1   Introduction: canonical agreement

Agreement is a widespread and varied phenomenon. In some of the world's languages it is pervasive, while in others it is absent. Despite extensive research, agreement remains deeply puzzling. There was a time when it was treated mainly as a tool for researching other syntactic phenomena. Yet there has also been a tradition of recognizing it as a challenging problem in its own right. Indeed agreement presents serious problems for all our theories of syntax. It is therefore worth looking first at the reasons for the continuing interest in agreement (§1.1). Part of this comes from the way in which it involves so many components of grammar (§1.2). The terminology has become somewhat confused, so I clarify the terms I shall use (§1.3). The substantial part of this chapter lays out the canonical approach to agreement (§1.4), which will form the basis for my typology. I then outline the way in which the book is structured (§1.5), and present background information which should be of value to the reader (§1.6).

1.1   The special interest of agreement

Consider the following idea:

Hypothesis I: Grammatical information will be found only together with the lexical item to which it is relevant. (False)

This hypothesis suggests a situation which is iconic, functional, sensible and understandable. Compare *dog* and *dogs*, where number is marked in accordance with the hypothesis, or *compute* and *computed*, where tense is similarly marked. This entire book presents evidence to show that Hypothesis I is also wrong.

It is surprising that grammatical meaning can be 'displaced' (Moravcsik 1988: 90), in other words, that one word can carry the grammatical meaning relevant to another. This is what happens in agreement:

(1)   Mary makes pancakes.

Here *makes* is singular because *Mary* is an individual; even if she makes pancakes frequently, the number of 'pancake making events' will not affect the agreement of the verb. The verb form tells us how many Marys there are, not how many makings there are. Thus the number information on the verb is displaced. This

1

displaced information, or 'information in the wrong place', is not a minor issue. Agreement affects different components of grammar, as we shall see in the next section.

1.2   The place of agreement

Take another simple example like:

(2)   The cooks make pancakes.

We need to specify that the form *make* ~ *makes* varies according to the subject (there is no effect if we change the object *pancakes* to *bread*, for example). Clearly, then, agreement is a matter of **syntax**, since the syntactic role of the items involved is of importance. But now compare:

(3)   The committee has agreed.
(4)   The committee have agreed.

Here there is a choice in some varieties of English, notably in British English. That is, there is a choice here, but not with *Mary* in (1) above. Why not? Because *Mary* is an individual, whereas *committee* may be conceptualized as an entity or as several individuals. Clearly, then, agreement is also a matter of **semantics**.

Particularly if we start from English data, we might think that agreement is all a matter of semantics, an idea put most consistently in Dowty & Jacobson (1989). We could argue that the singular verb in (1) results from semantic compatibility with a singular actor, and the plural in (2) similarly from a plural actor. However, there are three types of problem with such a view.

Consider first these examples from Morgan (1984: 235):

(5)   More than one person has failed this exam.
(6)   Fewer than two people have failed this exam.

Here we can see that the agreement of the verb depends on the grammatical number of the subject (shown by *person* versus *people*) and not on the meaning of the sentence (semantic plural in (5) and singular in (6));[1] another type of supporting example is given in §5.6.3.

There is a more general second argument that agreement cannot be entirely semantic which involves agreement in grammatical gender, in languages like Russian:

Russian
(7)   Lamp-a   stoja-l-a   v   ugl-u
           lamp(F)-SG   stand-PST-F.SG   in corner-SG.LOC
           'The lamp was standing in the corner.'

---

[1] For the form of pronouns with such phrases see Gil (2001).

In this example there is no semantic reason for *lampa* 'lamp' to be of feminine gender.[2] A similar argument can be made with grammatical number in English. The use of plural agreement with English *scissors* does not, for many linguists at least, have a semantic justification.

The third argument is that even when there are semantic reasons for a particular type of agreement, the domain in which this is possible is determined by syntax. *The committee have agreed* is fine in British English (as in (4)), which suggests that *committee* takes agreement according to its meaning. And yet **these committee* is quite unacceptable. It is syntax which determines when agreement according to meaning is possible. We shall see many more examples of such mismatches in agreement in chapter 5. And evidence from acquisition also supports the syntactic basis of agreement in English (§9.3). Thus an adequate theory requires reference both to syntactic and to semantic information (Pullum 1984).

Now consider for contrast:

(8)   The committee agreed.
(9)   Mary made pancakes.
(10)   The cooks made pancakes.

Here we see no evidence of agreement. Past tense verbs in English do not show agreement. Clearly, then, agreement is a matter of **morphology** (word structure) since we require the morphology to provide the opportunity for agreement to be indicated. Indeed agreement is arguably the major interface problem between morphology and syntax, and hence appears particularly difficult when viewed from the heartland of either component.

There is a single exception to the statement about the past tense in English, namely the verb *be* which distinguishes number in the past (*was* ∼ *were*). This is something that has to be stated individually for this verb, in its lexical entry. We conclude that agreement is a matter which may have to be specified in the lexicon; it is a matter of **lexicology**.

It is tempting to try to treat all such specific irregularities within the lexicon, but some apply so broadly that this approach cannot be right. Consider this example:

Russian (19th century, from Turgenev's *Nakanune* 'On the Eve', 1860)

(11)   «Mamen´ka **plač-ut**, — šepnu-l-a   ona vsled uxodivš-ej
        Mother   cry-3PL   whisper-PST-F.SG she after   leaving-F.SG.DAT

        Elene,   a   papen´ka **gnevaj-ut-sja** . . .»
        Elena.DAT and   father   be.angry-3PL-REFL

        'Your mother is crying', she whispered after Elena, who was leaving,
        'and your father is angry . . .'

The speaker is a maid, talking in turn about her mistress and her master. Here the plural verbs with singular subjects indicate that the speaker is showing respect

---

[2] Dowty & Jacobson (1989: 98–101) discuss the problem of gender and attempt to meet the objection, by suggesting that a real-world property of objects is the word which is used by convention to denote that class of objects. This is hardly convincing, in my view.

for the people referred to. There are all sorts of items which could appear in this construction. They cannot be restricted to particular lexical items, rather a range of noun phrases may be involved. The generalization involves the situation: this agreement occurs when the speaker wishes to show respect (to the referents of the noun phrases agreed with). Hence agreement can be a matter of **pragmatics**.

Agreement is increasingly recognized as of interest not just for core areas of linguistics like syntax and morphology, but also more widely, in work on acquisition and in psycholinguistics, for instance, which are topics I take up in the final chapter. Given this interest from 'outside', it is particularly important that we should be talking about the same thing. Unfortunately, the terminology is muddled, and important choices in analysis are made sometimes as much by tradition as by argument. I therefore will pay attention to key terms and to the analytic choices available.

1.3   Defining terms

I have just argued for the need for clarity in terminology. What then is it that unites the examples of agreement we have considered so far? Anderson (1992: 103) points out that agreement is 'a quite intuitive notion which is nonetheless surprisingly difficult to delimit with precision'. Indeed, while several definitions have been proposed, none is fully satisfactory; see the suggestions by Keenan (1978: 167), Lehmann (1982: 203) and Lapointe (1988). There is detailed discussion of definitional issues in Mel'čuk (1993) and a formal approach can be found in Avgustinova & Uszkoreit (2003). We shall start from a suggestion by Steele:

> The term *agreement* commonly refers to some systematic covariance between a semantic or formal property of one element and a formal property of another. Steele (1978: 610)

This covers the instances we have seen. The essential notion is covariance. It is not sufficient that two items happen to share properties; the sharing must be systematic, and we see this by the fact that as one element varies so will the other.

Some terms will be useful at this stage, to allow us to generalize about different types of agreement. We call the element which determines the agreement (say the subject noun phrase) the **controller**. The element whose form is determined by agreement is the **target**. The syntactic environment in which agreement occurs (the clause for instance) is the **domain** of agreement. And when we indicate in what respect there is agreement, we are referring to agreement **features**. Thus number is an agreement feature, it has the values: singular, dual, plural and so on. This is diagrammed in Figure 1.1.

Features are directly reflected in agreement. There can be other factors (like word order) which have an effect on agreement but are not directly reflected like
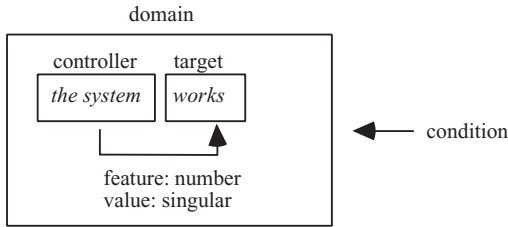
Figure 1.1   *Framework of terms*

features. Such factors are called agreement **conditions**. Thus, within a particular domain, a target agrees with a controller in respect of its feature specifications (that is, the features and their values); this may be dependent on some other condition being met.

These terms are now fairly standard among those working on agreement. For controller, the term 'trigger' or 'source' is sometimes found. 'Category' may be found in place of 'feature', and 'conditioning factor' for condition. For 'probe' and 'goal' see §4.2.5. As our terms suggest, there is a clear intuition that agreement is asymmetric. In *Mary laughs*, most accept that *laughs* is singular because *Mary* is singular. However, it does not follow that we should model it in this way. Older accounts of agreement captured the intuition directly by copying feature specifications from the controller to the target. More recent accounts use techniques like unification, and model the asymmetry less directly. This issue is considered in §1.4.3, and discussed more fully in §4.1.

I shall further clarify what is covered by agreement. First I deal with the term 'concord' (§1.3.1) and then I examine the relation of agreement to government (§1.3.2). The main way forward, however, will be using the notion of canonical agreement (§1.4), which will allow us to work with the full range of agreement, from the core instances of the phenomenon to those at the fringe.

1.3.1   Agreement and concord

These innocent terms have led to considerable confusion. For many linguists they are synonymous; the trend is towards the use of 'agreement',[3] which is the term I shall use. Some others have distinguished the terms, but they have done so in contradictory and potentially confusing ways. Since some of this confusion remains in the literature, I shall outline two positions, so that readers can be alert to the issues. Readers for whom this is not an issue should go straight to §1.3.2.

---

[3] In a survey of the topic, Moravcsik (1978: 333) gives 'agreement (or concord) phenomena'. Similarly Trask (1997: 10) has '**agreement** (also **concord**)'. Somewhat earlier, Lyons (1968: 239) had '*concord* (or 'agreement')'; this suggests that 'agreement' is on the rise, an impression supported by Anderson (1992: 103), who writes 'just what is "agreement" (or as it is often called in the traditional literature, "concord")?'

Some linguists, following Bloomfield (1933: 191–4), treat agreement as the superordinate term. According to Bloomfield (1933: 191), 'In a rough way, without real boundaries, we can distinguish three general types of agreement.' These are 'concord' or 'congruence', which includes agreement within the noun phrase and the agreement of predicate verbs, government and cross-reference.[4] As was predictable, this system of terms has not survived unchanged. One development has been to restrict concord to the noun phrase, which means that the domain of agreement is the key part of such definitions.[5]

In contrast to the position of Bloomfield, and developments from it, Greenberg (1978: 50) treats concord as the wider term:

> It would be useful, then, to distinguish the wider notion of concord from agreement, the latter being a subtype in which the choice of alternative concord elements depends on the class to which the stem of the governing item belongs, whether marked by an affix or not.

Greenberg would include matching in case within the noun phrase as an instance of concord. When, however, matching is determined by 'the class to which the stem of the governing item belongs', then we have agreement. Greenberg cites gender here, and is clearly talking of what we would term a lexical feature.

Note the contrast between this definition and Bloomfield's. Most obviously the subset relations are different: for Bloomfield concord is a subset of agreement, while for Greenberg agreement is a subset of concord. But the criteria on which the relation is based differ too. Bloomfield and several followers draw a distinction according to domain: concord exists in a 'smaller' domain than cross-reference. For Greenberg the distinction is based on the type of feature involved: agreement involves lexical features, while concord can involve matching of other features.[6]

Thus no distinction is drawn consistently between the terms 'agreement' and 'concord', indeed they are used in opposing ways. I shall therefore use just

---

[4] Bloomfield puts certain pronominal constructions and pro-drop together as cross-reference, and includes them with concord and government under 'agreement', but he treats antecedent-anaphor relations separately.

[5] For example:

> The term *concord* traditionally distinguishes this pattern of agreement within DP from the canonical specifier-head type: agreement theory as developed in Chomsky 1993 and related work accounts only for the latter.   (Carstens 2000: 323)

Note the distinction; what for Bloomfield counted as concord is cut down to agreement within the determiner phrase, and part of what he treated as concord is treated as the 'canonical' type of agreement. The difference in the definitions depends on what is considered the domain of agreement.

[6] The domain is not a defining criterion here, indeed Greenberg later talks of 'three types of concordial phenomena' and distinguishes in what he calls a 'somewhat rough and heuristic fashion' between agreement within the noun phrase, predicate agreement and anaphoric use (1978: 75–6). There are other ways in which the terms are used. Thus Lehmann (1982: 206, 249–50) also distinguishes agreement from concord: agreement is the core syntactic phenomenon, which he defines, and the term 'concord' is then used for instances of semantic compatibility, certain classifier-noun relations, for example. But still others use both terms without definition.

'agreement', as the more current term. There is no particular reason to determine
my terms primarily according to the domains of agreement or to the features
involved: both should be a part of the account (as will be the case in my 'canoni-
cal' approach). Any subdivision of agreement, whether or not 'concord' is used
as the term, will require a careful definition, since there is no generally accepted
terminology here.

1.3.2   Agreement and government

In the clearest instances of agreement (those I shall later treat as
'canonical'), agreement can be distinguished from government rather readily.
The differences can be illustrated by this example taken from a corpus of spoken
Russian.

Russian conversation (Zemskaja & Kapanadze 1978: 251)

(12)   Zna-eš́   kak-oj   mne   vsegda   dava-l-a
           know-2SG what-M.SG.ACC 1SG.DAT   Always   give-PST-F.SG

           sovet   moj-a   mam-a ?
           advice (M)[SG.ACC] my-F.SG.NOM   mother(FEM)-SG

           'Do you know what advice my mother always gave me?'

The subject is *moja mama* 'my mother', and the verb agrees with it. In agreement
the feature specification of the target is in the relevant respects the same as that
of the controller (here feminine singular).[7] In turn the verb governs the split noun
phrase *kakoj sovet* 'what advice'.[8] For government it is simply the presence of
the verb *davat´* 'give' which requires the accusative case for this noun phrase;
changing the form of the verb to, say, the present, does not affect its government
requirement (this is point 1 in (13) below). Another way of expressing this is
to say that the agreement controller has the feature specification required of the
target (i.e. the subject is indeed feminine and singular in my example), while the
governor does not (the verb is not accusative), as in point 2 below. The controller
of agreement is usually nominal, while targets are of various sorts; conversely,
the governor can be varied, but items which are governed are nominal (point 3).
The features involved in agreement, typically gender, number and person have
direct semantic relevance, to varying degrees (discussed further in §4.2.4), while
government canonically involves case, which is not directly involved in semantic
interpretation (point 4). And finally, if there are multiple targets for an agreement
controller, they will in the canonical instance share the same values (when they
realize the same features); thus *moja* 'my' and *davala* 'gave' are both feminine
singular.[9] However, when a single governor governs two governees, they will

---

[7] This is a further important aspect of 'systematic covariance' in the definition above, namely that
it is in respect of the same feature. Thus if the case of an argument varies according to the aspect
of the verb, this would not qualify as agreement any more than does normal government (thanks
to Atle Grönn for pointing out this issue).

[8] We discuss the glossing of phrases like this showing syncretism in §1.6.3 below.

[9] As we shall see in §1.4.4, hybrid controllers are non-canonical in this regard.

normally have different feature values; thus the noun phrase *kakoj sovet* 'what advice' is accusative, while *mne* 'to me' is dative, as in point 5.

(13)   Summary of differences: canonical agreement and canonical government

	AGREEMENT	GOVERNMENT
1. feature specification of target/governee is determined by:	feature specification of controller	presence of governor
2. controller/governor:	has the relevant feature specification	does not have the relevant feature specification[10]
3. element which is normally nominal:	controller	governee
4. features involved are:	gender, number, person, i.e. 'direct' features (§4.2.4)	case, i.e. an 'indirect' feature
5. multiple targets/governees are:	same as each other	different from each other

In the canonical instances agreement and government are rather different, agreement being characterized by matching, and government lacking this.[11] However, they share the characteristic of being syntactic relations of an asymmetric type. Indeed, in recent work in Minimalism, the operation Agree is given a major role, covering both agreement and case government (see Chomsky, 2000: 101). I shall here restrict myself to agreement in the narrower sense, retaining the sharper notion of the covariance of features, not found in government. Adopting the broader definition would blur this important distinction. While I have treated the canonical instances, there are difficult phenomena falling between these idealizations, as we shall see when we consider data from Kayardild (§4.5.2). We return to the issue of agreement in case in §4.4.1, and for 'collaborative agreement', which involves an interaction with case, see §3.3.5.

1.4   Canonical agreement

To clarify some of the conceptual problems and misunderstandings that have characterized the topic of agreement I shall adopt a 'canonical' approach. This means that I shall take definitions to their logical end point and build a

[10] Strictly speaking, it does not *necessarily* have the relevant feature specification; it may have it coincidentally. For example, if we have a verb which governs the genitive, a participle formed from it may be in the genitive. The fact that this participle then governs the genitive is still a matter of it being present, and does not depend on its being in the genitive.

[11] For extended discussion of definitions of agreement and government see Schmidt & Lehfeldt (1995).

theoretical space of possibilities. Only then do I ask how this space is populated. It follows that canonical instances, which are the best and clearest examples, those most closely matching the 'canon', may well not be the most frequent. They may indeed be extremely rare. However, they fix a point from which occurring phenomena can be calibrated. Then I discuss weakenings of the criteria, which allow for less canonical instances. As these instances no longer fully match the definitions, they will include some which not all linguists would accept as instances of agreement. At several points I introduce here interesting phenomena which are then taken up in more detail in later chapters.

To start from an instance of canonical agreement, consider agreement in gender in the Italian noun phrase:

Italian (Pierluigi Cuzzolin, personal communication)[12]

(14)   il   nuov-o   quadr-o
         DEF.M.SG new-M.SG   picture(M)-SG
         'the new picture'

(15)   i   nuov-i   quadr-i
         DEF.M.PL new-M.PL   picture(M)-PL
         'the new pictures'

(16)   la   nuov-a   tel-a
         DEF.F.SG new-F.SG   painting(F)-SG
         'the new painting'

(17)   le   nuov-e   tel-e
         DEF.F.PL new-F.PL   painting(F)-PL
         'the new paintings'

I shall discuss canonical aspects of such examples in turn. As a brief summary, the canonical aspects of these examples are as follows:

controller: is present, has overt expression of features, and is consistent in the agreements it takes, its part of speech is not relevant (this is a vacuous criterion in (14)–(17))
target: has bound expression of agreement, obligatory marking, doubling the marking of the noun, marking is regular, alliterative, productive; the target has a single controller and its part of speech is not relevant
domain: agreement is asymmetric (the gender of the adjective depends on that of the noun), local, and the domain is one of multiple domains
features: lexical (in one instance), matching values, not offering any choice in values
conditions: no conditions

For some readers examples like (14)–(17) will seem familiar; however, it is worth reflecting on how interesting they are. Each is a clear counter-example to

---

[12] Glossing conventions are discussed in §1.6.3. Inherent features (§4.2.3) are given in parentheses. Thus gender is glossed with the noun stem; it is true that *-a* on the noun often implies feminine gender by the assignment rules (§4.3.1) of Italian, but this is not necessarily so, as with *poeta* 'poet' (masculine); similarly *-o* often implies masculine, but this is not always the case, as with *mano* 'hand' (feminine). This glossing may seem over-careful. However, when discussing agreement it is important to distinguish between what is inherent and what is contextual.

Hypothesis I. As we shall see, the different canonical aspects of agreement converge, so that agreement in gender of the modifier with the noun in the noun phrase is confirmed as the canonical instance. Phenomena which extend the instances 'outwards' are now grouped under the five components (Figure 1.1) of my account of agreement.

1.4.1   Controllers

Several of the criteria relate to the controller. An important one is that canonical controllers are present.

> *C-1: controller present > controller absent* (where '>' means 'more canonical than')

Compare these two similar examples:

Russian
(18)   ty   čita-eš´
        2SG.NOM   read-2SG
        'you are reading'

Serbian/Croatian/Bosnian
(19)   čit-aš
        read-2SG
        'you are reading'

In such sentences in Russian the controller is typically present, while in Serbian/Croatian/Bosnian[13] typically it is not. I treat as canonical what is sometimes called 'grammatical agreement' rather than 'anaphoric agreement' (Bresnan & Mchombo 1987, Siewierska 1999, Bresnan 2001a: 151). An effect of adopting criterion 1 is that, for the construction we are discussing, the canonical type is restricted to relatively few languages, since the omission of subject pronouns (often referred to as 'pro drop') is common. It is important to stress that canonical is not necessarily what is 'normal' or 'common'. Several familiar examples of languages where pronominal subjects are normally included come from northern Europe (English and German being obvious examples).[14]

While discussions of 'dropping' concentrate on pronouns, I am making a more general point here: it is more canonical for any controller to be present rather than absent. For agreement of the adjective with the noun within the domain of noun phrase, it is more canonical for the noun to be present; similarly in

---

[13] I use this designation for the South Slavonic varieties spoken in Bosnia-Hercegovina, Croatia, Montenegro and Serbia, since linguistically they show considerable similarity. An account of the breakup of Serbo-Croat is given in Greenberg (2004).

[14] A particularly interesting less familiar example is Skou (New Guinea), which has elaborate agreement marking (to be discussed in §3.2.3 and §3.2.4) and which normally includes subject pronouns: the third person pronouns are regularly included and first and second person pronouns are present more often than not (Mark Donohue, personal communication). Siewierska (2004b: 268–70) suggests in addition: the Papuan languages Au, Ekari, Koiari and Vanimo, and the Austronesian languages Anejom, Fehan and Labu.