



Universiteit
Leiden

The Netherlands

Syntactic diagnostics in the study of human language

Cheng, L.L.; Corver, N.; Cheng, L.L.S.

Citation

Cheng, L. L., & Corver, N. (2013). Syntactic diagnostics in the study of human language. In L. L. S. Cheng (Ed.), *Diagnosing syntax* (pp. 1-17). Oxford: Oxford University Press.
doi:10.1093/acprof:oso/9780199602490.003.0001

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3245250>

Note: To cite this publication please use the final published version (if applicable).

Syntactic Diagnostics in the Study of Human Language

LISA LAI-SHEN CHENG AND NORBERT CORVER

1.1 Gnosis through diagnosis

Scientists try to find out what is the nature of things, that is, what is their internal hidden structure. In order to determine this structure, the scientist tries to identify properties and characteristics of the object of study. This process of identifying the nature ('the hidden structure') of things on the basis of their properties and characteristics (i.e. signs) is called *diagnosis*. The conclusion reached at the end of this discovery process is called *a diagnosis*. Thus, *γνωσις* ('knowledge, insight') is obtained by *διαγνωσις* ('discernment, the ability to discern'): the nature of an object or phenomenon is understood by means of the ability to discern relevant features of that object or phenomenon.

The term *diagnosis* is probably most familiar from the medical sciences. A physician must know the human body and its functions in terms of normality (so-called homeostasis). For understanding homeostasis, the physician must know four central domains of diagnostic medicine: anatomy (the structure—architecture—of the human body), physiology (how the body works), pathology (what can go wrong with the anatomy and physiology), and psychology (thought and behaviour). Once the physician has knowledge of what is normal (e.g. the normal architecture and functioning of the heart or circulatory system), he or she can then determine a patient's particular (degree of) departure from homeostasis. This determination of the patient's bodily state in comparison with the homeostatic state is called a diagnosis. The physician diagnoses a physical condition, disease, or disorder by identifying medical *signs*—that is, objective indications of some medical fact or quality, for example, hypertension, clubbing of the fingers, elevated blood pressure—during a physical examination of a patient and by interpreting the patient's *symptoms*—that is, the subjective experiences ('complaints which indicate a disease') of the patient, such as fatigue, pain, nausea, or loss of vision. An important tool in determining the

correct diagnosis is the use of medical *tests* (i.e. *diagnostic procedures*), such as measuring blood pressure, checking the pulse rate, sensation and muscle testing, urine and blood tests, and ECG (electrocardiogram).

Diagnosis is also part of the linguistic sciences. Linguists try to discover the nature of human language by identifying its characteristic properties. This search for diagnostic properties can be found in different approaches to the study of human language. Traditional grammar provided us with diagnostic features for identifying parts of speech (i.e. the categorial nature of a lexical item) and identifying grammatical functions such as subject and object. For example, a lexical item can be identified as being 'verbal' on the basis of the morphological *sign* of tense specification, and a noun phrase can be identified as fulfilling the role of 'subject' on the basis of its morphosyntactic *sign* of agreement with the finite verb. Also in (American) structuralist grammar, linguistic diagnosis played an important role. With the aim of making the study of language a scientific study, one tried to establish a set of structural procedures which could describe a language completely in terms of distributional relations. For example, a lexical item could be identified as being nominal on the basis of its 'distributional' *sign* of being able to combine with a determiner. Important structuralist contributions to the study of human language, more specifically syntax, are the well-known constituency tests on the basis of which the syntactic unity of a string of words can be diagnosed. The syntactic *signs* of preposing (1b), substitution (1c), and coordination (1d), for example, indicate that the string of words *the old man* in (1a) forms a constituent:

- (1) a. She recognized *the old man*.
- b. *The old man* she recognized.
- c. She recognized *the old man* and I recognized *him* too.
- d. She recognized *the old man and his dog*.

One may consider diagnostics to yield a categorization, that is, a grouping of objects or phenomena that share certain properties. In the discussion above, we have indicated diagnostics to put 'nouny' things into one category and 'verby' things into another category based on diagnostic criteria like tense specification and agreement. The same can be done with respect to phenomena. Based on the diagnostic criterion of subject- V_{fin} -inversion, one may distinguish English *wh*-questions (*How smart is he?*) from English *wh*-exclamatives (*How smart he is!*), and based on the possible presence of a *by*-phrase one may distinguish passive constructions (*The car was washed (by my children)*) from middle constructions (*This car washes easily (*by children)*).

As is well-known, categorization is not always rigid and straightforward due to the fact that there may be conflicting signs between two objects or phenomena that appear to belong to the same class. Ross (1972, 1973), for example, argued that the traditional view of syntactic categories as being discrete elements that can be rigidly

distinguished from each other is incorrect. Instead of a fixed, discrete inventory of syntactic categories, he proposed a quasi-continuum, where the distinction between one category and another one is not discrete, but ‘squishy’, that is, one of degree. According to this view, syntactic categories are not, or not necessarily, distinguished from each other by the absolute presence versus absence of some property. Categories, rather, differ from each other in relative terms: that is, category α (say V) may have property X to the highest degree, category β (say A) to a lesser degree, while for category δ (say N), this property is completely absent. Along the same lines, phenomena or constructions may be more similar or less similar depending on the properties they share. For example, even though passives and middles differ from each other in the presence versus absence of a *by*-phrase, they share the property of a ‘suppressed’ subject argument. This makes middles closer to passives than, for example, interrogatives (e.g. *Who washed the car?*), where argument-suppression does not play any role.

1.2 Diagnosis in generative grammar

Diagnosis is also part of generative grammar, which aims at identifying the (mental) architecture of the faculty of language, or as Noam Chomsky calls it: ‘a language organ’.

The faculty of language can reasonably be regarded as a ‘language organ’ in the sense in which scientists speak of the visual system, or immune system, or circulatory system, as organs of the body. Understood in this way, an organ is not something that can be removed from the body, leaving the rest intact. It is a subsystem of a more complex structure. We hope to understand the full complexity by investigating parts that have distinctive characteristics, and their interactions. Study of the faculty of language proceeds in the same way. (Chomsky 2000b: 4)

Being a subsystem of a more complex structure implies that the ‘language organ’ interfaces with other systems. In order to understand the nature (i.e. the hidden architecture or design) of the language system, one needs to identify the properties of human language as they occur at the interface with these other systems. Starting from the Aristotelian dictum ‘Language is sound with meaning’, Chomsky (1993, 1995b) distinguishes two types of performance systems that access the information provided by the language organ: the sensorimotor (SM) systems and the conceptual-intentional (CI) systems. The former access the linguistic sound information which is provided by the phonological (PF) representation (the representation of sound within the language organ). The latter access the linguistic meaning information as provided by the semantic (LF) representation (the representation of (structure-related) meaning within the language organ). In order to be accessible to the interface systems, the sound and meaning properties must be legible (interpretable) by the language-external systems. Thus, for our understanding of the nature of the

human language organ, it is crucial to discern, that is to diagnose, the *interface properties* of human language, that is to say, the linguistic properties that provide instructions to the systems with which the language organ interacts. *Interface diagnosis* provides insight into the nature of human language by studying it from the perspective of its interaction with other, language-external, systems.

Besides the interfacing dimensions of sound and meaning, there is a third core dimension of human language, viz., the dimension which is hidden behind the word *with* in the dictum 'Language is sound with meaning' (cf., Chomsky 2010) and which stands for the structural dimension (syntax) of human language. Linguistic expressions do not only have sound and meaning, they also have a syntax. In other words, they are structured objects. A major concern of generative syntax has been, and still is, to determine the nature of these syntactic structures. More specifically, what is the nature of the symbols (the computational atoms) that function as constituents within the structured object, and what is the nature of the computational system (the generative procedure) that combines the atoms into complex syntactic structures. Answers to these questions are obtained through a process of *computational diagnosis*, that is, the process of obtaining insight into the computational system of human language and the syntactic structures generated by this system by identifying the computational properties of human language.

In trying to determine the nature of human language, linguists are confronted with an important characteristic of the language organ: its diversity. On the surface, it looks as if the language organ of a speaker of English is different from the language organ of a speaker of Chinese. For example, the former has a *wh*-question formation rule which places a *wh*-phrase at the beginning of the sentence, while the latter simply leaves the *wh*-phrase in its base position (i.e. *wh*-in-situ). In order to obtain knowledge about the dimensions of cross-linguistic diversity, the linguist tries to identify the properties along which languages may vary. If linguists examine surface phenomena, languages appear to vary in an unbounded and arbitrary way. A guiding assumption of generative-linguistic research, however, is that cross-linguistic variation is to a large extent bounded and non-arbitrary. In other words, there is a system of variant properties (so-called parameters; cf., Chomsky 1981; Baker 2001). Research on the nature of cross-linguistic variation involves the identification of variant properties in human language. At the same time, it is important to examine whether diagnostic tools developed for one language can be extended to other languages. Thus, *cross-linguistic diagnosis* can take place along two different dimensions.

Linguistic diagnosis is obviously not restricted to the language of typical (i.e. 'normal') adult speakers. Obtaining knowledge about the language organ is also possible through diagnostic investigation of the growth of the organ (developing grammars) and the malfunction of the organ as a result of brain injury (aphasia). Also the use of language in production or comprehension allows us to identify

properties of the language organ. We will characterize this domain of diagnosis as *psycho-/neuro-linguistic diagnosis*.

In sum, generative linguistics aims to understand the nature of the human language faculty. In order to do this, it tries to identify properties and characteristics of the language organ in its various states ('normal adult state', different states of growth, pathological states), in its various uses (production and comprehension), and on the basis of its internal system (atoms and computation), its interaction with systems external to the language organ (interfaces), and the dimensions of variation across different language-particular grammars (i.e. parameterization).

In this book, the study of syntactic diagnostics will be undertaken from these different angles. More in particular, syntactic diagnosis will be investigated from the following points of view:

- Syntax (computational diagnosis)
- Syntax at the interface (interface diagnosis)
- Psycho-/neuro-linguistics (psycho-neuro-linguistic diagnosis)
- Language diversity (cross-linguistic diagnosis)

From each of these four diagnostic perspectives the following five core domains of natural language syntax will be investigated in the various chapters of this book:

- Head movement (Part I)
- Phrasal movement (Part II)
- Agreement (Part III)
- Anaphora (Part IV)
- Ellipsis (Part V)

Thus, each part will contain four contributions on a core syntactic domain, each from a different (diagnostic) perspective. Every theme will be closed off by a brief chapter which, starting from the four contributions, reflects on the issue of diagnosis in that specific syntactic domain. This chapter addresses questions such as: Which diagnostic properties are reliable, which ones are not? Where do we find convergence/divergence on the status of some diagnostic property? To what extent does some phenomenon which used to be diagnosed as being syntactic really belong to that domain of the grammar?

1.3 Some examples of linguistic diagnosis in generative grammar

From a historical perspective, it seems appropriate to give some illustrations of the search and use of syntactic diagnostics in generative linguistics. Arguably, one of the finest illustrations of syntactic diagnosis in generative linguistics is Chomsky's (1977b) 'On wh-movement'. In this article, Chomsky shows how the application of

the syntactic operation *wh*-movement can be identified on the basis of the following diagnostic properties (i.e. *syntactic signs*):

- (2) a. It leaves a gap.
b. Where there is a bridge, there is an apparent violation of Subjacency, the Propositional Island Constraint, and the Specified Subject Constraint.
c. It observes the Complex NP Constraint.
d. It observes the *Wh*-island Constraint.

Property (2a) follows from the assumption that *wh*-movement moves a phrase, which in turn implies that a gap (i.e. a category voided of phonetic content) is left behind in the extraction site. Property (2b) refers to the fact that *wh*-movement can circumvent a violation of Subjacency, the Propositional Island Constraint and the Specified Subject Constraint in structural contexts involving so-called bridge verbs. An important assumption here is that *wh*-movement can proceed successive-cyclically (i.e. stepwise on successive cycles, making use of all the intermediate COMP-positions that separate the extraction site from the ultimate landing site). The island properties (2c) and (2d) are implied by Subjacency.

Chomsky formulates the issue about the diagnostic function of the *wh*-movement properties in (3) as the central research question of 'On *wh*-movement' (Chomsky 1977b: 86, (50)).

- (3) 'Where we find the configuration [(2)] in some system of data, can we explain it on the assumption that the configuration results from *wh*-movement?'

The aim of 'On *wh*-movement' is thus to investigate whether or not the configuration in (2) can serve as a diagnostic for *wh*-movement. Much of the article then consists of demonstrating that each of the clausal constructions in (4) has (fully or to a great extent) the properties in (2).

- (4) a. They have more enemies than we have.
b. This book, I really like.
c. It is this book that I really like.
d. I wonder who to see.
e. I found a book for you to read.
f. John is tall enough for us to see.
g. John is easy (for us) to please.

Let us illustrate the diagnostic method by means of the infinitival indirect interrogative sentence in (4d), which features the *wh*-phrase *who* in the left periphery of the embedded clause. The relevant diagnostic patterns of *wh*-movement are given in (5), corresponding to the diagnostic properties in (2):

- (5) a. I wonder [who_i to see t_i]
 b. I wonder [who_i to order Mary [to promise [to visit t_i]]]
 b'. I wonder [who_i to persuade Mary [that she should promise [to visit t_i]]]
 c. *I wonder [who_i to insist on [NP the principle [S' that Bill should visit t_i]]]
 d. *I wonder [what $_i$ to ask her [when $_j$ to buy t_i t_j]]

Example (5a) illustrates that movement of *who* leaves behind a gap in the direct object position after *to see*. (5b) and (5b') show that *who* can be moved from the complement position of *visit* to the COMP position of the highest embedded clause thanks to the bridge properties of the verbs *to promise*, *to persuade*, and *to order*. Examples (5c) and (5d) show, respectively, that infinitival indirect questions obey the Complex NP Constraint and the *Wh*-island constraint.

In an infinitival indirect question like (4d)/(5a), a *wh*-word, *who*, is overtly present at the surface, which, of course, is very helpful for identifying *wh*-movement. Identifying *wh*-movement in constructions lacking an overt *wh*-phrase is arguably more exciting, since there is no direct (overt) element signalling the presence of a displaced *wh*-phrase. Take, for example, a comparative construction like (6a). Even though no overt *wh*-expression is present in the comparative clause, the *wh*-diagnostics lead us to analyse this construction in terms of *wh*-movement: that is, a non-overt *wh*-phrase is moved to COMP (Spec,CP in present-day terms). Consider in this regard the following examples, where *t* (trace) indicates the original extraction site:

- (6) a. John met more linguists [than you met t]
 b. John met more linguists [than we thought [you said [Bill believed [Sue met t]]]]]
 c. *John met more linguists than I believed [NP the claim [S' that Sue had met t]]]
 d. *John met more linguists than I wondered [S' when Sue had met t]

Examples (6a–d), again, illustrate the diagnostic properties of *wh*-movement: (6a) illustrates the gap property; (6b) shows that an apparently unbounded dependency can be created if the necessary bridge requirements are fulfilled; (6c) exemplifies the Complex NP Constraint effect; and (6d) shows that comparative formation is subject to the *wh*-island constraint. The presence of the configuration of *wh*-properties suggests that a rule of *wh*-movement is also involved in the derivation of comparatives. According to 'On *wh*-movement', this rule moves an (overt) *wh*-phrase to COMP where it is subsequently deleted by a (local) *wh*-phrase deletion rule. This example from comparatives nicely shows that the diagnostic method can help us identify properties of linguistic expressions (*in casu* comparatives) which are not directly manifested at the sound surface (because of the silence of the displaced element).

The *wh*-diagnostics applying to comparatives provide a nice illustration of how it is possible to diagnose the presence of the computational rule of *wh*-movement on the basis of a number of *wh*-properties and without the presence of *wh*-elements. Another well-known example of syntactic diagnosis comes from Perlmutter's (1978) pioneering work on the set of verbal predicates that are traditionally described as intransitive verbs (see also Burzio 1986). Perlmutter divides 'intransitives' into two natural classes: unergative predicates and unaccusative predicates. The 'subject'—'subject' in the sense of the noun phrase that bears nominative Case and agrees with the finite verb—of the former starts out as an external argument of the verbal predicate, while the 'subject' of the latter starts out as an internal argument of the verbal predicate. Research on the so-called unaccusativity hypothesis led to the identification of *diagnostic signs for unaccusativity* (unaccusativity diagnostics) that could be used in diagnosing the verbal status of 'intransitive' verbs. These diagnostics include: (i) auxiliary selection in certain languages (e.g. Dutch *hebben* 'to have' with unergatives versus *zijn* 'to be' with unaccusatives); (ii) the (im)possibility of forming impersonal passives; and (iii) the (im)possibility of participle–adjective conversion. These are illustrated in examples (7)–(9) in Dutch.

- (7) a. Jan heeft gelachen
 Jan has laughed
 b. Jan is gearriveerd
 Jan has arrived
- (8) a. Er werd gelachen
 there was laughed
 'There were people laughing'
 b. *Er werd gearriveerd
 there was arrived
- (9) a. *de gelachen jongen
 the laughed boy
 'the boy who laughed'
 b. de gearriveerde reizigers
 the arrived travellers
 'the travellers who arrived'

A third illustration of the use of diagnostic criteria comes from the distinction between lexical categories and functional categories. Abney (1987) lists a number of characteristic properties that have been noticed for functional categories:

(10) **General properties of functional heads**

- a. They constitute closed lexical classes.
- b. They lack descriptive content.
- c. They can be sisters only to one kind of category.
- d. They are generally phonologically and morphologically dependent.
- e. They are usually inseparable from their sister projection.

On the basis of these criteria, it is clear that a definite article like *the* falls within the class of functional atoms: it is the only member of the set of definite articles in English and clearly constitutes a closed class; it has an abstract meaning and lacks 'concrete' descriptive contents; it typically combines within nominal expressions (*the car*; **the about*, **the envious*, **the destroys*); it typically does not get stress and cannot occur independently (**I met the*); it is inseparable from its complement (**Intelligent students I met the* or **The I met intelligent students*).

1.4 On the quality of diagnosis

The quality of diagnosis is, of course, dependent on the quality of the diagnostic features. In the medical sciences, a distinction is made between symptoms and signs. The former are used for bodily features that are noticed by the patient (e.g. pain on the chest, fatigue) and consequently have a more vague and more subjective status. The notion 'sign' is used for bodily features (e.g. blood cell count, elevated blood pressure) that have been identified by the physician, possibly through the use of diagnostic medical techniques, and consequently have a more objective status. Thus, the distinction between symptoms and signs is based on the source of the information. A different way of looking at this distinction is based on the use that the information serves. As King (1982) puts it:

The modern usage of signs and symptoms emphasizes merely the source of the information, which is not really too important. Far more important is the use that the information serves. If the data, however derived, lead to some inferences and go beyond themselves, those data are signs. If, however, the data remain as mere observations without interpretation, they are symptoms, regardless of their source. Symptoms become signs when they lead to an interpretation. (King 1982: 89)

Borrowing the notions of 'symptoms' and 'signs' from the medical sciences, the question may be raised as to whether some stated diagnostic syntactic feature has the status of a symptom or the status of a sign. That is, how objective is some syntactic property as a diagnostic criterion? For example, the apparent unboundedness of *wh*-movement may be called a symptom. It is the theoretical concept of successive cyclicity (i.e. the interpretation of long-distance movement in terms of a theory of

local, successive cyclic steps) which makes it an interpretable/meaningful diagnostic property, that is, a *sign*.

The quality of diagnosis arguably also relates to what we may call the granularity of the diagnostic system. That is, the more fine-grained the diagnostic system is, the better the diagnosis of phenomena. Consider, for example, the surfacy diagnostic (2a). To what extent is the presence of a gap a truly distinguishing feature of *wh*-movement? Gaps in the sense of phonetically empty 'material' (i.e. *e*) also occur in other types of displacement environments (e.g. A-movement or head movement; see (11)) and even in non-displacement contexts, such as control structures (see however, Hornstein (2001)) and ellipsis (see (12)). As such, the presence of a gap is not a truly distinguishing feature of *wh*-movement.

- (11) a. John_i seems [*e*_i to sleep]
 b. Will_i he *e*_i sleep?
- (12) a. John_i tries [*e*_i to sleep]
 b. A: Did he sleep? B: He tried to *e*.

Therefore, it is important to reach a deep(er) level of understanding of the diagnostic tool at issue. For example, distinguishing different types of gaps—such as traces (copies), PROs, and ellipsis sites resulting from deletion of material—may turn a surfacy diagnostic (a symptom) into a more fine-grained and meaningful diagnostic property (a sign). *Wh*-movement is not simply characterized by the presence of a gap, but by the presence of a *wh*-trace (Chomsky 1973), which in the current Copy theory of movement (Chomsky 1993) is a copy of the element moved to Spec,CP that is deleted in the phonological component. In short, a coarse-grained diagnostic system based on the surfacy notion of gap may not be very helpful in understanding linguistic constructions or phenomena. A more fine-grained system in which different types of empty categories, each with its own properties, are distinguished yields a better diagnosis. In short, granularity matters in linguistic diagnosis.

A further illustration that granularity matters in syntactic diagnosis comes from the distribution of *wh*-phrases. In *wh-ex-situ*-languages, *wh*-phrases typically occupy a structural position in the left periphery of the clause (Spec,CP) as a result of (overt) *wh*-movement (i.e. the displacement property; (I)nternal-Merge). This left-peripheral placement of the *wh*-phrase, a word order property, may be used as a diagnostic for identifying *wh-ex-situ*-constructions. It has been argued, however, that in certain *wh-ex-situ*-languages there are *wh*-phrases in the left periphery of the clause that do not occupy this position as a result of overt *wh*-movement (I-Merge) but simply as a result of base-generation (i.e. External Merge) in this left peripheral position. More specifically, in various studies it has been argued that this holds for reason *wh*-adverbials (Rizzi 1990; Collins 1991; Hegarty 1992). In short, a diagnostic system which treats all *wh*-phrases (e.g. *who*, *what*, *how*, *why*, ...) on a par may not

be fine-grained enough. A more precise diagnosis can be made if a higher level of granularity can be obtained within the class of *wh*-elements.

It goes without saying that the fixation of some grammatical property as a diagnostic sign is also a matter of successful use by the linguistic practitioner and collective reasoning by the community of professionals. As Pesetsky puts it in his contribution to this volume (Chapter 7): ‘the only difference between an *argument for X* and a *diagnostic for X* is confidence’. The more confident we feel about a previous finding, the more likely we are to call it a diagnostic. When within the community of syntacticians, there is overall agreement about the ‘identifying value’ of some grammatical feature *X*, then this diagnostic feature *X* will be an important diagnostic tool in future grammatical analysis. As such, exchange of thoughts on what are reliable diagnostic features (i.e. signs) of syntactic phenomena in human language(s) leads to a certain amount of standardization (standardized diagnosis) within the study of natural language syntax.

Knowledge and command of the diagnostic procedure in one’s field of expertise also presupposes (i) knowledge of *conflicting* diagnosis (i.e. where does one and the same diagnostic test yield conflicting results, or where do two diagnostic criteria yield opposite results?) and (ii) knowledge of *non*-diagnostics (i.e. which property is an apparent and not a real sign of some syntactic phenomenon or construction?). As for the latter, we refer the reader to the contribution by van Craenenbroeck and Lipták (Chapter 25, this volume), who show that the restriction of sluicing to *wh*-questions is not a reliable test for diagnosing this construction cross-linguistically. As for conflicting diagnosis, we might refer the reader to the use of coordination as a test for identifying constituents (see Carnie 2008: 21). It is generally assumed that only constituents can be coordinated. Thus, (13a) is well-formed since *sells expensive cars* and *rents cheap trucks* each form a syntactic unit (i.e. VP). Example (13b), on the contrary, is ill-formed since *sells expensive* does not form a constituent.

- (13) a. Peter [sells expensive cars] and [rents cheap trucks]
 b. *Peter [sells expensive] and [rents cheap cars]

The coordination test, however, might hint at the constituency of the subject-DP and the verb in sentences like (14), that are known as right-node-raising sentences.

- (14) [Peter sells] and [Bill rents] expensive cars

Thus, the coordination test appears to provide conflicting evidence at the surface. The coordination pattern in (13a) suggests that the verb and its direct object-DP form a constituent whereas the coordination pattern in (14) hints at the constituency of the verb and the subject-DP. Fortunately, one generally need not rely on a single property for diagnosing the syntax of some construction or phenomenon. In this particular case, evidence from other constituency tests supports the constituency of *V* and the direct object-DP. Consider, for example, the following examples:

- (15) a. Peter says he will *sell expensive cars* and [*sell expensive cars*]_i he will *t_i*.
b. Peter [*sells expensive cars*] and Bill does *so* too.
- (16) a. *Peter says he will sell expensive cars and *he sell* will expensive cars
b. **Peter sells* expensive cars and does *so* cheap trucks too

As illustrated by (15a,b), the verb and its direct object can be moved as a syntactic unit (i.e. VP-fronting) and be replaced by a pro-form *so*. As shown in (16a,b), these operations cannot be successfully applied to the verb and its subject-DP. This brief discussion of conflicting diagnosis clearly shows that one should not rely too much on a single diagnostic test. A rich diagnostic procedure consisting of several tests contributes greatly to one's understanding of the 'hidden structure' of a syntactic phenomenon or construction.

1.5 Contributions to syntactic diagnosis

As pointed out at the end of Section 1.2, the contributions on syntactic diagnosis in this volume are organized around five core domains of natural language syntax, namely: phrasal movement, head movement, agreement, anaphora, and ellipsis. Each of these syntax topics will be studied from four (diagnostic) perspectives, which correspond to four different chapters on that specific topic: syntax, syntax at the interface, psycho-/neuro-linguistics, and language diversity. A fifth chapter in which a broader perspective is taken on syntactic diagnosis in core domain X closes off each part/domain. In what follows, we give a brief summary of each of the contributions.

1.5.1 *Diagnosing head movement*

In 'Head movement as a phonological operation' (Chapter 2), Platzack establishes a diagnostics of Head Movement (HM) that helps to distinguish it from instances of phrasal movement. He then presents an analysis of head movement which treats it as a purely phonological process rather than a syntactic process. Building on some ideas from Distributed Morphology, he develops an analysis according to which head movement is the result of two Spell out principles that determine how the links in the extended projection chain of a verbal root is spelled out. He points out that his phonological account is compatible with the head movement diagnostics and does not have the drawbacks of syntactic head movement.

In 'Getting morphemes in order: Merger, affixation and head movement' (Chapter 3), Harley addresses the question of what the diagnostics for syntactic head movement are, and whether these individual diagnostics are in fact reliable cues. She shows that a tempting pair of diagnostics, viz., affixation and morpheme order, must be applied with extreme caution. As she points out, it is clear that just as

there can be head movement without affixation, there can be affixation without head movement. She also shows that the issue of which morpheme orders can properly be taken as ‘mirroring’ the syntax is somewhat more complex in implementation than is commonly assumed.

Friedmann’s contribution, ‘Verb movement to C: From agrammatic aphasia to syntactic analysis’ (Chapter 4), uses neurolinguistic data to diagnose head movement, and specifically, to diagnose the landing site of the verb movement that creates verb-second sentences in Hebrew. Using results from agrammatic aphasia, she shows that movement of the verb to ‘second position’ in Hebrew should be analysed as movement of the verb to C, rather than T.

Zeller’s contribution, ‘In defence of head movement: Evidence from Bantu’ (Chapter 5), discusses data from the Bantu languages, Shona and Kinyarwanda, which provide evidence that head movement operations form part of narrow syntax. He first of all shows that a head movement analysis is better suited to explain the verb-initial word order of Shona object relative clauses than an alternative based on remnant phrasal movement. He subsequently argues that an analysis of object markers in Kinyarwanda locative constructions as syntactically incorporated pronominal heads is superior to theories that treat object markers as non-incorporated phrases or agreement morphemes. On the basis of the arguments presented in his analysis of Shona and Kinyarwanda, he identifies various general head movement ‘diagnostics’, that is, criteria that can help determine whether or not a given word or affix order is derived by head movement.

1.5.2 Diagnosing phrasal movement

In his contribution ‘Phrasal movement and its discontents: Diseases and diagnoses’ (Chapter 7), Pesetsky starts from the observation that a diagnostic is simply an argument in which one has particular confidence, put to practical use. He sketches the logical space of possible arguments for phrasal movement and gives various illustrations of such arguments. Hartman’s (2012) discussion of intervention effects is cited as an instance in which an established property of movement, viz., intervention effects in A-movement constructions, diagnosed the distribution of movement in a more poorly understood construction, namely English *tough* movement. Pesetsky’s article ends with the question of whether phrasal movement exists in the first place and places it in the context of the history of its discovery and current syntactic approaches that dispense with it.

In his contribution ‘Diagnosing covert movement: The Duke of York and reconstruction’ (Chapter 8), Lechner aims (i) to give a survey of diagnostic strategies which have been employed in testing symptoms of covert displacement, and (ii) to develop two new diagnostic tests which aid in identifying covert movement

operations. These new tests are related in that both involve covert movement of categories that lack descriptive content.

In 'Arguments for Long Distance movement in Long Distance questions in child language' (Chapter 9), Demirdache examines the arguments put forth in the literature on child language for long movement in Long Distance questions. She reviews three diagnostics for Long Distance movement provided by comprehension studies and three diagnostics for Long Distance movement provided by elicited production studies. She concludes among other things that there is converging experimental evidence for successive-cyclic movement in child language, but also for partial movement, which is a confounding factor.

In 'Diagnosing covert A-movement' (Chapter 10), Polinsky and Potsdam discuss diagnostics for covert A-movement and ways in which it can be distinguished from non-movement. They propose that covert A-movement is found in subject-to-subject raising in the Northwest Caucasian language Adyghe and compare the Adyghe construction with unaccusatives in Russian, which they show do not involve covert A-movement. They demonstrate that a range of mostly theory-independent phenomena can be used to determine whether covert A-movement occurs in a given construction, using Adyghe and Russian as contrasting test cases.

1.5.3 *Diagnosing agreement*

Chung's contribution, 'The syntactic relations behind agreement' (Chapter 12), addresses the issue of whether morphological agreement is a diagnostic sign of the syntactic relation Agree (Chomsky 2000b, 2001b). As she points out, the strongest position one could take on this issue is that morphological agreement *coincides* with Agree: it is found when, and only when, Agree operates. Chung argues that the strongest position is untenable. Morphological agreement and Agree do not coincide; there are languages with no morphological agreement in which Agree operates 'as usual'. From this she concludes that morphological agreement is an imperfect sign of Agree.

Matushansky's contribution, 'Gender confusion' (Chapter 13), also addresses the question of whether the phenomenon of morphological agreement is diagnostic for the syntactic operation Agree. Like Chung, she answers this question negatively and shows that the various formulations of Agree do not account for all the configurations in which morphological agreement takes place and that Agree is incapable of dealing with the phenomenon of mixed agreement, where one and the same NP may trigger morphological agreement for different feature values.

In the chapter 'Agreement in the production of subject and object *wh*-questions' (Chapter 14), Guasti starts from the observation that object extraction (in questions) is more difficult than subject extraction in acquisition and processing across languages. She hypothesizes that this difficulty stems from an interference in

the AGREE relation between AgrS and the thematic subject. In her study she argues that in several cases, morphological agreement can be a diagnostic of (i) syntactic agreement; (ii) AGREE; and (iii) spec-head agreement depending on whether it is full morphological agreement or impoverished morphological agreement; and (iv) the relative position of subjects and verbs.

Ouhalla's contribution, 'Agreement unified: Arabic' (Chapter 15), examines various instances of morphological agreement in Arabic—more specifically, verb-agreement with the subject in clauses, participle/adjective-agreement with the subject in clauses and with the modified noun in noun phrases, relative Comp-agreement with the relativized noun in noun phrases, and demonstrative agreement with the noun in noun phrases—and reaches the conclusion that they are all diagnostic of a single syntactic derivational process, namely Agree subject to closest c-command, incorporating the mechanism of Feature Valuation (Chomsky 2000b, 2001b).

1.5.4 *Diagnosing anaphora*

Anagnostopoulou and Everaert's contribution, 'Identifying anaphoric dependencies' (Chapter 17), examines the validity of several standard anaphora diagnostics, such as strict/sloppy identity, split antecedents, deictic reference, a command restriction on the anaphoric dependency, and domain-sensitivity. They conclude that although there are reliable ways of diagnosing 'anaphors' (although perhaps not as many as one might have thought), how to diagnose a 'pronominal' is less clear.

In his contribution '~~Condition B~~' (Chapter 18), Tancredi focuses on Condition B and argues that this condition as an independent stipulation can be dispensed with from the syntax by slightly changing our understanding of how pronominal interpretation obtains. Arguing that Condition B effects should be accounted for outside of syntax obviously implies that this condition is no longer available for *syntactic* diagnosis.

In their contribution, 'A processing view on agrammatism' (Chapter 19), Avrutin and Baauw argue that an approach that accounts for impaired syntactic performance in terms of a processing deficit which complicates the use of syntactic operations is more fruitful than an approach which tries to account for impaired syntactic performance in purely syntactic terms. On the basis of condition A and B effects in agrammatic speakers, they show that an approach based on processing deficiency can adequately explain a wide range of superficially unrelated errors and help us identify, that is, 'diagnose', the exact nature of specific syntactic constructions.

In 'Tagalog anaphora' (Chapter 20), Richards shows how binding diagnostics are particularly useful for our understanding of a language like Tagalog, which is characterized by comparatively free word order in the postverbal domain. Taking binding theory as his 'diagnostic toolkit', Richards tries to show what the proper syntactic analysis of topicalization is making use of the difference between Tagalog scrambling and topicalization.

1.5.5 Diagnosing ellipsis

In ‘Polarity items under ellipsis’ (Chapter 22), Merchant attempts to bring some conclusions from recent work on ellipsis to bear on the question of the nature of polarity items. He shows that polarity items license the ellipsis of their non-polarity counterparts and vice versa. If ellipsis is subject to a syntactic identity constraint on non-inflectional features, then the conclusion must be drawn that pairs like *some* ~ *any* are syntactically identical in the relevant sense. This identity is implemented as lexical identity, where the morphological differences among such items is analysed as the result of agreement operations valuing inflectional features differently.

In ‘Syntactic diagnostics for extraction of focus from ellipsis site’ (Chapter 23), Winkler argues that extraction from ellipsis site is restricted to Contrastive Focus. She presents an account of such extraction based on Krifka’s (2006) Association with Focus Phrase theory, and develops a set of diagnostics for extraction from ellipsis site which distinguishes contrastive ellipses from phrasal ellipses.

In ‘A Recycling approach to processing ellipsis’ (Chapter 24), Frazier points out that, in order to be able to identify reliable grammatical properties of ellipsis, it is necessary to distinguish grammatical effects from processing effects. She develops a theory of ellipsis processing in which specific types of ‘mismatch ellipsis’—that is, ellipsis where the antecedent does not syntactically match the elided constituent—may be repaired using the revision procedures independently needed in an account of garden-path processing. The Recycling account predicts the circumstances under which mismatch ellipsis is relatively good. This grammar plus processor view of ellipsis relieves the grammar of the burden of explaining the acceptability of mismatch examples and, it is argued, the burden of explaining the acceptability of island violations in certain ellipsis sentences.

In ‘What sluicing can do, what it can’t, and in which language: On the cross-linguistic syntax of ellipsis’ (Chapter 25), van Craenenbroeck and Lipták present a new, cross-linguistically refined theory of sluicing and examine the predictive force of this new theory in various domains of *wh*- and focus syntax. They start out by showing that the restriction of sluicing to *wh*-questions is not a reliable test for diagnosing this construction cross-linguistically. They put a new generalization in place, which informally states that the types of sluicing in any given language track the overt syntax of *wh*-movement in that language. This new generalization is put to work in the second part of the chapter, where they show, on the basis of Italian, Venetian, and Bulgarian, that the availability of non-*wh*-sluicing can provide new evidence for or against syntactic accounts positing that *wh*-movement and focusing target the same left peripheral position.

1.6 Conclusion

‘Diagnosis is not the end, but the beginning of practice.’ According to these words of the physiologist Martin H. Fischer (see Fabing and Marr (1937)), the identification of a disease or disorder by means of a diagnostic procedure precedes the most relevant part of medical practice, namely prescription of a treatment that is intended to stop a medical condition from progressing any further. A physician implements a health-care programme in the form of instructions that govern the plan of care for an individual patient. This programme may include instructions to the patient (Take medication X! Do the following physical exercises for ten minutes during the next three weeks!), instructions to a nurse (Collect blood from patient X for blood testing!), instructions to a pharmacist (Prepare a compounded drug! Determine the right dosage of this drug for patient X!), and instructions to the clinical laboratory worker (Run an extra urinalysis! Run a serum concentration test!).

For linguistics also, it holds that diagnosis is not the end, but the beginning of practice. It is the beginning of making important descriptions, generalizations, and theorizing. After having diagnosed the presence of a linguistic property P in language L, there are many different ways to go forward. One might look for other properties in the language that can be correlated with P, or we can investigate whether P is also attested in other languages (of the same family or not), and whether P helps in shedding new light on a particular phenomenon. Ultimately, it leads to the question of *why*? Why does P exist? And pursuing the answer of this question can give us better insight into the nature of language and of the human mind.

‘Practice’ is the act of engaging in an activity again and again, for the purpose of improving or mastering it. In order to be able to understand the nature (‘the hidden structure’) of syntax (i.e. syntactic knowledge), the syntactician must keep practising, developing, and refining his diagnostic abilities, and syntactic diagnostics in general. Being a good diagnostician is a prerequisite for being a good syntactician. We hope that the chapters in this book will help the current and future practitioners.

Acknowledgement

The chapters of this volume are updated versions of talks that were presented at the workshop *Diagnosing Syntax* that we organized at Leiden University and Utrecht University, the Netherlands, in January 2009. We gratefully acknowledge the financial support of the Royal Netherlands Academy of Arts and Sciences (KNAW), Leiden University Centre for Linguistics (LUCL), and Utrecht Institute of Linguistics (UiL-OTS) for financial and structural support. Anonymous thank-you’s are due to our anonymous referees. Finally, our thanks go to Robin Langendijk for her assistance in preparing the manuscript.