functors) receive such a structure only by morphological affixation; the distribution of AGR-features within a clause is then checked by a general control-agreement principle. Again, this theory is both too general and too restricted with respect to agreement.

In the Chomskyan framework of Principles and Parameter Theory, later replaced by Minimalist Syntax (Chomsky 1995), the bearers of agreement features are functional categories represented by independent syntactic projections (Laka 1988, Pollock 1989, Chomsky 1991). Consider the schematic structure in (8). This structure describes subject-verb agreement if the AGR-node is identified with AGRS (in some versions: TENSE, or VOICE), which is specified for certain ϕ-feature values, as well as for case features (such as NOM). In order to be marked for subject agreement morphologically, the verb stem must move to AGRS, while the subject DP moves into SPEC.AGRS, where it checks nominative case (both movements leave a trace, which is coindexed in (8)). Similarly, AGRO is responsible for object-verb agreement as well as accusative case, and AGRPOSS is responsible for possessor-noun agreement as well as genitive case. Regardless of various modifications in detail, the basic idea remained the same: there is one syntactic configuration that determines both case and agreement. Both agreement and case are checked between AGR and SPEC.AGR (and both must be licensed by a theta-role of the verb or noun).
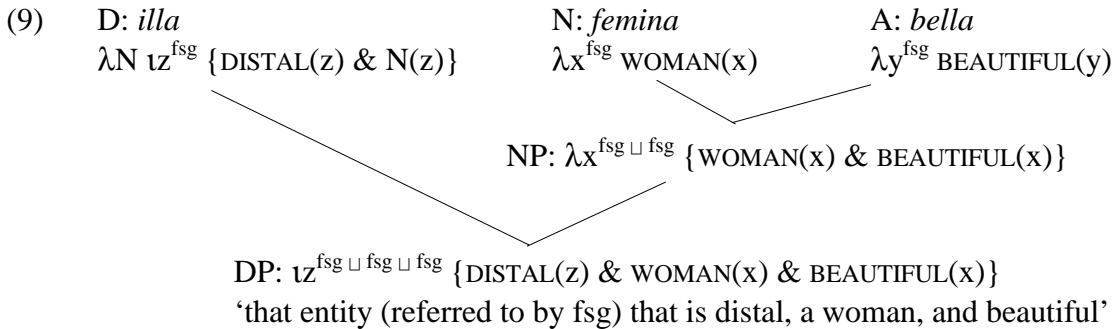
(8)
```
                        AGRP
                    3
         SPEC.AGR         AGR'
         DPᵢ          3
              AGR           VP
            2           2
         AGR     Vⱼ    tᵢ  ...  tⱼ
         φ+case
         features
```

Although such a configuration accounts for agreement relations that go together with argument structure, such as subject-verb, object-verb, and possessor-noun agreement, it is far too rigid for other kinds of agreement. It is unclear how agreement with attributes, adverbs or secondary predi-

cates can be described. Moreover, agreement is here restricted to local or locally transferred phenomena, as in the preceding approaches. The anapher-antecedent relationship certainly is non-local, and there are more instances of non-local agreement (see below) which hardly can be captured by the syntactic account.

Following the idea that ϕ-features keep record in the predication of individual referents, agreement is a mechanism that rests on the normal composition procedures; there is no need for a special syntax of agreement. In the recent literature, two uniform treatments of agreement have been proposed, both locating agreement at the syntax-semantics interface: the account of Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994), and the lexicon-based account (Wunderlich 1994). Within HPSG, all grammatical information of lexical items, as well as those of complex word forms or phrases, is encoded by means of feature structures. The ϕ-features that are shared in an agreement relation (regardless of whether they are lexically inherent or morphologically introduced) form a substructure INDEX (as part of CONTENT). When a clause is composed by unification of feature structures, the relevant INDEX structures are unified as well.

In the lexicon-based theory of agreement, which has become part of Lexical Decomposition Grammar (LDG; Wunderlich 1997), all morphological and phrasal composition is determined by the Theta Structure of lexical items, represented by a list of lambda abstractors related to the Semantic Form (a partial semantic representation). In this theory, ϕ-features are associated with (indexed to) individual variables under a binder, such as the λ-abstractor or the Russellian iota operator 'ɩ'. In the process of composition, these indexes are unified at the common binder of the variables in question. This is illustrated in (9), representing the composition of Latin *illa femina bella* 'that beautiful woman' from three individual entries; unification is represented by the sign '⊔'.

(9)   D: *illa*   N: *femina*   A: *bella*
       $\lambda N\ \iota z^{fsg}$ {DISTAL(z) & N(z)} λx^{fsg} WOMAN(x) λy^{fsg} BEAUTIFUL(y)

                        NP: $\lambda x^{fsg \sqcup fsg}$ {WOMAN(x) & BEAUTIFUL(x)}

              DP: $\iota z^{fsg \sqcup fsg \sqcup fsg}$ {DISTAL(z) & WOMAN(x) & BEAUTIFUL(x)}
              'that entity (referred to by fsg) that is distal, a woman, and beautiful'

It is possible for some of the elements to be underspecified for f or sg, but any other specification would lead to ungrammaticality (such as **illa femina bellus^{msg}*). Notice that the indexed information becomes relevant only when semantic composition takes place; therefore this account indeed locates agreement at the syntax-semantics interface. Index unification itself is symmetric; all asymmetries result from the fact that N+A forms a noun phrase, and D+NP forms a determiner phrase.

This approach can easily be extended to all other kinds of agreement, taking into account the specific mode of composition. If one forms the assertion *illa femina bella modesta est* 'that beautiful woman is modest', one predicates 'be modest' (λu^{fsg} MODEST(u)) of the result of (9), and if one continues with the pronoun *ea^{fsg}* 'she', *illa femina bella* is a possible antecedent (referring to the same person) because the ϕ-features are compatible.

Considering examples of long-distance agreement one can easily see that the way in which the complex expression is composed semantically becomes relevant. If one tries to describe these

examples purely syntactically, complex assumptions about the interaction of several AGR-nodes with argument positions are necessary, assumptions that ultimately only can imitate semantic composition in deep-structural syntax.

In the Italian example (10), the subject agrees locally with the auxiliary, but also non-locally with the participle embedded in a dependent infinitive clause. The agreement facts thus indicate that 'have fear' and 'be arrived' predicate of the same individual referent. If one assumes that *ha paura* is a subject control verb, the identity of arguments follows from the lexical entry of control verbs. By functional application the argument variable of 'be arrived' then fuses with the argument variable of *ha paura*, and the ϕ-features must unify, as shown in (11).

(10)   Italian participles
    a. Giovanni^{3msg} ha^{3sg} paura   di essere arrivato^{msg} troppo tardi.
        G.   has   fear   to be arrived   too late
        'Giovanni fears to have arrived too late'
    b. I ragazzi^{3mpl} hanno^{3pl} paura di essere arrivati^{mpl} troppo tardi.
        'The boys fear to have arrived too late'

(11)   *ha paura*   *di essere arrivato troppo tardi*
       $\lambda P \, \lambda x^{3sg} \, \text{FEAR}(x,P(x))$ $\lambda y^{msg} \, \text{ARRIVED\_TOO\_LATE}(y)$

            $\lambda x^{3sg \, \sqcup \, msg} \, \text{FEAR}(x,\text{ARRIVED\_TOO\_LATE}(x))$
            'the property (of 3msg entities) of fearing to have arrived too late'

Another even more spectacular instance of non-local agreement is shown in Hindi verbal complexes formed from an auxiliary, a subject control verb, and an embedded infinitive. All three verbs agree with the subject in (12a), but with the embedded object in (12b). Since the perfect morphology assigns ergative, which, however, blocks agreement in Hindi, the only candidate left to agree with the verb is the embedded object. A verb such as 'be' or 'want' can thus agree with an argument that does not belong to its own argument structure. Only when the verbal complex is formed, the argument structure of the lowest verb is inherited to the result.

(12)   Hindi verbal complexes
    a. Raam^{msg} roTii^{fsg} khaanaa^{msg} caahtaa^{msg} thaa^{msg}.
        Ram   bread   eat.INF   want.IMPF   be.PAST
        'Ram habitually wanted to eat bread'
    b. Raam ne roTii^{fsg} khaanii^{fsg} caahii^{f} thii^{fsg}.
        Ram ERG bread   eat.INF   want.PERF be.PAST
        'Ram had wanted to eat bread'

The analysis of (12b) is sketched in (13). The verb form *khaanii* instantiates the feature complex fsg, which may be associated with either argument (whereby association with the subject 'u' is the default), whereas the perfect form *caahii* instantiates f, which cannot be associated with the argument role assigned ergative. In this context, only the second alternative of *khaanii* is possible because it allows unification.

(13)   *khaanii*   *caahii*
$\lambda y\ \lambda u^{fsg}$ EAT(u,y)
$\lambda y^{fsg}\ \lambda u$ EAT(u,y) $\lambda P^f\ \lambda x$ WANT(x,P(x))
                                           ERG

$\lambda y^{fsg\sqcup f}\ \lambda x$ WANT(x,EAT(x,y))
            ERG

Other types of verbs that form verbal complexes in Hindi are raising-to-subject verbs, such as 'seem' in (14a), and raising-to-object verbs, such as 'see' in (14b). Here again the ergative assigned by a perfect form blocks subject agreement.

(14)   Hindi verbal complexes
   a. Raam ne roTii^{fsg} khaayii^f lagtii^f thii^{fsg}.
      Ram ERG bread   eat.PERF   seem IMPF be.PAST
      'Ram seemed to have eaten bread'
   b. Raam ne billii^{fsg} aatii^f dekhii^f thii^{fsg}.
      Ram ERG cat   come.IMPF   see.PERF   be.PAST
      'Ram had seen a cat coming'

Only an analysis that keeps record of all the argument variables in the verb complex formation is able to explain non-local agreement of the type just illustrated. Nothing is spectacular in such an analysis. All that one needs are quite general assumptions about the representation of lexical items and the way in which these entries are processed, together with few language-specific principles, in this case concerning the interaction of case and agreement.

   Hindi also allows for some intransitive verbs to assign ergative. (15a) illustrates an example where only one argument is present, which by virtue of the ergative is blocked for agreement. Therefore, all the verbs in the verbal complex must be realized in their default form, which is the msg-form. Most languages among their verb forms have one form that serves for this so-called neutral agreement. In such a case, the feature complex introduced by the verb form cannot be associated with an argument variable. In German, as in many other languages, only nominative forms can agree with the verb; compare (16a) where agreement takes place, with the alternative (16b) where it is blocked because of accusative assignment, and therefore only a 3sg verb form is grammatical.

(15)   Hindi
   Siitaa^{fsg} ne   chiiNkaa^{msg} lagtaa^{msg} thaa^{msg}.
   Sita ERG   sneeze.PERF   seem.IMPF be.PAST
   'Sita seemed to have sneezed'

(16)   German
   a. Ich$^{1sg}_{NOM}$ friere^{1sg}.   b. Mich$^{1sg}_{ACC}$ friert^{3sg}/ *friere^{1sg}.
      'I am cold'   'I am cold'

Not all languages exhibit agreement. Clearly, a language such as Vietnamese with hardly any morphology cannot have agreement, but also Malayalam (a Dravidian language spoken in Kerala, SW India), which shows rather rich derivational morphology as well as morphological case, nevertheless

lacks all agreement other than anapher-antecedent. Moreover, the extent to which agreement is used differs from language to language. Not only may the number of distinctions in the gender, number, and person categories, as well as the richness of inflectional affix inventories vary, it also depends on language-specific factors in which contexts agreement is obligatory, optional, or even excluded.

Hungarian marks plural on nouns, adjectives, and verbs. However, adjectives only agree with a noun when they are used predicatively, see (17a,b). Surprisingly, in the context of a numeral, which inherently contains the concept of plurality, the noun must not be marked for plural (17c). Furthermore, a predicative adjective or verb only agrees with a subject phrase which is overtly marked for plural, see (17d,e).

(17)   Hungarian plural agreement
   a.   gyors hajók^{pl} / *gyorsak^{pl} hajók^{pl} b.   Ezek^{pl} a hajók^{pl} gyorsak^{pl} / *gyors^{sg}.
        'fast ships'   'These ships are fast'
   c.   öt hajó^{sg}/ *öt hajók^{pl} d.   Az öt hajó^{sg} gyors^{sg}/ *gyorsak^{pl}.
        'five ships'   'The five ships are fast'
   e.   Az öt nagynéni^{sg} sört iszik^{3sg}/ *isznak^{3pl}.
        DET five aunt beer drinks
        'The five aunts drink beer'

Since Hungarian has the plural forms *gyorsak* 'fast' and *hajók* 'ships', these forms must somehow be blocked in (17a,c). Ortmann (2000) proposes that the agreement facts in (17) follow from the interaction of violable economy and expressivity principles. Slightly simplified, if one assumes the economy constraint 'Avoid multiple expression of plurality in NPs' to rank above the expressivity requirement 'Mark plural on the noun!', both *gyors hajó* and *gyorsak hajók* would violate one of these constraints, and *gyors hajók* turns out to be the optimal form in a context in which more than one ship is meant. In contrast, the singular form *öt hajó* is preferred because the numeral *öt* already expresses plurality; in this case, it is better to violate the lower-ranked constraint rather than the higher-ranked one. The tableau (18) illustrates that the winner exhibits the fewest violations of higher-ranked constraints.

(18)

several ships are meant	Avoid multiple expression of plurality in NPs!	Mark plural on the noun!
a. gyors hajó^{sg}		violated
⇒ gyors hajók^{pl}		
gyorsak^{pl} hajók^{pl}	violated	
b. ⇒ öt^5 hajó^{sg}		violated
öt^5 hajók^{pl}	violated	

If plural is unmarked on the noun, the noun phrase is formally specified as sg and thus incompatible with a plural verb form, as shown in (17e).

As we have seen, Hungarian (unlike English and German) avoids NP-internal agreement. There are also languages that avoid subject-verb agreement, even if they have the suitable affixes. Kurdish, for instance, marks plural only on the verb, see (19). In contrast, Georgian marks plural on the verb only if the subject refers to animate beings, see (20).

(19) Kurdish: Mirov /*mirovan^{pl} hat-in^{3pl}.
            man   come.past-3pl
            'The men came'

(20) Georgian
    a. Burt-eb^{pl}-i   goravs^{3} / *goraven$^{3pl\text{-}anim}$.   b. Knut-eb^{pl}-i   goraven$^{3pl\text{-}anim}$ /*goravs^{3}.
       ball-pl-NOM   roll   kitten-pl-NOM   roll
       'The balls are rolling'   'The kittens are rolling'

These few examples may have convinced the reader that the ways in which agreement is realized vary considerably from language to language. Besides languages that lack agreement altogether, there are also languages that avoid agreement in certain domains, as well as languages that opt for multiple agreement, thus allowing more freedom in word order, as seen in the following example from Latin.

(21) Grandia^{npl} per multos$^{mpl}_{ACC}$ tenuantur^{3pl} flumina^{npl} rivos$^{mpl}_{ACC}$. (Ov. *Rem.* 445)
    'Big rivers dissolve into numerous streams'

The morphology and syntax of agreement is certainly one of the aspects in which languages widely differ, this topic thus being one of the most interesting parts of linguistic typology.

References

Anderson, Stephen R. (1988) Inflection. In Michael Hammond and Michael Noonen (eds.) *Theoretical Morphology. Approaches in Modern Linguistics*, 23-43. San Diego: Academic Press.

Chomsky, Noam (1991) Some notes on economy of derivation and representation. In R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Corbett, Greville G. (1991) *Gender.* Cambridge University Press.

Corbett, Greville G. (1994) Agreement. In R.E. Asher (ed.) *Encyclopedia of language and linguistics*, vol.1, 54-60. Oxford: Pergamon Press.

Corbett, Greville G. (1995) Agreement. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.) *Syntax: An International Handbook of Contemporary Research*, vol. 2, 1235-1244. Berlin: de Gruyter.

Corbett, Greville G. (1998) Morphology and agreement. In Andrew Spencer and Arnold M. Zwicky (eds.) *The Handbook of Morphology*, 191-205. Oxford: Blackwell.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag (1985) *Generalized Phrase Structure Grammar.* Oxford: Basil Blackwell.

Laka, Itziar (1988) Configurational heads in inflectional morphology: the structure of the inflected forms in Basque. *International Journal of Basque Linguistics and Philology* 22:343-365.

Lapointe, Stephen G. (1988) Toward a unified theory of agreement. In M. Barlow and C. A. Ferguson (eds.) *Agreement in natural language: approaches, theories and descriptions*, 67-87. Stanford, Calif.: CSLI.

Lehmann, Christian (1982) Universal and typological aspects of agreement. In Hansjakob Seiler and Franz J. Stachowiak (eds.) *Apprehension: Das sprachliche Erfassen von Gegenständen.* II, 201-267. Tübingen: Narr.

Ortmann, Albert (2000) Where plural refuses to agree: feature unification and morphological economy. *Acta Linguistica Hungarica*. To appear.

Pollard, Carl, and Ivan A. Sag (1994) *Head-driven Phrase Structure Grammar*. The University of Chicago Press.

Pollock, Jean-Yves (1989) Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.

Wunderlich, Dieter (1994) Towards a lexicon-based theory of agreement. *Theoretical Linguistics* 20:1-35.

Wunderlich, Dieter (1997) Cause and the structure of verbs. *Linguistic Inquiry* 28:27-68.