

Diagnosics for Constituents: Dependency, Constituency, and the Status of Function Words

Timothy Osborne
Zhejiang University
Hangzhou
China
tjo3ya@yahoo.com

Abstract

This contribution delivers two messages: 1) that the tests for constituents that are widely employed in linguistics and syntax textbooks are more congruent with dependency-based syntax than with constituency-based syntax and 2) that these same tests support the conventional analysis of function words, that is, the analysis that takes most function words (auxiliary verbs, adpositions, subordinators) to be heads over the content words with which they cooccur. The latter issue is important at present, since a recent annotation scheme is choosing to subordinate all function words to the content words with which they cooccur.

1 Two messages

Most English language textbooks on syntax and linguistics rely on tests for constituents to introduce the concept of syntactic structure. Tests such as coordination, proform substitution, topicalization, answer fragments, clefting, VP-ellipsis, pseudoclefting, etc. are used to demonstrate the presence of constituents, and thus, the presence of sentence structure. The tests show that words are being grouped together into phrases, and smaller phrases are grouped into ever larger phrases, until the largest phrase, the sentence, is reached. The tests are very widely employed, so widely that they enjoy a prominent spot in the syntactician's toolbox; they are basic tools with which the syntactician works.

An interesting aspect of most tests for constituents, however, is that they identify much less syntactic structure than most constituency grammars assume. In this respect the data delivered by the tests are relatively congruent with dependency grammars (DGs), since by its very nature dependency-based syntax posits much less syntactic structure than constituency-based syntax. Interestingly, the DGs currently in existence

rarely draw attention to this fact, that is, they rarely draw attention to the fact that the dependency-based understanding of syntactic structures is strongly supported by the basic tests that are, ironically, so widely employed by constituency grammars.

Tests for constituents can also be employed to shed light on debates about the best hierarchical analysis of various syntactic structures, for instance concerning the hierarchical status of function words. The tests are consistent with the traditional DG analysis of function words, namely that auxiliary verbs are heads over content verbs and prepositions are heads over their nouns.

This contribution draws attention to the two points just mentioned. It delivers two messages: 1) most commonly used tests for constituents are much more consistent with dependency-based syntax than with constituency-based syntax and DGs can and should draw attention to this fact, and 2) the tests reveal that auxiliary verbs are heads over content verbs and prepositions are heads over their nouns.

The data examined in this contribution are limited to English. This is due mainly to the fact that the most widely employed tests for constituents are employed in English language textbooks, applied to the syntactic structures of English. Section 6 below reflects on this aspect of the tests, considering the extent to which they can be employed in other languages.

2 Constituents

The term *constituent* is associated with *constituency* grammars, the morphological relatedness of the two words, *constituent* and *constituency*, being obvious. In this respect the first message delivered in this manuscript might seem contrary to basic terminology, this terminology suggesting that dependency and constituency are distinct principles of syntactic organization and that the constituent unit is not compatible with depen-

dependency syntax in general. I view the terminology in this area as a historical accident, and this accident has, in my view, played out to the detriment of DG, since it has obscured the fact that dependency syntax is actually more consistent with the data delivered by diagnostics for constituents than constituency syntax.

The *dependency vs. constituency* terminology as it is understood and employed today is perhaps due most to Hays' (1964) seminal article *Dependency theory: A formalism and some observations*. This early article seems to be most responsible for introducing and establishing the dependency concept and for contrasting dependency with constituency. Hays employed both terms, *dependency* and *constituency*, whereby he was emphasizing that the dependency formalism was distinct from the constituency formalism. The constituent concept at that time had already been long established; it goes back at least as far as Bloomfield (1933: 160ff.), and it is associated perhaps most with *the immediate constituent analysis* developed by Wells (1948).

The noteworthy aspect of Hays (1964) article is the terminology that he used when describing dependency trees. It is instructive to consider exactly what he wrote in this area:

“A SUBTREE is a connected subset of a tree. A complete subtree consists of some element of a tree, plus all others connected to it, directly or indirectly, and more remote from the origin of the tree...”

An IC [immediate constituent] structure and a dependency structure, both defined over the same string, correspond relationally if every constituent is coextensive with a subtree and every complete subtree is coextensive with a constituent. (Two structural entities are coextensive if they refer to the same elements of a terminal string.)” (p. 520)

The noteworthy aspect of this passage is the term *complete subtree*. Hays chose to denote a given word plus all the words that that it dominates a *complete subtree*.

Hays did not simply call the relevant unit a *constituent*. In other words, Hays was introducing a distinct terminology across dependency- and constituency-based systems. Had he employed the term *constituent* for both types of structures, the nature of the dependency vs. constituency debate might be quite different today than it is, since the terminology would have aided the comparison and evaluation of the two

competing approaches to syntactic structures.

Other dependency grammarians who followed Hays realized that constituents can be acknowledged in both dependency and constituency-based systems. Hudson (1984: 92) wrote the following in this regard:

“The general connection between dependency structure and constituent structure is that a constituent can be defined as some word plus all the words depending on it, either directly or indirectly (in other words, that word plus all the dependency chains leading up to it).”

Starosta (1988: 105) picked up on Hudson's point; Starosta wrote:

“...and a constituent is any word plus all its direct or indirect dependents”

Hellwig (2003: 603) is more explicit with his statements in this area:

“Contrary to other dependency grammars, the notion of constituent is endorsed in DUG [Dependency Unification Grammar]. However, it is a specific constituent structure that results from dependency analysis. Let us define a constituent as the string that corresponds to a node in the dependency tree together with all the nodes subordinated to that node (directly or mediated by other nodes). Then, any dependency tree can be dispersed into smaller trees until nodes with no dependents are reached. Each of these trees corresponds to a constituent of the sentence or phrase in question.”

The three passages just cited agree that constituents can be acknowledged in dependency-based structures.

Had Hays (1964) used the term *constituent* to denote the *complete subtrees* of dependency hierarchies, the realization may have long set in by now that dependency-based syntax is much more consistent with most tests for constituents than constituency-based syntax.

3 Tests for constituents

The most widely employed tests for constituents in syntax textbooks are listed next, the order given reflecting the frequency of use:

1. Coordination
2. Topicalization

3. Proform substitution
4. Answer fragments
5. Clefting
6. VP-ellipsis
7. Pseudoclefting

Coordination is the most widely employed of these tests. There are, however, major problems with coordination as a diagnostic for constituents, since phenomena such as right node raising (RNR) (e.g. *[Fred likes], but [Sue dislikes], the Chinese beer*) and so-called non-constituent conjuncts (e.g. *Fred sent [Sue to the store] and [Jim to the post office]*) appear to involve the coordination of nonconstituent strings. Due to such problems, coordination is not employed below.

The other six diagnostics, however, are more consistent about the strings that they suggest are and are not constituents. They too are very widely employed. Just how widely is documented with the following lists of syntax and linguistics textbooks that use them:

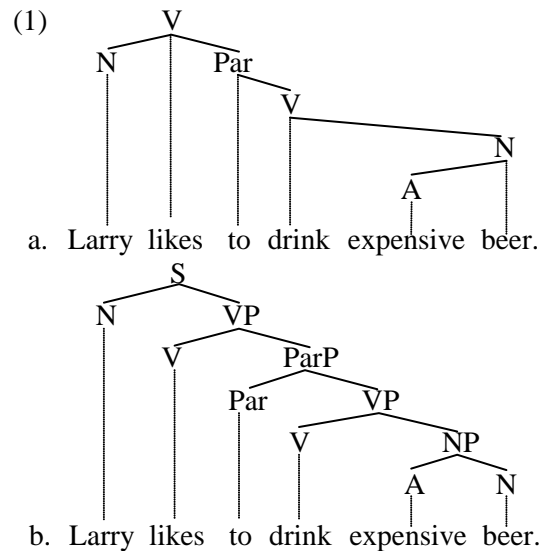
Topicalization
Allerton 1979:114, Radford 1981:213, Burton-Roberts 1986:17, Radford 1988:95, Haegeman 1991:35, Napoli 1993: 148, Borsley 1991:24, Ouhalla 1994:20, Fabb 1994:4, Haegeman and Guéron 1999:46, Fromkin et al. 2000:151, Lasnik 2000:10, Börjars and Burridge 2001:26, van Valin 2001:11, Poole 2002:32, Adger 2003:65, Sag et al. 2003:33, Eggins 1994:72, Radford 2004:72, Kroeger 2005:31, Haegeman 2006:79, Culicover 2009:84, Müller 2010:6, Sobin 2011:31, Sportiche et al. 2014:68.
Proform substitution
Allerton 1979:113, Radford 1981:64, Radford 1988:98, Thomas 1993:10, Fabb 1994:3, Ouhalla 1994:19, Radford 1997:109, Haegeman and Guéron 1999:46, Fromkin et al. 2000:155, Lasnik 2000:9, Börjars and Burridge 2001:24, van Valin 2001:111, Poole 2002:29, Eggins 1994:131, Radford 2004:71, Tallerman 2005:142, Haegeman 2006:74, Kim and Sells 2008:21, Culicover 2009:81, Carnie 2010:20, Müller 2010:5, Sobin 2011:32, Carnie 2013:98, Sportiche et al. 2014:50
Answer fragments
Radford 1981:72, Burton-Roberts 1986:16, Radford 1988:91, Haegeman 1991:28, Radford 1997:107, Haegeman and Guéron 1999:46, Börjars and Burridge 2001:25, Eggins 1994:134, Kroeger 2005:31, Tallerman 2005:125, Haegeman 2006:82, Kim and Sells 2008:20, Carnie 2010:18, Müller 2010:6, Sobin 2011:31, Carnie 2013:98, Sportiche et al. 2014
Clefting
McCawley 1988:64, Akmajian et al. 1990:150, Borsley 1991:23, Napoli 1993:148, McCawley 1998:64, Haegeman and Guéron 1999:49, Börjars and Burridge 2001:27, Adger 2003:67, Sag et al. 2003:33,

Tallerman 2005:127, Haegeman 2006:85, Kim and Sells 2008:19, Carnie 2013:98, Sportiche et al. 2014:70
VP-ellipsis
Radford 1981:67, 1988:101, Ouhalla 1994:20, Radford 1997:110, McCawley 1998: 67, Fromkin et al. 2000:158, Adger 2003:65, Kroeger 2005:82, Tallerman 2005:141, Payne 2006:163, Culicover 2009:80: Sobin 2011:58
Pseudoclefting
Borsley 1991:24, McCawley 1998: 64, Haegeman and Guéron 1999:50, Kroeger 2005:82, Haegeman 2006:88, Payne 2006:160, Culicover 2009:89, Carnie 2013:99, Sportiche et al. 2014:71

A large majority of these sources overlook DG entirely, only four of them have anything to say about DG: Borsley (1991:30f.) briefly mentions DG in passing; van Valin (2001: 86–109) grants DG more space – he devotes a chapter to it; Sag et al. (2003:535f.) grant DG less than a page; and Carnie (2010:175–8, 268f.) devotes about four pages to DG.

4 Using the tests

To illustrate what the tests reveal about syntactic structures, the following two analyses of the sentence *Larry likes to drink expensive beer* are used:



Using the concept of the constituent unit established above (i.e. a complete subtree), there are six constituents in the dependency tree (1a) (6 nodes = 6 constituents) and there are eleven constituents in tree (1b) (11 nodes = 11 constituents). These constituents are listed as follows:

6 constituents in (1a)

Larry, expensive, expensive beer, drink expensive beer, to drink expensive beer, and Larry likes to drink expensive beer

11 constituents in (1b)

Larry, likes, to, drink, expensive, beer, expensive beer, drink expensive beer, to drink expensive beer, likes to drink expensive beer, and Larry likes to drink expensive beer

Thus the constituency tree (1b) assumes five more constituents than the dependency tree (1a).

A pertinent observation here concerns the status of phrases in the competing analyses. The phrasal constituents in the constituency tree (1b), those labeled with ...P, are also constituents in the dependency tree (1a), the one exception being the VP *likes to drink expensive beer*, which is not a constituent in (1a). However, four of the sub-phrasal constituents shown in (1b) (*likes, to, drink, and beer*) are not constituents in the dependency tree (1a). These observations point to a key difference in how phrases are understood across dependency and constituency structures. Most sub-phrasal constituents in constituency structures are not constituents in dependency structures to begin with, whereas most phrasal constituents in constituency structures are also constituents in dependency structures.

Most constituency tests easily identify nouns like *Larry* and noun phrases like *expensive beer* as constituents. This point is illustrated next by focusing on *expensive beer*:

Topicalization

- (2) a. ...but **expensive beer** Larry does like to drink.

Proform substitution

- b. Larry likes to drink **it**.
(*it = expensive beer*)

Answer fragments

- c. What does Larry like to drink?
– **Expensive beer**.

Clefting

- d. It is **expensive beer** that Larry likes to drink.

Pseudoclefting

- e. What Larry likes to drink is **expensive beer**.

The tests also converge identifying the nonfinite VP *drink expensive beer* as a constituent:

Topicalization

- (3) a. ?...but **drink expensive beer** Larry does like to.

Proform substitution

- b. Larry does like to **do so**.
(*do so = drink expensive beer*)

Answer fragments

- c. What does Larry like to do?
– **Drink expensive beer**.

Clefting

- d. *It is **drink expensive beer** that Larry likes to.

VP-ellipsis

- e. Sam likes to drink expensive beer, and Larry also likes to **drink expensive beer**.

Pseudoclefting

- f. What Larry likes to do is **drink expensive beer**.

Five of the six tests converge, agreeing that *drink expensive beer* should have the status of a constituent. Concerning clefting, the reason why it contradicts the other five tests is an open question.

The message currently being established is more easily arrived at if the points of agreement and disagreement are acknowledged across the two analyses. The dependency- and constituency-based analyses in trees (3a–b) agree with respect to six of the constituents shown. These six constituents are therefore not controversial, so the discussion can skip to the other five constituents, i.e. to the five constituents where the two analyses disagree. The constituency tree views *likes, to, drink, beer, and likes to drink expensive beer* as constituents, whereas the dependency tree views them as non-constituents.

The six constituency tests are almost unanimous in rejecting the status of these five strings as constituents. This point is illustrated first with the finite verb *likes*:

Topicalization

- (4) a. *...and **likes** Larry to drink expensive beer.

Proform substitution

- b. *Larry **does/so/does so** to drink expensive beer. (*does/so/does so = likes*)

Answer fragments

- c. What does Jim feel about drinking expensive beer? – ***Likes**.

Clefting

- d. *It is **likes** that Larry to drink expensive beer.

VP-ellipsis

- e. *Jim likes to drink expensive beer, and Larry **likes** to drink expensive beer.

Pseudoclefting

- e. *What Larry does concerning drinking expensive beer is **likes**.

The six tests converge; they agree that *likes* should not have the status of a constituent.

A second example solidifies the message. The tests agree that the finite VP string *likes to drink expensive beer* should not have the status of a constituent

Topicalization

- (5) a. *...and **likes to drink expensive beer** Larry.

Proform substitution

- b. [?]Sid **does so**.
(*do so = likes to drink expensive beer*)

Answer fragments

- c. What does Larry do?
– ***Likes to drink expensive beer**.

Clefting

- d. *It is **likes to drink expensive beer** that Larry does.

VP-ellipsis

- e. *Jim likes to drink expensive beer, and Larry **likes to drink expensive beer**, too.

Pseudoclefting

- f. *What Larry does is **likes to drink expensive beer**.

An analysis in terms of VP-ellipsis is not available for example (5e), although one in terms of stripping is available – the star indicates badness of VP-ellipsis. The six tests mostly converge; they mostly agree that the finite VP string *likes to drink expensive beer* should not have the status of a constituent.

There is no reason to belabor the point. The reader can extend the tests for him- or herself to the other three strings for which there is disagreement (*to*, *drink*, and *beer*). The tests further support the dependency tree (1a); they agree that these strings should not be granted the status of constituents.

To summarize, the tests point to the meaningfulness of phrases: phrases can serve as top-

ics, they can be replaced by proforms, they can be clefted and pseudoclefted, they can appear as answer fragments, and they can be elided. The tests contradict the existence of sub-phrasal constituents. Sub-phrasal constituents are an artifact of constituency-based syntax. Phrase structure grammars must posit their existence to maintain a constituency-based approach to syntactic structures. The fact that many of the most widely employed tests for constituents do not support their existence is a big problem for constituency-based syntax in general.

5 Function words

The message just delivered in the preceding section should not be controversial among DGs. The fact that dependency-based syntax is more congruent with empirical tests for syntactic structures should be a welcome insight. There are, though, points of disagreement among DGs where the tests can help. In particular, the tests can help decide points of contention when DGs disagree about the best analysis for a given structure. Indeed, the tests provide guidance concerning the status of many function words in the syntactic hierarchy. This contribution now focuses on the status of function words.

There is, namely, some disagreement concerning the best analysis of function words among DGs. Certainly the dominant position in most of the theoretically-oriented DG literature is that auxiliary verbs are heads over content verbs, adpositions are heads over their nouns, and subordinators are heads over their verbs.¹ More recently, a quite different approach to dependencies has been put forth. The Universal Stanford Dependencies (USD) (de Marneffe et al. 2014) is now advocating an annotation scheme that consistently subordinates function words to content words. Thus according to this annotation scheme, auxiliary verbs are dependents of main verbs, adpositions are dependents of nouns, and subordinators are dependents of verbs.

The USD position in this area does receive some support from Matthews (1981) and from the Prague school, both of which also subordinate auxiliary verbs to content verbs in surface syntax. Matthews and the Prague school disagree with USD concerning the status of adpositions

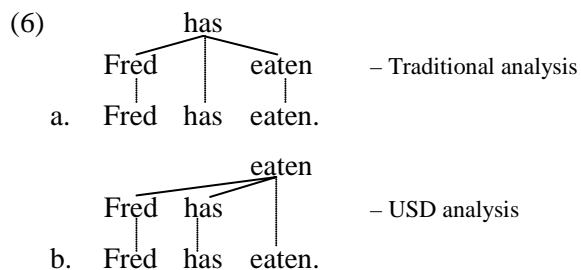
¹ The following linguists and sources all pursue the conventional analysis: Kunze 1975, Starosta 1988, Lobin 1993, Engel 1994, Jung 1995, Heringer 1996, Groß 1999, Eroms 2000, Hellwig 2003, Mel'čuk 1988, 2009, Hudson 1976, 1984, 1990, 2007, 2010.

and subordinators, however, since they position adpositions above their nouns and subordinators above their verbs.

In any case, the diagnostics for constituents discussed and illustrated above can shed light on the status of function words. In particular, they deliver strong support for the more traditional stance; they hence contradict the USD annotation scheme. The critique of USD presented below must be understood in a broader context, though. USD parsing actually advocates more than one annotation scheme; it advocates the unorthodox one just mentioned, which subordinates all function words to their associated content words, as well as two others, one of which is more traditional in that it positions most function words above the content words with which they co-occur. The points about function words established in the following two sections are directed at the former, more prominent annotation scheme of USD.

5.1 Auxiliary verbs

The traditional approach and the USD approach are contrasted with the following trees:



The analysis in (6a) shows *eaten* as a constituent, whereas the analysis in (6b) shows *has* as a constituent.

The six tests mostly converge in support of the a-analysis. They mostly agree that *eaten* is a constituent:

- (7) a. ...and **eaten** Fred certainly has.
- Topicalization
- b. Fred has **done so**. (*done so = eaten*)
- Proform substitution
- c. What has Fred done? – **Eaten**.
- Answer fragments
- d. *It is **eaten** that Fred has.
- Clefting
- e. *Sue has eaten, and Fred has **eaten**, too.
- VP-ellipsis
- f. Sue has eaten, and Fred has **eaten**, too.
- Pseudoclefting
- g. What Fred has done is **eaten**.

Five of the six tests agree that *eaten* should be viewed as a constituent. Concerning clefting, the reason why it disagrees with the other five tests is an open issue that is not addressed here.

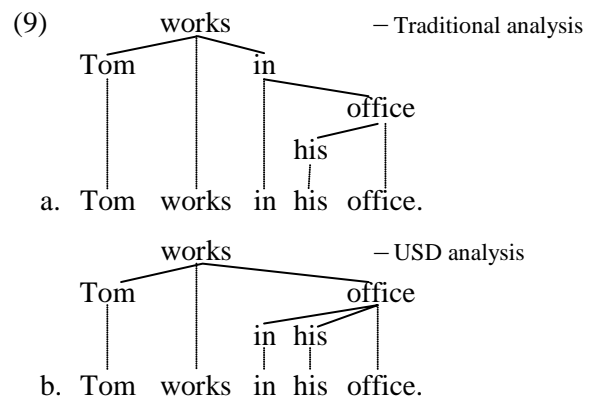
The six tests are also unanimous in their agreement that *has* is not a constituent:

- (8) a. *...and **has** Fred eaten.
- Topicalization
- b. *Fred **does so** eaten. (*does so = has*)
- Proform substitution
- c. What concerning Fred and eating?
– ***Has**.
- Answer fragments
- d. *It is **has** that Fred eaten.
- Clefting
- e. *Sue has eaten, and Fred **has** eaten, too.
- VP-ellipsis
- f. *What Fred eaten is **has**.
- Pseudoclefting

Note that the example of topicalization should maintain the declarative force of the original sentence – the star therefore indicates that the sentence cannot be construed as a statement. Based on these results, one can conclude that the six tests for constituents provide no evidence for the status of *has* as a constituent.

5.2 Prepositions

The six tests strongly support the subordination of nouns to prepositions. This point is established with the following two competing analyses of a simple sentence containing a preposition:



The traditional analysis in (9a) takes *his office* to be a constituent, whereas the USD analysis takes *his office* to be a non-constituent.

Five of the six tests agree that *his office* is a constituent:

Topicalization
(10) a. ...but **his office** Tom does work in.

Proform substitution
b. Tom works in **there/it**.
(*there/it = his office*)

Answer fragments
c. What (room) does Tom work in?
– **His office**.

Clefting
d. It is **his office** that Tom works in.

Pseudoclefting
e. The room Tom works in is **his office**.

VP-ellipsis is not applicable in this case because no verb is involved. The other five tests agree that *his office* should be viewed as a constituent.

The USD analysis shown with (10b) takes the preposition *in* alone to be a constituent. The tests are unanimous, however, insofar as *in* alone is not a constituent:

Topicalization
(11) a. *...but **in** Tom works his office..

Proform substitution
b. *Tom works **there** his office.
(*there = in*)

Answer fragments
c. What does Tom do concerning working and his office? – ***In**.

Clefting
d. *It is **in** that Tom works his office.

Pseudoclefting
e. *Where Tom works his office is **in** .

Based on these results, there is no motivation for granting the preposition *in* the status of a constituent.

In sum, the five applicable diagnostics clearly support the traditional analysis of prepositions: they are heads over their nouns.

5.3 Subordinators and determiners

Reaching a conclusion about the hierarchical status of subordinators and determiners using the six tests for constituents is much more difficult to do, because the tests typically do not support any analysis at all, at least not when applied to English sentences. In this respect other considerations must be accessed to help determine the hierarchical status of these two additional types of function words.

Concerning subordinators (e.g. *after, because, before, if, that, when, where, whether, why*, etc.), the fact that a couple of them also serve as prepositions is an indication that they should receive a similar analysis as prepositions; the subordinators *before, after, with, and for* also serve as simple prepositions. Thus since there is strong evidence supporting the status of prepositions as heads over their nouns, the same sort of analysis can be extended to these subordinators, and then by analogy to subordinators in general.

Concerning determiners, however, the debate concerning their status in the syntactic hierarchy is ongoing. This debate has split the syntax world into two camps since the 1980s: determiner phrase (DP) vs. noun phrase (NP). For the most part, the six tests for constituents do not shed much light on this debate, since they in general fail to identify either determiners or their nouns as constituents.

There are, however, a couple of limited cases that one can interpret as evidence in favor of the traditional NP analysis, a point now illustrated here using the sentence *Susan's house is beautiful*:

Proform substitution
(12) a. **Her** house is beautiful.
(*her = Susan's*)

Answer fragment
b. Whose house is beautiful? – **Susan's**.

These two examples demonstrate that proform substitution and answer fragments can be interpreted as identifying the determiner *Susan's* as a constituent. The other four tests (topicalization, clefting, VP-ellipsis, and pseudoclefting) do not support these results, however. Furthermore, the answer fragment in (12b) can be seen as involving noun ellipsis (N-ellipsis); the noun *house* has been elided, leaving just the determiner. This observation weakens any conclusion about the constituenthood of the determiner *Susan's* based on (12b).

In sum then, the hierarchical analysis of prepositions can be extended to subordinators, since there is much overlap in the forms and distributions of these two classes of function words. Concerning determiners, however, the tests deliver only rather weak evidence for the position that they are dependents of their nouns.

6. Other languages

An objection can be raised against the reasoning produced above. This objection points to the English-centered focus of the diagnostics discussed. The data produced have been from English alone. This fact raises the concern that the conclusion may not extend to other languages, and thus the diagnostics for constituents may not be very insightful from a cross-linguistic perspective. This objection is conceded here, but only in part.

There are a couple of points to keep in mind when assessing the objection. The first is that the most prominent schools of syntax internationally have been founded and are/were led primarily by native speakers of English (e.g. Noam Chomsky, Ivan Sag, Carl Pollard, Joan Bresnan, Ronald Langacker, etc.). The arguments and insights of these linguists are produced primarily in English, using examples primarily from English. Thus the syntax of English has had and continues to have a far greater influence on our understanding of syntax on the international stage in general than that of any other language. In this regard, the fact that tests for constituents developed for English sentences contradict the syntactic theories of the schools of syntax just alluded to should carry a lot of weight.

The second point to keep in mind concerns the sources that are using the tests. The textbooks that employ the tests are intended for students of linguistics. These texts are then used around the world in numerous countries by students of English in language departments at colleges and universities. Thus often the first exposure to syntactic theory that aspiring linguists receive comes in the form of textbooks written in English, using primarily English examples. This situation is suggestive of the great influence that these texts are having on the development of syntactic theory internationally. The message, then, is again that the tests as applied to English are having a disproportionate influence on the development and direction of syntactic theory in general.

The third point to consider is the extent to which the tests are in fact applicable to other languages. Some of the tests employed above should be valid for many other languages. This is particularly true of proform substitution and answer fragments. Most if not all languages have proforms, and these proforms can be used to identify syntactic structure in a manner similar to how proform substitution has been employed above. Similarly, most if not all languages allow

question-answer pairs and the answer fragments that are produced can deliver important clues about syntactic structure no matter the language.

Ideally, each language needs to develop its own inventory of diagnostics for syntactic structure, based on its idiosyncrasies. Certainly some of the diagnostics above can be adopted directly into other languages (proform substitution, answer fragments), and others can perhaps be adapted in one way or another so that they can also be employed (clefting, pseudoclefting, ellipsis). When a given diagnostic does not seem to provide insights about syntactic structure, one should ask why this is so. The fact that the diagnostic is not helpful can then serve as an indicator about what is going on with the particular syntax of that language.

7. Concluding points

To conclude this contribution, two further objections that come to mind against the reasoning developed above are briefly countered. The first of these concerns the fact that diagnostics for constituents are fallible; at times the results they deliver are contradictory. This is perhaps most evident with determiners in English. Dependency- and constituency-based theories of syntax alike view determiners as constituents, yet most of the tests above fail to identify them as such. While this point must be conceded, at no time has the presentation above claimed that the diagnostics are infallible. Indeed, the tests are at times quite fallible. But what this contribution has claimed is that most diagnostics for constituents consistently fail to identify sub-phrasal strings as constituents. Since this is precisely what dependency-based models predict, the dependency models are preferable in this area. On the whole, they make much more accurate predictions about sentence structure with much less effort.

The second further objection that can be raised against the messages delivered above concerns the critique of the USD annotation scheme. No attempt has been made here to refute the main motivation for the USD scheme, this motivation being uniformity of annotation across diverse languages. Subordinating function words to content words establishes hierarchies of content words that are directly linked to each other, and these hierarchies are then relatively consistent across diverse languages. While this objection must also be conceded, this concession should not be misinterpreted, since this contribu-

tion never intended to refute this supposed strength of the USD annotation scheme.

The authors of the USD scheme claim that USD embodies “linguistic quality” (de Marneffe et al. 2014: 4589) – as opposed to accuracy of parsing. The message delivered above is that diagnostics for constituents contradict this claim to linguistic quality. Indeed, the diagnostics reveal the opposite, namely that the USD scheme cannot claim linguistic quality concerning the tests. Given the prominent role that the tests play in modeling syntactic structures, the lack of linguistic quality is in fact a major drawback of the USD approach.

References

- David Adger. 2003. *Core Syntax: A Minimalist Approach*. Oxford University Press.
- Adrian Akmajian, Richard Demers, Ann Farmer, and Robert Harnish. 1990. *An Introduction to Language and Communication*, 3rd edition. The MIT Press, Cambridge, MA.
- David Allerton. 1979. *Essentials of Grammatical Theory: A Consensus View of Syntax and Morphology*. Routledge & Kegan Paul, London.
- Carl Baker. 1978. *Introduction to Generative-Transformational Syntax*. Prentice-Hall.
- Kersti Börjars and Kate Burridge. 2001. *Introducing English Grammar*. Arnold, London.
- Robert Borsley. 1991. *Syntactic Theory: A Unified Approach*. Edward Arnold, London.
- Noel Burton-Roberts. 1986. *Analysing Sentences: An Introduction to English Syntax*, 2nd edition. Longman, London.
- Andrew Carnie. 2010. *Constituent Structure*, 2nd edition. Oxford University Press, Oxford, UK.
- Andrew Carnie. 2013. *Syntax: A Generative Introduction*, 3rd edition. Wiley-Blackwell, Malden, MA.
- Noam Chomsky. 1957. *Syntactic Structures*. Mouton Publishers, The Hague.
- Elizabeth Cowper. 1992. *A Concise Introduction to Syntactic Theory: The Government-Binding Approach*. The University of Chicago Press, Chicago.
- Peter Culicover. 2009. *Natural Language Syntax*. Oxford University Press, Oxford, UK.
- Peter Culicover and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford University Press, New York.
- Ulrich Engel. 1994. *Syntax der deutschen Sprache*, 3rd edition. Erich Schmidt Verlag, Berlin.
- Hans-Werner Eroms. 2000. *Syntax der deutschen Sprache*. de Gruyter, Berlin.
- Nigel Fabb. 1994. *Sentence Structure*. Routledge, London.
- Victoria Fromkin (ed.). 2000. *Linguistics: An Introduction to Linguistic Theory*. Basil-Blackwell, Malden, MA.
- Thomas Groß. 1999. *Theoretical Foundations of Dependency Syntax*. Iudicium, Munich.
- Liliane Haegeman. 2006. *Thinking Syntactically: A Guide to Argumentation and Analysis*. Blackwell Publishing, Malden, MA.
- Liliane Haegeman and Jacqueline Guéron. 1999. *English Grammar: A Generative Perspective*. Blackwell Publishers, Oxford, UK.
- Peter Hellwig. 2003. Dependency Unification Grammar. In: Vilmos Ágel et al. (eds.), *Dependency and Valency: An International Handbook of Contemporary Research*, 593–635. Walter de Gruyter, Berlin.
- Hans Heringer. 1996. *Deutsche Syntax: Dependentiell*. Stauffenburg, Tübingen.
- Rodney Huddleston and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press.
- Richard Hudson. 1984. *Word Grammar*. Basil Blackwell, Oxford, UK.
- Richard Hudson. 1990. *An English Word Grammar*. Basil Blackwell, Oxford, UK.
- Richard Hudson. 2007. *Language Networks: The New Word Grammar*. Oxford University Press, Oxford, UK.
- Wha-Young Jung. 1995. *Syntaktische Relationen im Rahmen der Dependenzgrammatik*. Buske Verlag, Hamburg.
- Samuel Keyser and Paul Postal. 1976. *Beginning English Grammar*. Harper & Row, New York.
- Jong-Bok Kim and Peter Sells. 2008. *English Syntax: An Introduction*. CSLI Publications, Stanford.
- Paul Kroeger. 2005. *Analyzing Grammar: An Introduction*. Cambridge University Press.
- Jürgen Kunze. 1975. *Abhängigkeitsgrammatik*. *Studia Grammatica* 12. Akademie Verlag, Berlin.
- Howard Lasnik with Marcela Depiante and Arthur Stepanov. 2000. *Syntactic Structures Revisited: Contemporary Lectures on Classic Transformational Theory*. The MIT Press, Cambridge, MA.
- Henning Lobin. 1993. *Koordinationsyntax als prozedurales Phänomen*. *Studien zur deutschen Grammatik* 46. Gunter Narr Verlag, Tübingen.

- Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silvaire, Katrin Haverinen, Filip Ginter, Joakim Nivre, Christopher D. Manning. 2014. Universal Stanford Dependencies: A cross-linguistic typology. *LREC* 14, 4585–92.
- James McCawley. 1998. *The Syntactic Phenomena of English*, 2nd edition. University of Chicago Press, Chicago.
- Igor Mel'čuk. 1988. *Dependency Syntax: Theory and Practice*. State University of New York Press, Albany.
- Stefan Müller. 2010. *Grammatiktheorie*. Stauffenberg Verlag, Tübingen.
- Donna Napoli. 1993. *Syntax: Theory and Problems*. Oxford University Press, New York.
- Jamal Ouhalla. 1994. *Transformational Grammar: From Rules to Principles and Parameters*. Edward Arnold, London.
- Thomas Payne. 2006. *Exploring Language Structure: A Student's Guide*. Cambridge University Press.
- Geoffrey Poole. 2002. *Syntactic Theory*. Pelgrave, New York.
- Paul Postal. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge University Press, Cambridge, MA.
- Radford, Andrew. 1988. *Transformational Grammar: A First Course*. Cambridge University Press, Cambridge, UK.
- Andrew Radford. 1997. *Syntactic Theory and the Structure of English: A Minimalist Approach*. Cambridge University Press, Cambridge, UK.
- Andrew Radford. 2004. *English Syntax: An Introduction*. Cambridge University Press.
- Ian Roberts. 1997. *Comparative Syntax*. Arnold, London.
- Ivan Sag, Thomas Wasow, and Emily Bender. 2003. *Syntactic Theory*, 2nd edition. CSLI Publications, Stanford.
- Nicholas Sobin. 2011. *Syntactic Analysis: The Basics*. Wiley-Blackwell, Malden, MA.
- Dominique Sportiche, Hilda Koopman, and Edward Stabler. 2014. *An Introduction to Syntactic Analysis and Theory*. Wiley Blackwell, Malden, MA.
- Stanley Starosta. 1988. *The Case for Lexicase: An Outline of Lexicase Grammatical Theory*. Pinter Publishers, London.
- Maggie Tallerman. 2005. *Understanding Syntax*, 2nd edition. Hodder Education, London.
- Linda Thomas. 1993. *Beginning Syntax*. Blackwell, Oxford, UK.
- Robert van Valin. 2001. *An Introduction to Syntax*. Cambridge University Press, Cambridge, UK.