Dependency categorial grammar and coordination*

MARTIN PICKERING and GUY BARRY

Abstract

This paper proposes a linguistic theory based on dependency relations between words and applies it to the analysis of coordinate constructions. First, we present arguments for "flexible" constituency, where constituents may overlap on a single reading of a sentence. In the context of dependency grammar, we then define the theory of dependency constituency, where words that are all linked by dependencies form a constituent, and formalize this approach within categorial grammar. Next we apply this theory to the analysis of coordinate constructions, with the goal of determining what strings of words can be coordinated. We survey a range of approaches to coordination, at first concentrating on the assumption that it should only be possible to coordinate constituents of the same category. We show that this account is inadequate even with flexible constituency. Therefore we suggest that coordination may be bound by a parallelism restriction, under which conjuncts must consist of equivalent sequences of constituents, and examine this in relation to various theories of constituency. After considering coordination in dependency grammar, we propose that an adequate account is provided by restricting coordination to equivalent sequences of dependency constituents. Finally, we demonstrate how the conjoined expressions should themselves be analyzed in order to be incorporated into complete sentences.

1. Introduction

1.1. *Theories of constituency*

The notion of a constituent, or the constituency analysis of a sentence, has been fundamental to most modern linguistics. However, the assumptions about what a constituent can be have been very restricted. We shall

Linguistics 31 (1993), 855–902

0024–3949/93/0031–0855 $2.00

outline these assumptions in the context of the very simple sentence *John loves Mary* and shall show how it is possible to liberalize the notion of constituency to allow other possible analyses.

The standard analysis for our example has the following set of constituents:

(1) {John loves Mary, John, loves Mary, loves, Mary}

Each word is a constituent, as is the complete sentence. In addition the string *loves Mary* is a constituent, a verb phrase. We shall call this the *traditional* constituency analysis and shall use the term (loosely) to refer to any standard "textbook" analysis of the constituent structure of a sentence.

This analysis is adopted in preference to the following two nontraditional analyses:

(2) a. {John loves Mary, John loves, John, loves, Mary}
b. {John loves Mary, John, loves, Mary}

The point is that (1) is contrasted with (2a) and (2b) only; no other constituency analyses are considered "possible." Let us examine what these three analyses have in common, and why discovering the constituent structure of *John loves Mary* is deemed to involve deciding between these analyses and no others.

The general intuition about constituency is that it serves to pick out the sets of words that go together. But modern linguistics goes further than this: it assumes that constituency analysis has a hierarchical structure. On this conception, constituents must "nest" within each other, so that each constituent (apart from the full sentence) is immediately contained within exactly one larger constituent. Sentences can be broken into parts, and each part can be broken into smaller parts, and so on, until the atomic elements (usually words) are reached. This conception is assumed by the vast majority of modern linguistics and derives ultimately from the work of Bloomfield (1933).

This conception makes two strong assumptions: (i) constituency is EXHAUSTIVE, that is, it is necessary for each word to be a constituent (we shall ignore units smaller than words), and that the entire sentence is a constituent; (ii) constituency is RIGID, that is, no two constituents may partly overlap, so that the intersection of two constituents must either be empty (i.e. the constituents are disjoint), or equal to one of them (i.e. one constituent is entirely contained within the other).

One very important reason for assuming that constituency is rigid and exhaustive is that context-free phrase-structure grammar (CFPSG) gives rise to such a constituent structure for any sentence generated by the

grammar; note that it provides further restrictions as a result of not allowing crossing branches. Although it is possible for a particular set of CFPSG rules to give rise to more than one analysis for a sentence, it is universally assumed that this can only happen when the sentence is (semantically) ambiguous, so that it is not possible for an unambiguous sentence to have more than one analysis. Hence it is not possible for CFPSG to generate overlapping constituents.

It is possible to abandon exhaustivity, for instance as in the following analysis:

(3) {John loves Mary, John, Mary}

Such an account was proposed by Gaifman (1965); we shall briefly consider his proposal in section 2 below. Abandoning exhaustivity seems very hard to motivate, and apart from the reference to Gaifman below, we shall not consider it further.

In this paper, we shall claim that the assumption that constituency is rigid is arbitrary, and that in fact it is linguistically advantageous to abandon it. We shall refer to any theory that allows partly overlapping constituents as a *flexible* theory of constituency, or just *flexible constituency*. In a flexible theory, *John loves Mary* could contain the following constituents:

(4) {John loves Mary, John loves, John, loves Mary, loves, Mary}

Here *John loves* and *loves Mary* partly overlap, with their intersection being the word *loves*. In general, flexibility can massively increase the number of possible constituent analyses, although in this case the only exhaustive analysis requiring flexibility is the one in (4). Of course all of the analyses in (1) could also be generated by an appropriate flexible theory. Flexible constituency is a consequence of proposals by Sampson (1975) and McCawley (1981). For example, Sampson suggests allowing phrase-structure representations to permit nodes to branch upward as well as downward, creating what he calls semitrees. In other words, nodes may have more than one mother. A consequence of this is that constituents may overlap. However, Sampson is more interested in the implications of his proposals for transformations than in the effects of admitting flexible constituency itself. In contrast, flexible constituency is explicitly proposed within the framework of "extended" categorial grammar (e.g. Steedman 1987; Moortgat 1988). We shall introduce a version of extended categorial grammar in section 2. Let us now examine some tests for constituency, in order to demonstrate that the case for flexible constituency is strong.

1.2. *Tests for constituency*

We contend that rigid constituency is too restrictive to serve as the basis for syntax, and hence that flexible constituency should be adopted. We now present evidence that most reasonable tests for constituency are more consistent with flexible constituency than with rigid constituency.

The most primitive test for constituency is replacement. Broadly speaking, this says that a string of words forms a constituent of a particular category if it can replace another constituent of that category. Of course this entails that we already know of at least one constituent in that category. Let us assume that constituency is exhaustive. This means that each word is a constituent (as is the sentence).

There are two versions of the replacement test, which we may call the "existential" and the "universal." Under the existential replacement test, a string is deemed to be a constituent of a particular category if it can replace a word of that category IN SOME ENVIRONMENT; under the universal replacement test, a string is deemed to be a constituent of a particular category if it can replace a word of that category IN ANY ENVIRONMENT. Under either version of the test, the expression formed by replacement should have the same distribution as the original expression. It is often the case that replacement is partly interpreted in semantic terms, so that the meanings of the sentence before and after replacement differ only with respect to the changed element. However, there is no way to be precise about this without a detailed theory of semantics.

"Existential" replacement works in many cases but turns out to be too weak to serve as a useful test. For instance, we can assume that *Mary* is a constituent in *John loves Mary* (as we are only considering exhaustive theories). Since *Mary* can be replaced by *the tall woman* to give *John loves the tall woman*, and the sentences *John loves Mary* and *John loves the tall woman* are both declarative sentences, we may conclude that *the tall woman* is a constituent. However, the same argument can be used to conclude that *Mary a great deal* is a constituent, since it can replace *Mary* in *John loves Mary* to yield *John loves Mary a great deal*, which is also a declarative sentence.

The stronger test of "universal" replacement gets around this particular problem. Since *Mary* can be replaced by *the tall woman* in all contexts where it occurs, we may again conclude that *the tall woman* is a constituent. However, there are contexts in which *Mary* cannot be replaced by *Mary a great deal*; for example, *Mary loves John* is grammatical, but not **Mary a great deal loves John*.

But universal replacement does not give evidence for traditional constituency. For example, *loves* can be replaced by *spoke to* wherever it occurs; for example, *John loves Mary* changes to *John spoke to Mary* or *I know who John loves* changes to *I know who John spoke to*, even though *spoke to* is not a traditional constituent. Likewise, it is possible to replace *loves* with *introduced Bill to*, giving *John introduced Bill to Mary* and *I know who John introduced Bill to*. Notice that a rigid theory consistent with this data would have to regard *to Mary* as a nonconstituent.

In fact, universal replacement can be used to give evidence against rigid constituency. The sentence *John entered the room* can be changed into *John went into the room* by replacing *entered* with *went into*, and similar substitutions can be made in other sentences, hence suggesting that *went into* is a constituent. But similarly, *John went in* can be changed into *John went into the room*, by replacing *in* with *into the room*, and similar substitutions are possible in other sentences (when *in* is used intransitively), and so *into the room* is a constituent. Since *into the room* and *went into* overlap in *John went into the room*, this test supports a flexible theory of constituency. It is possible to criticize any individual use of this test; what we are interested in demonstrating is that if replacement can be used as a test for constituency at all, then it presents better evidence for flexibility than for rigidity.

Now let us consider standardly used syntactic tests for constituency. We shall show that, between them, they give a strong *prima facie* case for adopting flexible constituency. Perhaps the strongest evidence is provided by coordination, on the basis of the assumption that all conjuncts have to be constituents. From *John loves Mary*, both *John loves* and *loves Mary* can serve as conjuncts (we notate conjuncts using square brackets):

(5) a. [John loves] and [Bill hates] Mary.
    b. John [loves Mary] and [hates Sue].

Of course it is possible to assume phrase-structure constituency, and to assume that (at least) one of these examples is not directly generated by the rules. Traditionally, (5a) was generated by the transformation known as *right-node raising*. But it seems odd that a relatively standard sentence like (5a) cannot be primarily generated by the grammatical theory. Without additional mechanisms such as transformations, rigid constituency cannot provide an account of any pair of coordinations where the conjuncts overlap in the related simple sentence, though these examples are extremely common. Flexible constituency does not have this problem, because it allows overlapping expressions to be constituents. However, we shall revise the assumption that all conjuncts must be constituents later in the paper.

There are a number of other reasons to opt for flexible constituency.

On the assumption that sentence fragments have to be constituents, it is possible to find evidence for partly intersecting constituents:

(6) a. Q: What will John do? A: Remind Fred to buy a card.
b. Q: Who will remind Fred to buy a card? A: John will remind Fred.

It seems reasonably clear that *remind* in the answer to (6b) has undergone ellipsis, because *John will remind Fred* is not felicitous on its own. Hence we do not appear to be simply dealing with two different subcategorizations of *remind*. Instead, both answers appear to function as constituents on the same analysis of *remind*.

If on the other hand we assume that expressions that can undergo ellipsis have to be constituents, then similar evidence can be found:

(7) a. Q: Who will remind Fred to buy a card? A: John will.
b. Q: What will John remind Fred to do? A: To buy a card.

Here, *remind Fred to buy a card* has undergone ellipsis in (7a), whereas *John will remind Fred* has undergone ellipsis in (7b). Further evidence comes from fronting:

(8) a. John loves Mary.
b. Mary, John loves.

Assuming that these sentences are reorderings of each other, they suggest that both *Mary* and *John loves* are distributional units, and hence constituents. This would not be the traditional conclusion, because fronting is only taken as evidence for the constituenthood of the fronted string. However, there is no reason to make this assumption. On the assumption that *loves Mary* is a constituent (as in traditional phrase-structure grammars), a flexible theory is supported.[1]

An extreme case of flexible constituency is *structurally complete constituency* (see Moortgat 1988), in which every continuous substring of a sentence is a constituent (and every constituent is continuous). However, it would be hard to provide linguistic motivation for structurally complete constituency, since some strings do not appear to serve as constituents under any test. For example, the string *Sue and Fred* in *John loves Sue and Fred hates Jane* does not pass any of the above tests for constituency:

(9) a. *John loves [Sue and Fred] and [Mary and Bill] hates Jane. (coordination)
b. *Q: John loves whom and who hates Jane? A: Sue and Fred. (fragments)
c. *Q: Who does what to Sue and Fred does what to whom? A: John loves hates Jane. (ellipsis)
d. *Sue and Fred, John loves hates Jane. (fronting)

The sentence also fails the replacement test; there would have to be some word α such that *John loves* α *hates Jane* was a grammatical sentence, and *Sue and Fred* could replace α wherever it occurred. The only word that will fit the first requirement is a conjunction, and clearly *Sue and Fred* cannot replace a conjunction in an arbitrary environment (e.g. **John Sue and Fred Tom left*).

Of the above tests coordination appears to be by far the freest. In the other cases, there are expressions that some tests are unable to identify as constituents but others do; for example, it is impossible to use a single case of fronting to determine that subject and finite verb phrase are constituents (e.g. **loves Mary, John*), but a finite verb phrase can replace a single word (*John ran* becomes *John loves Mary*), and *loves Mary* can undergo ellipsis (Q: *Who loves Mary?* A: *John*). With coordination, however, there appears to be the problem that almost any string can serve as a conjunct in an appropriate context. As we shall see later, there are restrictions on possible coordinations, which appear to be tied up with the account of constituency employed. These have to do with what strings can serve as conjuncts in the context of what other strings.

If the traditional arguments discussed above can be used to determine constituency, then the appropriate theory of constituency does not appear to be traditional constituency or indeed any rigid notion. Indeed, they provide evidence for assuming some version of flexible constituency. We shall now derive a specific proposal for flexible constituency based on the linguistic primitive of dependencies between words.

2. Dependency and categorial grammar

Having presented the case for flexible constituency, we shall now outline a flexible theory based on dependency relations between words. We first introduce this concept of dependency and then derive a flexible theory called dependency constituency. We then show how dependency constituency can be formulated within categorial grammar in terms of the theory called dependency categorial grammar. Much of this section is based on Barry and Pickering (n.d.).

2.1. *Dependency*

Dependency grammar (DG) is a grammatical framework in which the primitive syntactic links hold between words rather than phrases. There is a long tradition of basing linguistic description on links between words.

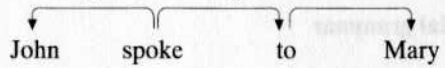
For example, such concepts as government and agreement have traditionally been viewed as primitive to the analysis of the structure of languages (especially classical languages) rather than the grouping of words into phrases. The development of the general notion of relations between words can be traced to Tesnière (1959); this work can be thought of as descriptive dependency grammar. Generative dependency grammar began with Gaifman (1965) and Hays (1964) and was heavily derivative from the development of context-free phrase-structure grammar. Since then there has been a small amount of work within generative dependency grammar, in particular Mel'cuk (1988) and Hudson (1984). A good introductory reference for dependency grammar is Matthews (1981: 78ff.).

In DG the links between words are called *dependencies* and relate two words called the *ruler* (otherwise known as the head[2] or controller) and the *dependent* (otherwise known as the controlled); we shall ignore units smaller than words. If a dependency links a ruler R and a dependent D, we say that D is *a dependent of* R (or *depends on* R), and that R is *a ruler of* D. Also we say that a word X is *subordinate* to another word Y if X is either a dependent of Y or a dependent of another word that is itself subordinate to Y.

We define a *dependency structure* for a sentence as a directed graph with the words at the nodes and the dependencies as edges. A *dependency diagram* is a diagram illustrating both the linear order of the words and the dependency structure. For example, the sentence in (10) may be given the dependency diagram in (11):

(10) John spoke to Mary.

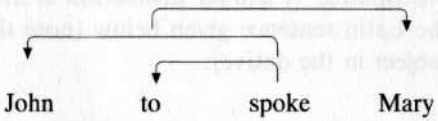
(11)



We write the words in order and join each pair of words in a dependency with an arrow leading from the ruler to the dependent. So (11) says that *spoke* has the dependents *John* and *to*, and *to* has the dependent *Mary*.

Note that if the directions on the arrows in (11) are ignored then it is always possible to trace a path from any word to any other word; such a structure is said to be *connected*. Also, the word *spoke* has no ruler, and every other word has exactly one ruler; *spoke* is therefore called the *root of the sentence*. In this paper we are concerned with *classical dependency grammar* (Gaifman 1965; Robinson 1970; see also Fraser 1989), in which every structure is connected, each sentence has a single root, and all words except the root have exactly one ruler (the *single-ruler requirement*); structures with these three properties are known as *directed trees*. In addition, every structure in classical dependency grammar obeys *adjacency*. Adjacency is the requirement that no word may intervene between the two words in a dependency unless it is subordinate to one of them. For example, (12) violates adjacency because *spoke* comes between *to* and *Mary*, and it is not subordinate to either of them:

(12)



How do we get the structure in (11)? First, we assume that the root of a sentence is the tensed verb (in the main clause). Then, we assume that the dependents of a word are those words whose categories are selected by it. In this case the root is *spoke*. The dependents of the verb are those words whose category the verb selects. In this case, *spoke* requires a noun and a preposition, so its dependents are *John* and *to*. The dependent of *to* is the word whose category *to* selects; it must be a noun and, in this case, is *Mary*.

This criterion of selection is meant as a rule of thumb. In general there is no infallible way of finding the dependency structure for a given sentence. In this respect dependency grammar is no better or worse off than phrase-structure grammar. However, selection as a test for dependency structure is supported by the use of agreement and government. If two words agree, then we have evidence that there is a dependency between them, and if one word governs another, then we have evidence that the governor is the ruler and the governed the dependent. When these tests are combined with the restrictions of classical dependency grammar, it is usually reasonably straightforward to determine the dependency structure for a given sentence.

Let us assume that it is straightforward to determine what lexical items constitute forms of a word, and we can assign them to (sub)categories in the same way that we can assign words to categories. For instance, *he* and *him* are forms of the same word (the masculine pronoun), differing in case. Verbs govern the case of the masculine pronoun, so that *he walks* is grammatical but not **him walks*. On this basis it is reasonable to claim that *walks* is the ruler and *he* is the dependent. This is backed up by the fact that a subjectless sentence such as the imperative *walk!* is possible but a verbless sentence is not. Number agreement can also be used as a test for dependencies. For instance, it is possible to say *he walks* or *they walk* but not **he walk* or **they walks*. This indicates that the forms of the verb and the pronoun covary; either both have to be plural or both singular. But it is not possible to say whether the plural form of the verb

constrains the pronoun to be plural or whether it is the other way around, and hence this gives no evidence for direction of dependency.

There is thus a distinction to be made between cases in which one word constrains the form of another, and cases in which the forms of two words are mutually constraining. A similar distinction is noted by Palmer (1971), who cites the Latin sentence given below (note that the verb *pareo* 'obey' takes its object in the dative):

(13) Viro bono paruit.
Man-SING-DAT good-MASC-SING-DAT obey-PERF-3RD-SING
'He obeyed the good man'.

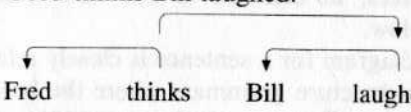
The verb *paruit* constrains the category of the noun *viro*. It also constrains the case of the noun *viro* and the case of the adjective *bono*. The cases of *viro* and *bono* are mutually constraining, as are the numbers. Finally, the noun *viro* constrains the gender of the adjective *bono*. On the basis of this evidence we can determine the dependency structure of the sentence. Nothing constrains the form of the verb so it must be the root. This leaves three possible classical dependency structures: (i) the verb is the ruler of the noun and the noun is the ruler of the adjective; (ii) the verb is the ruler of the adjective and the adjective is the ruler of the noun; or (iii) the verb is the ruler of both the noun and the adjective. But we can rule out both (ii) and (iii) because in neither structure is the adjective subordinate to the noun, and hence there would be no way for the noun to constrain the gender of the adjective. The fact that the verb constrains the case of the adjective is not realized as a dependency; the reason for this is that the verb constrains the case of the noun, and the noun and the adjective have to have the same case.

Like the criterion of selection, the criterion of constraint cannot always determine dependency structure mechanically. For instance, in *Mary saw herself*, there is at least a mutually constraining relation between the gender of *Mary* and *herself* (presumably *Mary* constrains the form of *herself*). But *Mary* and *herself* are both selected and constrained by *saw*, so these tests appear to imply too many dependencies. We would suggest that there is no dependency between *Mary* and *herself*, and that this relationship is handled by some other mechanism ("binding"). Clearly there is a strong theoretical component to the assumption of dependencies, just as there is in the assumption of rules in phrase-structure grammar. There is unlikely to be much contention about the structure of simple constructions.

Let us consider some further examples of dependency structures. Another example is (14), which has the dependency diagram in (15):

(14) Fred thinks Bill laughed.

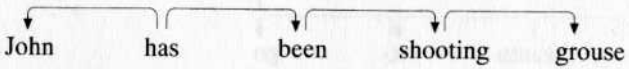
(15)



Here *laughed* is the *root* of the subordinate clause *Bill laughed*. This notion of root is the natural extension of the notion of root of a sentence; it is the word that has no ruler within the subordinate clause (and so every other word in the subordinate clause is subordinate to it). In this sentence *laughed* is a dependent of *thinks*, whereas in the sentence *Fred thinks that Bill laughed*, *laughed* is the dependent of *that*, and *that* is a dependent of *thinks*.

Note that if the tensed verb in a sentence is an auxiliary verb, then it is the root of the sentence, and so all nonfinite verbs will be subordinate to it. We also assume that the subject of the sentence is dependent on the finite verb, and all the other complements are dependent on the nonauxiliary verb. For example,

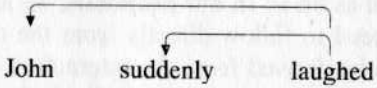
(16)



This view contrasts for example with the proposals of Matthews (1981).

Work in DG has standardly assumed that modifiers are dependent on the words they modify. They may therefore be regarded as subcategorized for, although their presence is optional. For instance, in *John suddenly laughed*, *laughed* is the ruler of *suddenly* as well as of *John*. This means that the dependency diagram is as below:

(17)



Notice that if *suddenly* were treated as the ruler of *laughed*, then adjacency would be violated.

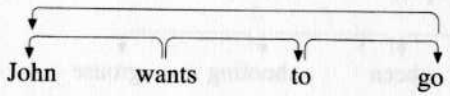
The only notion of constituency that has been discussed in relation to dependency grammar is the following one: a constituent consists of a word W and every word subordinate to W (see Gaifman 1965; cf. Matthews 1981: 84–93). We shall refer to this as the *full constituent rooted at W*. For example, in (11) the full constituents (with roots bracketed) are *(John)*, *(Mary)*, *(to) Mary*, and *John (spoke) to Mary*. In (15), the full constituents are *(Fred)*, *(Bill)*, *Bill (laughed)*, and *Fred (thinks) Bill laughed*. In (17), the full constituents are *(John)*, *(suddenly)*, and *John suddenly (laughed)*. This notion of constituency is rigid and nonexhaustive. In general, it generates structures that are "flatter" than

traditional phrase-structure trees, in contrast to the new notion of constituency that we define below.

In general, the dependency diagram for a sentence is closely related to its analysis in standard phrase-structure grammars where the head of a constituent is distinguished, such as $\bar{X}$-syntax (Jackendoff 1977). All full constituents are constituents in such an analysis, and the root of a full constituent in a dependency structure is the lexical head of that constituent in headed phrase-structure grammar (see Matthews 1981).

There have also been proposals for dependency grammars that extend beyond classical dependency grammar. An important current proposal is word grammar (Hudson 1984, 1990). The main formal difference is that the single-ruler requirement is abandoned, in order to allow additional dependencies that reflect properties of particular constructions. For example, word grammar proposes an additional dependency in "equi" constructions, so that *John* depends on *go* as well as on *wants* in *John wants to go*.

(18)



The other three dependencies are motivated by the same considerations as classical DG, but the extra dependency from *go* to *John* is motivated by the assumption that there is a sense in which *John* is the subject of *go* as well as of *wants*. (Word grammar also extends the notion of dependency by labelling the dependencies with grammatical relations.) Similarly, it proposes dependencies to reflect predication: *Mary* depends on *happy* in *Mary is happy*, as well as on *is*. In our proposals, we assume that semantic properties do not need to follow directly from the dependency relations. Instead, they can be derived from the interaction of the dependency relations and the lexical semantics of the individual words. For instance, *John wants to run* has dependencies from *wants* to *John*, *wants* to *to*, and *to* to *run* only, but the lexical semantics of *wants* indicates that its syntactic subject acts as the semantic subject of *run* (as well as the semantic subject of *wants*).

2.2. *Dependency constituency*

We shall now derive a new theory of constituency from dependency grammar called *dependency constituency*. This theory of constituency is flexible. A dependency constituent consists of words that are all linked by dependencies: in other words, if directionality is ignored, then a path

can be traced between any two words in the constituent without going outside the constituent. A dependency constituent is *continuous* if its words form a substring of the sentence, and *discontinuous* otherwise. Clearly single words will always be dependency constituents, as will the entire sentence (and hence dependency constituency is exhaustive).

Let us return to the dependency diagrams (11) and (15) and derive the dependency constituents for the associated sentences. In *John spoke to Mary*, the dependency constituents are *John spoke*, *spoke to*, *to Mary*, *John spoke to* and *spoke to Mary*, plus the complete sentence and the individual words. Note that some of these, such as *to Mary*, are phrase-structure constituents in traditional grammars, but others, such as *John spoke*, are not. Note also that some pairs of dependency constituents overlap, such as *spoke to* and *to Mary*, which illustrates that dependency constituency is flexible.

In *Fred thinks Bill laughed*, not all substrings are dependency constituents: *Fred thinks*, *Bill laughed*, and *thinks Bill laughed* are all dependency constituents, but *thinks Bill* and *Fred thinks Bill* are not. In addition, not all dependency constituents are continuous: for instance, the words *thinks* and *laughed* form a dependency constituent.

Every dependency constituent contains precisely one word that has no ruler within that dependency constituent; we may naturally extend the earlier terminology and refer to this as the *root* of the dependency constituent. (The root of a single-word dependency constituent is of course that word.) For instance, in (11) *John spoke*, *spoke to*, *John spoke to*, and *spoke to Mary* all have root *spoke*, and *to Mary* has root *to*. In (15) *thinks laughed* has root *thinks* and *Bill laughed* has root *laughed*.

Because dependency constituency is flexible, it can be used in a more adequate way than any rigid theory of constituency to describe linguistic phenomena such as coordination, sentence fragments, ellipsis, and fronting. In *John loves Mary*, *loves* is the ruler of both *John* and *Mary*, so both *John loves* and *loves Mary* are dependency constituents. This provides motivation for why both strings can be used as conjuncts, as in (5) above. Dependency constituency is also consistent with many examples of replacement; for example, *loves* can be replaced by *spoke to* to give *John spoke to Mary*. Dependency constituency is also consistent with the other tests outlined earlier; for example, *Remind Fred to buy a card* and *John will remind Fred* can both serve as sentence fragments (as in [6]) and undergo ellipsis (as in [7]), and *John loves* can be left behind after fronting (as in [8]). These considerations suggest that dependency constituency is a better foundation for syntax than phrase-structure constituency.