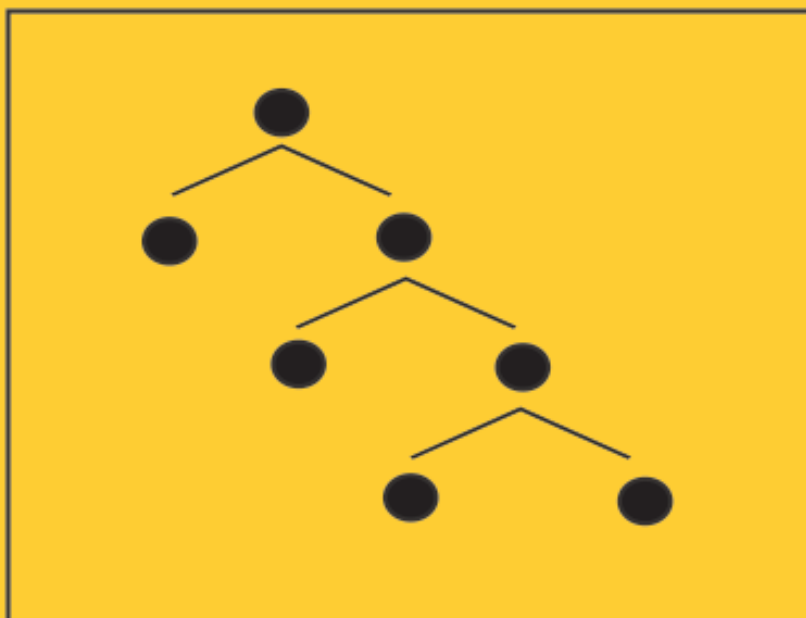




MATERIALS
250

SUSAGNA TUBAU

Syntactic Analysis of English



SUSAGNA TUBAU

Syntactic Analysis of English

Universitat Autònoma de Barcelona
Servei de Publicacions
Bellaterra, 2026

Primera edició: juliol de 2026

© del text:

Susagna Tubau

© d'aquesta edició:

Servei de Publicacions de la UAB

© imatge de la coberta:

Susagna Tubau, *Syntactic tree*

Edició:

Servei de Publicacions

Universitat Autònoma de Barcelona

Edifici A. 08193 Bellaterra (Cerdanyola del Vallès). Spain

Tel. 93 581 10 22

sp@uab.cat

<https://publicacions.uab.cat>

ISBN (paper): 979-13-88031-37-3

ISBN (digital): 979-13-88031-38-0

Dipòsit legal: B-14312-2026

Imprès a Espanya. Printed in Spain



Aquest llibre està publicat amb una llicència Creative Commons CC-BY-NC-ND.
El titular de l'obra autoritza a utilitzar els continguts sempre que es reconegui l'autoria.
No es permet fer un ús comercial, ni la generació d'obres derivades.

Table of contents

UNIT 1. INTRODUCTION: CORE CONCEPTS AND THEORETICAL ASSUMPTIONS.....	7
1.1. Objectives of the unit.....	7
1.2. Linguistic competence and grammar.....	7
1.3. (Un)grammaticality vs. (un)acceptability.....	8
1.4. Syntax as science and levels of adequacy.....	8
1.5. Suggested exercises and key.....	10
1.5.1. Exercises.....	10
Exercise 1.....	10
Exercise 2.....	10
Exercise 3.....	11
Exercise 4.....	11
1.5.2. Key to exercises.....	12
Exercise 1.....	12
Exercise 2.....	12
Exercise 3.....	12
Exercise 4.....	13
1.6. Suggested reading.....	13
UNIT 2. MORPHOSYNTACTIC FEATURES AND MERGE.....	15
2.1. Objectives of the unit.....	15
2.2. What are features?.....	15
2.3. Ways to think of a feature: privative vs. binary-valued.....	16
2.4. Category features, phi-features, and verbal features.....	17
2.5. Interpretable vs. uninterpretable features and the interfaces.....	21
2.6. Suggested exercises and key.....	23
2.6.1. Exercises.....	23
Exercise 1.....	23
Exercise 2.....	24
Exercise 3.....	25
2.6.2. Key to exercises.....	25
Exercise 1.....	25
Exercise 2.....	26
Exercise 3.....	26
2.7. Suggested reading.....	26
UNIT 3. PHRASE STRUCTURE AND DEPENDENCIES.....	27
3.1. Objectives of the unit.....	27
3.2. Constituency and phrase structure.....	27
3.3. Theta-roles and phrase structure.....	29
3.4. More on Merge and feature checking.....	29
3.5. The structure of phrases: Merge and adjunction.....	30
3.6. C-command.....	32
3.7. Ditransitives and little <i>v</i>	34

3.8. Suggested exercises and key	36
3.8.1. Exercises	36
Exercise 1.....	36
Exercise 2.....	36
Exercise 3.....	37
Exercise 4.....	37
Exercise 5.....	37
3.8.2. Key to exercises	37
Exercise 1.....	37
Exercise 2.....	38
Exercise 3.....	38
Exercise 4.....	38
Exercise 5.....	39
3.9. Suggested reading	39
UNIT 4. FUNCTIONAL CATEGORIES AND MOVEMENT I (TP)	41
4.1. Objectives of the unit.....	41
4.2. Sentences as Tense Phrases.....	41
4.3. Checking via Agree.....	44
4.4. Movement: what drives it?.....	45
4.5. The Pronouncing Tense Rule, Neg(ation) and <i>do</i> -support.....	48
4.6. The syntax of intransitive verbs	50
4.7. Suggested exercises and key	52
4.7.1. Exercises	52
Exercise 1.....	52
Exercise 2.....	52
Exercise 3.....	53
Exercise 4.....	53
Exercise 5.....	53
4.7.2. Key to exercises	54
Exercise 1.....	54
Exercise 2.....	54
Exercise 3.....	55
Exercise 4.....	55
Exercise 5.....	55
4.8. Suggested reading	56
UNIT 5. FUNCTIONAL CATEGORIES AND MOVEMENT II (CP)	57
5.1. Objectives of the unit.....	57
5.2. Complementisers (<i>that, whether, if</i>).....	57
5.3. Interrogatives: <i>yes/no</i> -questions.....	60
5.4. Suggested exercises and key	61
5.4.1. Exercises	61
Exercise 1.....	61
Exercise 2.....	62
Exercise 3.....	62
5.4.2. Key to exercises	62
Exercise 1.....	62
Exercise 2.....	62
Exercise 3.....	63
5.5. Suggested reading	63
6. BIBLIOGRAPHY.....	65

UNIT 1. Introduction: core concepts and theoretical assumptions

Unit 1 aims to introduce you to the concepts of *linguistic competence* and *grammar*, as well as to the notions of *grammaticality* and *ungrammaticality*. You will be taught that *(un)grammaticality* is different from what we understand as *(un)acceptability*. In this Unit we will also encourage the need to address syntax as a scientific undertaking, guided by the scientific method. You will also be introduced to different *levels of adequacy* and their significance for formalising a theory of language.

1.1. Objectives of the unit

- Define *linguistic competence* and its relation to *grammar*.
- Distinguish *(un)acceptability* from *(un)grammaticality*.
- Apply the scientific method to the analysis of linguistic data.
- Define different *levels of adequacy*.

1.2. Linguistic competence and grammar

Speakers can decide whether a given linguistic form is well-formed or not, in the languages they speak. Take, for instance, the sentence in (1):

- (1) Today started the class rather late the teacher.

What is the problem with this sentence?

It does not seem to be a well-formed sentence in English. If asked to correct it, most speakers (even those without explicit linguistic training) would probably coincide on (2) being an appropriate counterpart of (1).

- (2) Today the teacher started the class rather late.

In addition, if linguistically trained, speakers might say that the reason why (1) is not well-formed in English is the position of the subject (*the teacher*), which occurs after the verb in (1) but before it in (2).

The crucial question here is why all speakers of English would agree on (1) being incorrect but (2) being correct. The answer to this question is that they do so because they all possess a certain kind of linguistic knowledge, whether they are aware of it or not. Such linguistic knowledge is called *grammar*, and it is what allows speakers of a language to decide whether a given linguistic form such as a phrase or a sentence is well-formed or not in a given language. It is the mission of syntacticians to uncover the grammar of particular languages and to explain how speakers come to have such knowledge.

1.3. (Un)grammaticality vs. (un)acceptability

Because of our knowledge of English grammar, we can not only rule out (1) as ill formed but also have judgements on much more subtle data. Let us see this with an example. Three of the sentences in (3)-(6) are grammatical, but one is not. Can you identify which is the ungrammatical one?

- (3) I filled the bottle up.
- (4) I filled up the bottle.
- (5) I filled it up.
- (6) I filled up it.

And could you try to explain, in your own words, what it is that you know about English grammar that has enabled you to detect the odd one out?

Your first answer must have been that the ungrammatical sentence is (6), while (3), (4) and (5) are fine. Somehow, (6) breaks an English grammar rule, and you seem to be aware of it (even if you may not know how to formalise the actual rule that is being broken). You might have spotted that *fill up* is a phrasal verb, and that both *bottle* and *it* are complements of this verb. They are, however, slightly different complements: while *bottle* is a noun, *it* is a pronoun. So as both (3) and (4) are fine, you can conclude that noun complements can either follow or precede the particle of a phrasal verb, while pronominal complements can only precede the particle. This is a grammar rule, and you can further test that it applies to other instances of noun and pronoun complements by experimenting with other phrasal verbs. It is precisely this knowledge that allows you to have the correct *grammatical judgements* for (3), (4), (5) and (6).

To indicate the ungrammaticality of (6) syntacticians use an asterisk in front of the sentence. From now on, then, you will take starred sentences to have been judged as ungrammatical by native speakers of English.

- (7) *I filled up it.

The concept of *(un)grammaticality* must be distinguished from another concept that is also often used in syntax, namely *(un)acceptability*. While these two notions sometimes overlap, this might not always be the case. Take, for example, the sentence in (1), repeated here as (8). This is both ungrammatical (as it does not enforce the grammar rules for English word order) and unacceptable.

- (8) *Today started the class rather late the teacher.

But consider a sentence such as (9).

- (9) #The dog spoke wise words to me.

Is it grammatical?

It is, as it does not break any English grammar rules, but the literal meaning is odd. Speakers would find (9) grammatical but unacceptable unless it is used in the context of an imaginary world where dogs can speak wise words to humans. Unacceptability is indicated with the symbol '#'.

1.4. Syntax as science and levels of adequacy

Syntax is a scientific undertaking in the sense that by investigating the grammar of a given language, we can follow the scientific method in four steps. First, we need linguistic data. That is, we need to gather observations, which, in our case, will be sentences that are considered grammatical by native speakers of the language and others that are not. This is what we did with the examples in (3), (4), (5) and (6). We then tried to make generalisations. In (3), (4), (5) and (6), for example, we used the terminology *noun*, *pronoun*, *complement*,

phrasal verb and *particle* and came up with the generalisation that nominal complements can either precede or follow the particle of a phrasal verb, but that pronominal complements can only precede it. In step 3, we will hypothesise that this is the case for every phrasal verb, and in step 4 we will test this hypothesis with more data (which takes us back to step 1).

If you do that, you will soon come up with plenty of counterexamples. *Look after* is just one of them, but there are many others such as *look into*, *come across*, *go over*, etc. As you can see in (11), *look after* is a phrasal verb that does not allow a nominal complement to precede the particle, which is unexpected according to how we have so far characterised the syntax of phrasal verbs. By contrast, the ungrammaticality of (13) falls into place within our hypothesis.

- (10) I'm looking after my sister tonight.
- (11) *I'm looking my sister after tonight.
- (12) I'm looking after her tonight.
- (13) *I'm looking her after tonight.

To make sense of the data, it is time to refine our analysis and consider whether there may be differences between phrasal verbs. It seems that some phrasal verbs allow the verb to be separated from its particle (separable phrasal verbs), while others do not (non-separable phrasal verbs). Taking this new development into account, can you reformulate your earlier generalisation to capture all the facts about the grammar of phrasal verbs that we have seen thus far?

The new generalisation could be as follows:

(14) If a phrasal verb is separable, a nominal complement can either precede or follow the particle, while a pronominal complement can only precede the particle (e.g., *fill the bottle up/fill up the bottle* vs. *fill it up/*fill up it*). If a phrasal verb is non-separable, both nominal and pronominal complements must follow the particle (e.g., *look after my sister/*look my sister after* and *look after her/*look her after*).

In the approach to syntax that we are adopting, we aim to achieve what is known as *explanatory adequacy*, which involves not only seeking to observe the native speakers' grammatical judgements (that is, which of the sentences in the data sets should be starred and which ones should not), but also to explain why these judgements are as they are. There are other levels of adequacy that our theory might have that are not as powerful as explanatory adequacy. If our model has *observational adequacy*, it can list the grammatical sentences (but provides no information about the ungrammatical ones); if it has *descriptive adequacy*, our model formalises the rules that produce grammatical sentences but does not explain why they are grammatical. It seems, therefore, that only a model with *explanatory adequacy* has predictive power and provides information about the underlying structure of a language (that is, the grammar).

Grammar, or speakers' tacit knowledge of their language, is learnt in a very different way than other kinds of knowledge: normal-developing children around age three show command of the basic structures of their language without having been explicitly taught. In addition, they seem to acquire their native language going through similar steps (that is, children's deviations from adult grammar are not aberrant). These observations (namely that our first language is learnt without being taught, and that there seems to be a common acquisition path) have been taken to support the idea that humans are biologically endowed with a language faculty, and that this constrains the kinds of languages that a child can learn. Such language faculty has also been called *Universal Grammar* (Chomsky 1957, 1965, 1981, 1986 and 1995 among others) and is understood as a collection of principles that are common to human language. Yet, it is obvious that languages differ. For example, all languages might form sentences by using the same ingredients (a subject, a verb and, optionally, an object), but these occur in different orders. English and Catalan are S(ubject)-V(erb)-O(bject) languages, as shown by (15) and (16), while Basque, (17), is SOV.

- (15) [The boy]_S [reads]_V [a book]_O
- (16) [El nen]_S [llegeix]_V [un llibre]_O

- (17) [Haurrak]_S [liburu bat]_O [irakurtzen du]_V
 boy.the book one read auxiliary
 ‘The boy reads a book.’

Children are not explicitly taught which word order their language has; rather, they infer it from the linguistic input. In the process, they might try different word orders, even some which they have not really found in their adult input, as in (18). Eventually, though, they will set the right one in their grammar.

- (18) [Mama]_S [pitet]_O [posa]_V
 mummy bib puts on
 ‘Mummy is putting the bib on.’
 (adapted from Llinàs-Grau and Coll-Alfonso 2001)

Note that for our syntactic theory to be explanatorily adequate, we should be able to observe that languages differ with respect to word order, but also to explain why they do so, and how speakers come to acquire the right property for their native language.

1.5. Suggested exercises and key

1.5.1. Exercises

Exercise 1.

While it is easy to think of examples of grammatical sentences that are acceptable and of ungrammatical sentences that are unacceptable, examples also exist of

- (i) grammatical sentences that are unacceptable, and
 - (ii) ungrammatical sentences that are acceptable.
- (1) is an example of (i) and (2) is an example of (ii).

- (1) This book is mine while that that that professor took is yours.
- (2) But if this ever-changing world in which we live in makes you give in and cry, say live and let die.
(Live and Let Die, Paul McCartney)

Try to explain why (1) is unacceptable (though grammatical) and (2) is ungrammatical but acceptable.

Exercise 2.

In the lecture it was shown that a sentence such as (1) was ungrammatical in English, and that its grammatical counterpart was (2). By using the notion of *subject*, we claimed that the subject could not occur after the verb.

- (1) *Today started the class rather late the teacher.
- (2) Today the teacher started the class rather late.

Consider the sentences in (3) and (4). Are they grammatical? Ask a native English speaker if you are unsure.

- (3) Down the street marched the soldiers.
- (4) Marched the soldiers down the street.

Once you have the judgements, compare them to (5) and (6), which are grammatical in English. Can you figure out what allows a subject to be postverbal in English?

- (5) Here comes the sun!
- (6) I opened the door and there stood a mysterious woman.

Exercise 3.

Practise the scientific method by doing exactly what each Task asks you to do:

Task 1 (gathering data). Consider the following sentences and indicate which are grammatical and which are not.

- (1) I love myself.
- (2) I love yourself.
- (3) They love themselves.
- (4) They love ourselves.

Task 2 (making generalisations). Use the terminology ‘subject’, ‘object’, ‘singular’ and ‘plural’, ‘first, second and third person’ to write a generalisation that captures what you observed in Task 1.

Task 3 (hypothesising). By referring to ‘I’ and ‘they’ as *the antecedent*, and to ‘myself/yourself/themselves/ourselves’ as *the anaphor*, formulate a hypothesis that explains why some sentences in Task 1 are correct and some are incorrect. You can also use the terms *agree*, *person* (first, second, third) and *number*.

Task 4 (testing the predictions made by your hypothesis). Consider this new set of data. Does your hypothesis predict the facts? If not, revise your hypothesis.

- (1) *The woman talked to itself.
- (2) The woman talked to herself.

Task 5. If you had to revise your hypothesis, test it again with this new set of data. Does your revised hypothesis explain all the data? If not, revise it again.

- (1) *We talked to myself.
- (2) *I talked to ourselves.
- (3) We talked to ourselves.
- (4) I talked to myself.
- (5) *She talked to myself.
- (6) She talked to herself.
- (7) *You talked to myself.
- (8) You talked to yourself.

Exercise 4.

[Attempt this exercise after doing the suggested reading.]

Explain how the *Principle of Structure Dependency* can account for the ungrammaticality of (2) (and the grammaticality of (3)).

- (1) The woman’s cat escaped.
- (2) *The old’s woman cat escaped.
- (3) The old woman’s cat escaped.

Expand the noun phrase in (3) to further illustrate how the *Principle of Structure Dependency* functions.

1.5.2. Key to exercises

Exercise 1.

- (1) This book is mine while that that that professor took is yours.

The sentence has three consecutive *that*, which may result in speakers finding it unacceptable. Yet, no rules of English are broken here, as the sentence is analysed/interpreted as in (1’):

- (1’) This book is mine while [_{pronoun} that (book)] [_{subordinator} that [_{demonstrative} that professor took is yours]]
- (2) But if this ever-changing world in which we live in makes you give in and cry, say live and let die.
(*Live and Let Die*, Paul McCartney)

The sentence is acceptable, as most speakers would not reject it as a non-English sentence, but it is ungrammatical, as the preposition *in* has been used twice. In fact, there is a dispute concerning whether McCartney sings “But if this everchanging world *in* which we live *in*...” or “But if this ever-changing world *in* which we’re living...”. The cover that is sung by Guns N’ Roses undoubtedly has “*in* which we live *in*”.

Exercise 2.

- (3) Down the street marched the soldiers. [GRAMMATICAL]
(4) *Marched the soldiers down the street. [UNGRAMMATICAL]
(5) *Here comes* the sun!
(6) I opened the door and *there stood* a mysterious woman.

Factors that allow subjects to be postverbal in English: (i) the presence of a sentence initial adverb (e.g. *here, there*); (ii) certain types of verbs such as *march, come, stand*.

Exercise 3.

Task 1.

- (1) I love myself.
(2) *I love yourself.
(3) They love themselves.
(4) *They love ourselves.

Task 2. In the grammatical sentences a singular subject in the first person matches with a singular object in the first person, and a plural subject in the third person matches with a plural object in the third person. The ungrammatical sentences show mismatches of singular/plural (number) or person (e.g., a singular subject in the first person is put in the sentence with a singular object in the second person).

Task 3. The anaphor and its antecedent must agree in person (first, second, third) and number (singular, plural).

Task 4.

- (1) *The woman talked to itself. [third person, singular – third person, singular]. Our hypothesis does not predict the ungrammaticality.
(2) The woman talked to herself. [third person, singular – third person, singular].

Grammaticality is predicted by our hypothesis, but as it stands, we cannot distinguish (1) from (2) and yet (1) is ungrammatical and (2) is grammatical.

Gender seems to be relevant for anaphors. ‘The woman’ is feminine and ‘itself’ is neutral, while ‘herself’ is feminine.

Reformulation of the hypothesis: The anaphor and its antecedent must agree in person (first, second, third), number (singular, plural) and gender (masculine, feminine and neuter).

Task 5. Our hypothesis now covers all the cases in Task 4.

Exercise 4.

In the Suggested reading we learn that the *Principle of Structure Dependency* says that syntactic operations apply to hierarchical rather than linear structures. This means that, as shown in (2), it would be wrong to assume that the genitive is attached generally to the second word in the sentence, as could be wrongly concluded from (1). As shown in (3), the genitive is to be attached to the noun phrase, regardless of how many words this contains. Examples (4) and (5) further illustrate this point.

(4) The nice old lady's cat escaped.

(5) Mary's cat escaped.

(4) shows that if we make the noun phrase longer, the genitive still comes at the very end, while (5) shows that the genitive can be attached to noun phrases shorter than two words.

1.6. Suggested reading

Edelstein (2020): Chapter 1. *Got grammar?*

As a complement to the theoretical concepts seen in the unit, this reading adds the notions of *competence*, *performance*, *I-language*, *E-language*, *Universal Grammar*, and *economy*, and it gives a definition of *syntax* and *Minimalism*.

UNIT 2. Morphosyntactic features and Merge

Unit 2 aims to introduce the notion of morphosyntactic features as the building blocks for syntax. In this unit we will explore which morphosyntactic features are relevant for English syntax, how they relate to each other and how they interact with the articulatory-perceptual and conceptual-intentional system (the so-called interfaces).

2.1. Objectives of the unit

- Define *morphosyntactic features* and argue for their existence.
- Introduce and discuss the concept of *privative* and *binary-valued* features.
- Introduce *category features*, *phi-features*, and *verbal features*.
- Establish a basic inventory of morphosyntactic features relevant for English.
- Introduce the distinction between *interpretable* and *uninterpretable* features.
- Introduce the concept of *interface*, the *Principle of Full Interpretation* and *checking*.
- Introduce the syntactic operation *Merge*.

2.2. What are features?

Very much in the same way atoms (from Greek *atomos* ‘uncuttable’) were discovered to be further broken down into subatomic particles (electrons, protons, and neutrons), it soon becomes obvious that words cannot be the ultimate syntactic formatives. Consider, for example, the sentences in (1), (2), (3) and (4).

- (1) The students like syntax.
- (2) The student likes syntax.
- (3) *The students likes syntax.
- (4) *The student like syntax.

Example (1) will continue to be grammatical if the noun phrase ‘the students’ is replaced by any other (e.g., ‘the boys’, ‘John and Mary’, etc.) provided it is *plural* rather than *singular* (see (4)).

In a similar vein, (2) will also continue to be grammatical if ‘the student’ is replaced by any other singular noun phrase rather than a plural one (see (3)). If instead of the noun phrases we examine the verb, we will conclude that we also have a choice (e.g., *hate*, *love*, *study*, *take*, etc.) provided the verb we choose agrees with the subject. For English present tense, this means that no inflection appears with the verb when the subject is plural, (1). And what about the singular? Does it require the {-s} morpheme?

The answer is yes, and no. ‘Yes’ in light of (2) (and the ungrammaticality of (4)), but ‘no’ in light of the data in (5) and (6).

- (5) I like syntax.
- (6) *I likes syntax.

So, when is the morpheme {-s} required?

If you add more data to the equation, you will conclude that the morpheme {-s} is connected to present tense (see (9)), singular, and third person, (8). These properties that different types of words have are what we call *morphosyntactic features*.

- (7) You like syntax.
- (8) *You likes syntax.
- (9) *She likeds syntax.

Morphosyntactic features are relevant for syntax (hence the ‘syntactic’ part of the term), but they may also determine the shape of a word (hence the ‘morpho’ part of the term). Concerning this latter statement, namely that morphosyntactic features may be responsible for a word taking a given shape, let us consider the plural.

If you think about the plural forms of the nouns in (10), you can make sense of the data in (11). The morphology of the three words in (10) is affected in different ways by the feature *plural*.

- (10) car, man, child
- (11) cars, *caren / *mans, men / *childs, children

Morphosyntactic features, hence, can be defined as properties of words that may affect their morphology and that matter for syntax. They also affect the semantics (that is, the meaning) of words. The feature plural, for example, changes the meaning of the words in (10): if there is more than one *car/man/child*, we must use the plural forms in (11), rather than the singular ones in (10).

When a feature has an effect on semantic interpretation, it is called *interpretable*. We will come back to this term (and its counterpart *uninterpretable*) later in the unit. For the moment, our aim will be to establish the minimal set of features that suffices to explain English grammar. In doing so, though, we will have to answer a key question: namely, how do we think of a feature?

2.3. Ways to think of a feature: privative vs. binary-valued

Given that our purpose is to uncover the minimal set of morphosyntactic features that are relevant for English syntax, we need to start looking at data to establish what properties of words can be considered morphosyntactic features. Let us start by examining the case of nouns. We saw earlier that in English, things can be either singular or plural, and that this property has an impact both on form and on syntax. It makes sense to hypothesise, therefore, that [singular] and [plural] (hereinafter, we will use square brackets to indicate that a given property is a feature) are morphosyntactic features of English. According to this hypothesis, *car*, *man* and *child* would be specified as [singular], while *cars*, *men* and *children* would be specified as [plural].

Our hypothesis is problematic in several ways, though. First, if [singular] and [plural] are independent features, what prevents them from occurring simultaneously? To my knowledge, there are no words in English that are simultaneously singular and plural. Or are there?

At this point, you might have come up with words such as *fish* or *deer*. The plural form for these words is identical to their singular form. Note, however, that if we use them in the context of a sentence, they can no longer be assumed to be both singular and plural at the same time. In (12) *fish* is singular, and in (13) it is plural, and it cannot be the other way around because, otherwise, the sentences would be ungrammatical due to lack of verbal agreement.

- (12) The fish I bought keeps me company.
- (13) The fish I bought keep me company.

A second problem with our hypothesis is that it would also allow us to predict that there are nouns with neither of the two features. Again, no such nouns exist in English.

Thus, as nouns in English are either singular or plural, we will refine our hypothesis by trying to use just one feature instead of two. If we can still account for what we have observed for English nouns, then our new analysis will be superior to the former one; being simpler, yet explanatory.

Let us hypothesise that only [plural] is a morphosyntactic feature of English (but not [singular]). If present on a noun, then, we should obtain a morphological reflection of it (and/or a syntactic effect). That is, the nouns in (10) would be analysed as not carrying any feature related to number, whereas the grammatical words in (11) would be analysed as specified with the feature [plural]. The feature [plural], then, would either be absent (e.g., *car*, *man*, *child*) or present (*cars*, *men*, *children*) and would therefore be referred to as a *privative feature*.

Alternatively, we might consider that morphosyntactic features are *binary-valued* rather than privative. That is, a feature such as [plural] would have the [+] value for a noun like *cars*, but the [-] feature in *car*.

Finally, a last approach within the binary-valued view of features involves thinking of the relevant feature as [number] and taking their possible values to be ‘singular’ or ‘plural’, (14).

- (14) a. [number: singular]
b. [number: plural]

So how should features be treated? The choice is empirical (that is, determined by the data under study) and must be driven by the philosophical principle known as Occam’s Razor, according to which *non sunt multiplicanda entia sine necessitate* ‘Entities are not to be multiplied without necessity’ or, in other words, that the simplest theory is always the best. Of course, it might be the case that certain phenomena force us to increase the complexity of our analyses. The idea is, though, that we use the privative feature system whenever possible, and consider using some version of the binary-valued one only on empirically justified grounds.

2.4. Category features, phi-features, and verbal features

How was the feature [plural] motivated in the previous sections? Mostly, by paying attention to two crucial facts: (i) it resulted in distinguishable morphological forms, and (ii) it had a semantic effect. It is possible, however, to express features that have a semantic interpretation but do not result in a distinct morphology (e.g., some kinds of verbal feature) and also features that result in morphological variation but do not have a semantic interpretation (e.g., case features).

In this section we will see evidence for *category features*, *phi-features* (*person*, *number*, and *case* in English), and some *verbal features*. Let us start with category features.

Category features allow us to classify words into traditional major classes (noun, verb, adjective, preposition, etc.). They have a semantic basis (e.g., nouns denote objects in the world, verbs denote events or actions, etc.) and give rise to different morphological forms. They also have different distributions. For example, only nouns can appear after possessive pronouns, (21), only verbs are compatible with tense, (22), and only adjectives can be inflected for degree (comparative and superlative), (23).

- (21) a. my book
b. *my under
c. *my sing
d. *my tall

- (22) a. I played.
b. *I undered
c. *I windowed.
d. *I tallied.

- (23) a. *This is booker than that.
b. *This is the underest in the world.
c. *This is singer than that.
d. This is taller than that.
e. This is the tallest in the world.

How can we use morphosyntactic features to classify words into major word classes then? Let us assume that nouns can have the category feature [N], whereas verbs can have the category feature [V]. As we must try to account for the facts with as few features as possible (remember that the simplest theory is the best), let us try to derive all word classes just by using these two features. With two independent privative features, four categories are predicted. The first two correspond to nouns and verbs, but what do (24c) and (24d) correspond to?

- (24) a. nouns [N]
b. verbs [V]
c. [N, V]
d. []

It has been observed that adjectives relate to nouns (e.g., they may be distributed in the same positions, as in *The rich should be more generous*) but also to verbs (e.g. participle verbs can act as adjectives, as in *Imported goods may be exotic*) so it would make sense to argue that they carry both [N] and [V] features. The specification in (24d) (that is, the lack of both [N] and [V] features) would then correspond to the fourth of the major syntactic classes, namely prepositions, which are not nouns, or verbs, or adjectives.

The fact that prepositions do not have any category features is somewhat problematic, though (how can they be a word class without any features at all?), and triggers an alternative analysis of major word classes with noun and verb features as binary-valued. The characterisations in (24), therefore, would be as in (25), thus solving the theoretical problem posed by having a given category unspecified for category features.

- (25) a. nouns [+N, -V]
b. verbs [-N, +V]
c. adjectives [+N, +V]
d. prepositions [-N, -V]

Apart from category features, English also has *phi-features*. These include number (which we saw earlier), person and case. Let us consider what person features are relevant for English morphosyntax and how they combine with number features. To describe the pronominal paradigm in the chart in (26), we might want to use the following privative features (in combination with the number feature [plural]): first person, [1], second person, [2], and third person, [3].

(26)

I [1]	we [1], [plural]
you [2]	you [2], [plural]
he/she/it [3]	they [3], [plural]

Notice, though, that while this characterises the different pronominal forms in (26), it also makes the wrong predictions about possible feature combinations that are not attested in English. Which ones?

As shown in (27), no pronouns exist without any person specification at all in English, and neither do pronouns with combinations of person features such as (27c, d, e, f). This means that we need to rethink whether we need that many features or not.

- (27) a. []
b. [1, 2]
c. [2, 3]
d. [1, 3]
e. [2, 3]
f. [1, 2, 3]

In a second attempt to characterise the pronouns in (26) we may try to remove the feature [3] and see if we can still attribute each of the forms a unique feature characterisation. This results in (28).

(28)

I [1]	we [1], [plural]
you [2]	
he/she/it	they [plural]

Note that we need to distinguish pronouns with feature [1] based on number (*I* vs. *we*), and that if we only have one form, *you*, for [2], then this is because number is not specified when it comes to the second-person pronoun. The correct interpretation of *you* as referring to one person or more must be guessed by the context. Finally, concerning *he/she/it* and *they*, we can distinguish them by number (as in the case of [1] pronouns), but we do not need to define a person feature for them, as they are simply not [1] and not [2] (so they are third person by default).

The feature characterisation in (28) is superior to (26) according to Occam’s Razor, but it still makes a prediction that needs to be checked. If [1] and [2] are the only two person features, then, it is predicted that they may occur together in a pronoun. This is indeed the case in some languages of the world. For example, Dakota in North America, Marghi in Africa, and Ngandi in Australia (Adger 2003: 44) differentiate between *exclusive* plural pronouns (the addressee is left out of the *we* pronoun so the feature characterisation is [1], [plural]) and *inclusive* plural pronouns (the addressee is included in the *we* pronoun so the feature characterisation is [1, 2] [plural]) and select different pronouns for these two readings.

English does not have a lexical difference between these two kinds of *we* pronoun, but it is possible to create a context that confirms our hypothesis that features [1] and [2] can be combined to yield a different interpretation of a sentence with *we*. Consider the two examples in (29):

- (29) a. We, as in you and I, are going to the movies. Inclusive, [1] [2], [plural]
b. We are going to the movies (but you are not). Exclusive, [1], [plural]

In (29a), by adding the phrase *as in you and I*, we make it clear that the reading includes both the speaker, [1], and the addressee, [2]. By contrast, in (29b), by specifying *but you are not*, we make it clear that the addressee, [2], is excluded.

As part of phi-features, we also find *case features*. These are responsible for pronouns adopting different morphological forms depending on the syntactic function they perform. In (30), *I* is in the nominative case and carries the feature [nominative], *them* is in the accusative case and carries the feature [accusative], and *our* is in the genitive case and carries the feature [genitive].

- (30) I told them to fix our car.

Finally, is gender a morphosyntactic feature in English? We should test whether it affects the morphology of words, whether it has a semantic effect and whether it matters for syntax. As we see three different pronominal forms for third person singular pronouns, we conclude that gender affects the morphology of words. Gender also has a semantic effect. Yet, we do not see any syntactic differences triggered by the presence of the gender feature, (31). We conclude, therefore, that gender is not a morphosyntactic feature of English. Rather, it is a *semantic feature* in English (but it is a morphosyntactic feature in other languages such as Catalan or Spanish).

- (31) a. She is running/runs/ran away.
 b. He is running/runs/ran away.
 c. It is running/runs/ran away.

To end this section, we now turn to *verbal features*. As we can see in (32), verbs in English are inflected for past tense (but not for present). This triggers the consideration of [past] as the relevant feature. Past inflection can be by means of the suffix {-ed}, (32b), suppletion, (32d), or vowel change, (32f).

- (32) a. I usually look up words in the dictionary.
 b. I looked up the word in the dictionary.
 c. They go to the cinema whenever they can.
 d. They went to the cinema on Monday.
 e. You run five miles a day.
 f. You ran to the store to get milk.

Since the future in English does not involve inflection, but is expressed by means of modality (*will*), the periphrasis *be + going to*, or a special use of the present progressive, the English tense system can be reduced to the presence or absence of the feature [past].

Other verb features that can be detected on verbs are [participle] and [infinitive]. The feature [participle] can combine with the feature [past] to distinguish present participles, (33a), from past participles, (33b). The feature [infinitive] indicates that the verb is not specified for tense or aspect, and that it will not participate in agreement.

- (33) a. She is/was singing. Singing: [V], [participle]
 b. She has sung. Sung: [V], [past], [participle]

To summarise, we can list the most relevant morphosyntactic features in English in the chart in (34). Note that gender is not in the chart as it is a semantic feature in English.

(34)

Kind of morphosyntactic feature		Features
Category		[N], [V] or [±N],[±V]
Phi-features	Number	[plural] or [±plural] or [number: (value)]
	Person	[1], [2]
	Case	[nominative], [accusative], [genitive]
Verb features	Tense	[past]
	Other	[participle] [infinitive]

2.5. Interpretable vs. uninterpretable features and the interfaces

In the previous sections we showed that features can be considered formatives of syntax and saw different types. In this section, we will introduce the difference between *interpretable* and *uninterpretable features*, which is central to the syntactic model that we are developing in this course.

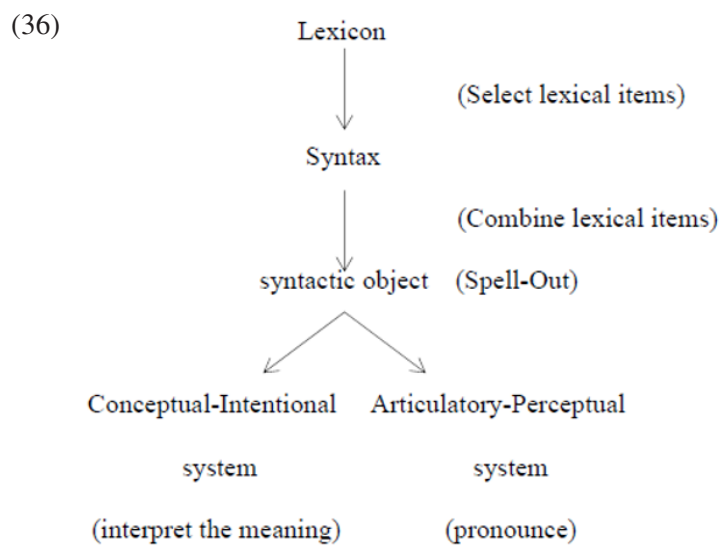
Recall that we claimed that [plural] is an interpretable feature because it has a semantic effect. In the same vein, category features [N] and [V] are interpretable on nouns and verbs, as they determine that they be interpreted as members of a class of words. Uninterpretable features, by contrast, do not have a semantic effect, but they are necessary for complex syntactic objects to be formed. Let us see why.

Take any transitive verb (recall that these are verbs that select an object). It is in its nature to be combined syntactically with an object. But how does this concern syntax?

If we assume that a verb such as *watch* is specified with an interpretable feature [V], which allows us to classify it as a verb, then we must also be able to know from its feature specification that it needs to be combined with a noun. Such a requirement is encoded by means of an uninterpretable noun feature, which we represent as [*u*N]. We can characterise the predicate *watch* as in (35):

(35) *watch* [V, *u*N]

Inasmuch as [*u*N] is useful for the syntax to ensure that we combine a noun with the verb *watch* to form a phrase, we will see next that uninterpretable features are a problem for the cognitive systems with which language interacts. The linguistic model that we propose will offer a solution to handle uninterpretable features, but let us first establish what the problem is by exploring the architecture of the model that we are assuming, which can be represented as in (36).



As shown in the diagram, humans are endowed with a mental dictionary of their language (the Lexicon), where the relevant lexical items (words, morphemes, features, etc.) are stored. To form a sentence, we select a discrete number of these lexical items, and then we combine them using syntax (that is, the set of rules of a given language). The outcome of syntax (the syntactic object) must then be interpreted by two other systems, which interact with language but are actually independent. On the one hand, the syntactic object must be interpreted for meaning by the Conceptual-Intentional system. This is the system we use for thinking, problem-solving, and to assign meaning to the world. On the other hand, the syntactic object must be uttered physically as sound by the Articulatory-Perceptual system. This is what we use to generate sound, whether linguistically meaningful or not (when we scream, for instance). As these two systems interact (or interface) with syntax, they are often called the *interfaces*. Spell-Out is the name that the action of handing over the syntactic object to the interfaces for interpretation/pronunciation receives in the literature.

Let us assume that only interpretable features can be properly ‘read’ at the interfaces after Spell-Out, but uninterpretable features cannot. This is known as the *Principle of Full Interpretation* and is taken to be a general linguistic constraint.

(37) *The Principle of Full Interpretation*

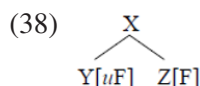
The syntactic structure that is sent to the interfaces contains no uninterpretable features.

This means that in the case of (35), only the feature [V] would be possibly interpreted by the Conceptual-Intentional system. What about the uninterpretable [*uN*] feature then? What do we do with it?

We will assume that uninterpretable features are instructions for syntax and, therefore, once they have served their purpose (e.g., combined a verb with a noun in the case of (35)), they are erased from the syntactic structure. Such an operation is called *feature checking*, and it ensures that the Principle of Full Interpretation is not violated.

But what do we mean by ‘combine’ when we say that a verb such as *watch* in (35) contains the instruction that it must be combined with a noun?

What we mean is that the verb in (35) can be used to build a more complex structure with a lexical item that carries the interpretable counterpart of the uninterpretable feature it carries. In plain words, we mean that the verb *watch* can be used to build a phrase with a noun (as they are specified with the interpretable feature [N], which is the interpretable counterpart of the uninterpretable feature [*uN*]). We call such an operation *Merge*, and we can represent it abstractly as in (38).



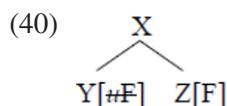
If we are to translate the structure in (38) to an actual linguistic string, then ‘Y’ would correspond to the verb *watch*, and ‘Z’ could be the noun *movies*. ‘X’ would be the phrase *watch movies*.

We close this unit with a question: what kind of phrase is *watch movies*? We are inclined to say that it is a verb phrase, as its distribution is closer to that of a verb, (37a-c), rather than to that of a noun, (39).

- (39) a. She likes to investigate_[V]
 b. *She likes to blackboard_[N]
 c. She likes to watch movies.
 d. The blackboard_[N] is useful.
 e. *The investigate_[V] is useful.
 f. *The watch movies is useful.

It seems, then, that in (38), one of the two elements that Merge is more important than the other, and this is seen in the fact that it names the resulting phrase. But which is it? That is, is the resulting ‘X’ a Z Phrase, or a Y Phrase in (38)? And how is this determined by the syntax?

If we assume that in (38) Merge is driven by the uninterpretable feature of Y, then this lexical item must be ultimately responsible for the creation of syntactic structure. This is why we will say that it is the *head* of the phrase, and that the syntax ‘knows’ this because heads are always the needy elements that cause Merge to happen. By Merging Y and Z in (38), [μF] is checked, as represented by the strikethrough in (40), and, hence, no longer a problem for the interfaces.



In Unit 3 we will learn more about Merge as a syntactic operation, and we will see that when applied successively it results in hierarchical syntactic structure that reflects constituency.

2.6. Suggested exercises and key

2.6.1. Exercises

Exercise 1.

According to Adger (2003: 42) ‘...syntactic relations like agreement access syntactic features and not phonological or purely semantic features.’ Can the *was/were* alternation in the following examples be used to justify this claim?

- (1) The pencil was red.
- (2) The pencils were red.
- (3) The fish was red.
- (4) The fish were red.
- (5) The scissors were red.
- (6) *The scissors was red.
- (7) *The pair of scissors were red.
- (8) The pair of scissors was red.

Exercise 2.

[This exercise has been adapted from Adger (2003: 57)]

Consider the data in (1)-(3).

- (1) Close the door!
- (2) Eat dirt!
- (3) Know yourself!

These are imperative clauses. They have been said to have missing subjects (compare to the data in (4) and (5)).

- (4) Frieda closed the door.
- (5) Kane ate dirt.

There are two possible hypotheses concerning the observation that imperatives do not have a subject:

- Hypothesis 1: Imperatives are full sentences and do have a subject, but this subject is not pronounced.
- Hypothesis 2: Imperatives are not full sentences. They do not have a subject.

Task 1. Look at the sentences in (6) and (7) and discuss how they support one of the two hypotheses above. [Tip: retrieve your answer for Exercise 4 in Unit 1].

- (6) Keep yourself clean!
- (7) Look after yourself!

Task 2. Look at the sentences in (8) and (9). Do these data back up your conclusion for Task 1, or do they contradict it? Explain your answer.

- (8) *Keep myself clean!
- (9) *Look after herself!

Task 3. English has question tags, as shown in (10)-(12).

- (10) Frieda closed the door, didn't she?
- (11) I can come, can't I?
- (12) You won't be there, will you?

Assuming the following Tag Question Generalisation in (13), construct an argument from the data in (14) and (15) either for Hypothesis 1 or Hypothesis 2.

(13) *Tag Question Generalisation*

A tag question is constructed from the auxiliary of the main sentence, which is negated if the main sentence is positive, and which is positive if the main sentence is negative, followed by a pronoun which has the same phi-features as the subject.

- (14) Close the door, won't you!
- (15) *Close the door, won't he!

Exercise 3.

Evaluate whether adverbs and adjectives form a separate class by examining the data given and answering the questions in each of the tasks. In this exercise the focus is on argumentation, so there is no need for you to use features in your answers.

1. Morphological evidence:

- (1) unhappy, unusual, unimportant
- (2) *unextremely, *unbeautifully, *unquickly

but

- (3) *unblue, *unclever, *unopen
- (4) unexpectedly, unusually, unreliably

- (5) happier, shorter, prettier
- (6) *extremelier, *beautifullier, *quicklier

but

- (7) more expensive, more intelligent, more difficult
- (8) more incredibly, more interestingly, more relevantly

2. Syntactic distribution:

- (9) *ridiculous expensive shoes
- (10) ridiculously expensive shoes

Task 1. What do (1) and (2) tell you about adjectives and adverbs?

Task 2. How do (3) and (4) change what you have concluded based on (1) and (2)?

Task 3. What do (5) and (6) tell you about adjectives and adverbs?

Task 4. How do (7) and (8) change what you have concluded based on (5) and (6)?

Task 5. According to (9) and (10), can an adjective modify another adjective? Can an adverb modify an adjective?

Task 6. Summarise your findings by considering whether morphological evidence and syntactic distribution facts allow us to classify adjectives and adverbs as a separate class or not.

2.6.2. Key to exercises**Exercise 1.**

Yes, they can. (1) and (2) show that the *was/were* alternation is related to the singular/plural nature of the subject. This is so even in (3) and (4), where the plural is not phonologically uttered. The pair in (5) and (6) shows that for a plural noun *scissors*, the choice is the plural form *were*, whereas for the singular *pair of scissors* the choice is the singular form *was*. The semantics of *scissors* and *pair of scissors* is similar.

Exercise 2.

Task 1. As was concluded in Exercise 4 in Unit 1, anaphors agree with their antecedent in person, number and gender. This tells us that if there is an anaphor, then there is an antecedent (that is, a subject for the imperative). Our observation supports Hypothesis 1.

Task 2. With the new data, which we can compare to (1)-(3), we observe that the silent subject can only be *you*, as it agrees with *yourself* in (1)-(3), but not with *myself* and *herself* in (8) and (9). Thus, the data support the conclusion in Task 1.

Task 3. Assuming the Tag Question Generalisation, we would expect (14) to be grammatical if Hypothesis 1 is correct (imperatives have a silent subject, and this is *you*), and (15) to be ungrammatical.

Exercise 3.

Task 1. Adjectives, (1), can be affixed with *un-*, but adverbs cannot, (2).

Task 2. It seems that what was concluded in Task 1 does not hold for all adjectives and adverbs. Some adjectives cannot be affixed with *un-*, (3), while some adverbs can, (4).

Task 3. Adjectives can be inflected for comparative, (5), while adverbs cannot, (6).

Task 4. What was concluded in Task 3 does not seem to hold for all adverbs. Some adverbs can be made comparative provided the analytic form *more* is used.

Task 5. An adjective cannot modify another adjective, (9), but an adverb can modify an adjective, (10).

Task 6. *Un-*affixation does not enable us to tell adjectives apart from adverbs: while some adjectives allow *un-*affixation and some do not, the same is true for adverbs.

Comparative inflection does not enable us to tell adjectives apart from adverbs completely, either, as both adjectives and adverbs allow comparatives. It has been observed, however, that comparatives are formed synthetically (*-er*) or analytically (*more*) depending on the adjectives, while adverbs only allow analytical comparatives. Syntactic distribution data allows us to tell adjectives and adverbs apart, as they show different behaviour with respect to modification. In conclusion, while morphological evidence does not clearly support the classification of adjectives and adverbs as different classes, syntactic evidence does.

2.7. Suggested reading

Koenenman & Zeijlstra (2017). Chapter 1. *Categories and Features*.

It offers a very comprehensive overview of the contents covered in Unit 2, and it contains brief exercises to back up the descriptive knowledge acquired in the two previous foundation grammar courses.

UNIT 3. Phrase structure and dependencies

Unit 3 focuses on connecting constituency and theta-roles to phrase structure. It also continues to explore Merge and feature checking as syntactic operations. The operation Merge is further characterised as First Merge and Second Merge depending on whether the resulting structure is that of head-complement or if it involves the projection of a Specifier. The operation Adjunction is also brought into the picture as a structure-building mechanism that is different from First and Second Merge. The notion of c-command and its role in syntactic dependencies is discussed, as is the syntax of ditransitive verbs as an argument in favour of the functional category little *v* and the need for head movement.

3.1. Objectives of the unit

- Revise the notion of *constituency* and connect it to *phrase structure*.
- Introduce some terminology related to *tree structure*.
- Revise the notion of *theta-roles* and connect it to *phrase structure*.
- Further explore the *Merge* and *feature checking* operations.
- Introduce *First Merge*, *Second Merge*, and *adjunction*.
- Introduce *minimal*, *maximal* and *intermediate* projections.
- Introduce the notion of *c-command* as central to syntactic dependencies.
- Argue for the functional category *little v* through the syntax of ditransitive verbs.
- Introduce *head movement* (V-to-*v*).

3.2. Constituency and phrase structure

As you learnt in previous courses, words are arranged into phrases, which are called *constituents* and can be diagnosed as such by means of several tests. Revise your notes from last year and review the different tests that can be used to diagnose if a group of words form a constituent or not. These tests are *transposition/movement*, *substitution*, *ellipsis*, and *coordination*.

Constituent structure corresponds to a node (the result of Merge) if we represent complex syntactic structure in the form of a tree. In a sentence such as (1), you can easily diagnose four different constituents. What are they?

(1) The students read the books.

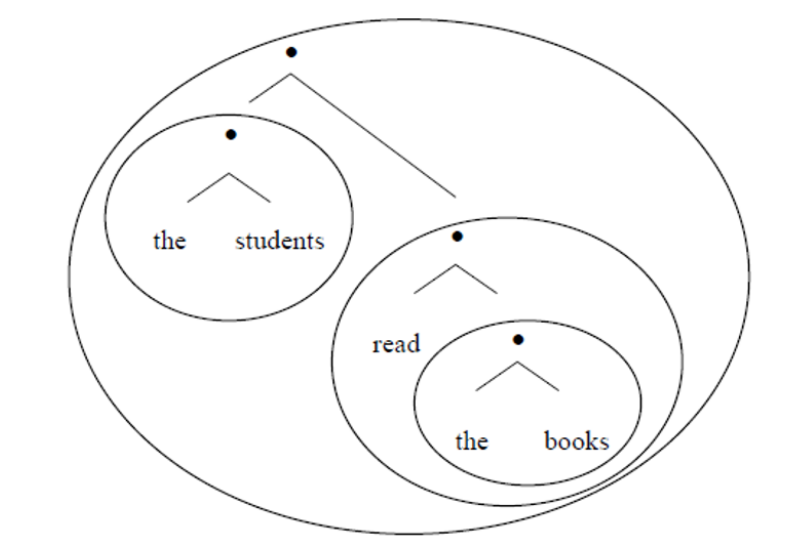
Constituency tests show that (1) is made up of the following constituents:

- (i) *the books*
- (ii) *read the books*
- (iii) *the students*
- (iv) *The students read the books*

They also show that *students read* is not a constituent, and that *read the* is not a constituent either.

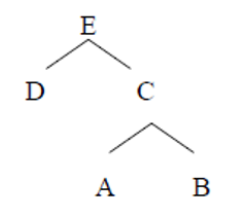
By successively applying Merge to pairs of lexical items bottom-up, we obtain the binary-branching tree in (2). In it, all resulting phrases, which have been marked with a dot, are diagnosed as constituents, as indicated, as well, by the circles.

(2)



At this point some terminology related to tree structure must be revised. Take an abstract binary-branching tree structure such as (3).

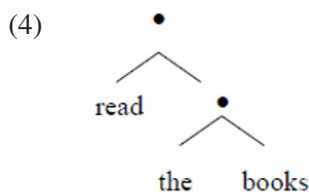
(3)



In (3), the joints are called *nodes*, and they are labelled with *node labels* (C, E). *Nodes* are connected by *branches*. Continuing with the tree metaphor, the top node (e.g., E) is called the *root node*. In terms of hierarchy, it is said that a given node *dominates* those below it in the tree. That is, in (3), C dominates A and B, and E dominates D and C. Nodes A, B and D do not dominate anything. A node that dominates others is referred to as *mother*, while two nodes dominated by the same *mother* (e.g., A and B, or D and C) are *sisters*. So E is the mother of D and C, as it dominates them, but what is its relationship with A and B? Does it also dominate them?

The answer is that it does, as these nodes are below it in the tree. However, it is useful to distinguish between *immediate domination* (e.g., E *immediately* dominates D and C, but it only dominates A and B) and *domination*.

Note that if you overlap the tree in (3) with the bottom part of the tree in (2), you get the phrase *read the books*. We will ignore phrase labels for the moment and keep the dots. Note, though, that former E (the top node in (4)) is a constituent that includes all the material that the top node dominates, and the same happens with the node on top of *the books*. As *read the* does not correspond to a node (that is, there is no dot resulting from the combination of these two elements), it is, predictably, not a constituent.



3.3. Theta-roles and phrase structure

As you also saw in previous courses, the element that specifies who/what takes part in the event expressed by a sentence is the *predicate* and the participants are the *arguments*. Verbs are classified as different kinds of predicates depending on the number of arguments they require (*one-place*, *two-place*, *three-place predicate*, etc.).

The semantic notions of arguments and predicates have corresponding syntactic functions such as subject, direct object, indirect object, etc. Thus, in this section we explore how *argument structure* translates into syntactic features and, ultimately, into hierarchical phrase structure.

It is assumed that verbs must be recorded with information about the syntactic categories they can be combined with. This is known as *categorical-selection* (*c-selection* for short) and is syntactically encoded in the predicate by means of uninterpretable features. Yet, it is also true that verbs also exert semantic control (apart from syntactic, in the form of c-selection) on the kinds of arguments they allow. For example, many verbs can only have an Agentive subject, (5), and not respecting such a requirement, which is called *semantic selection* (or *s-selection* for short) results in an anomalous interpretation of the sentence.

- (5) a. John watches TV.
b. #The chair watches TV.

It is a general principle that a constituent cannot have more than one theta-role (e.g., AGENT, THEME, GOAL, etc.), and that a theta-role cannot be left unassigned to a syntactic constituent. This is known as *The Unique Theta Generalisation*, (6), and in later Units it will have a central role in exploring the syntax of certain kinds of predicates.

(6) *The Unique Theta Generalisation*

Each theta-role must be assigned but a constituent cannot be assigned more than one theta-role.

3.4. More on Merge and feature checking

Recall that at the end of Unit 2, we introduced the idea that syntax gets instructions about what can be Merged with what by means of uninterpretable features. That is, if a syntactic object carries an uninterpretable feature of some kind, it must Merge with a syntactic object that carries an interpretable matching feature, so that the outcome satisfies the Principle of Full Interpretation, which bans uninterpretable features from the interfaces (that is, the Conceptual-Intentional and Articulatory-Perceptual systems).

Let us try to ‘translate’ the argument structures in (7)-(9) into syntactic feature specifications.

- (7) cry (verb) [1 <NP>]
 (8) watch (verb) [1 <NP>, 2 <NP>]
 (9) send (verb) [1 <NP>, 2 <NP>, 3 <NP>]

For all kinds of predicates in (7)-(9), an interpretable category feature [V] would be specified. For (7), by contrast, only an uninterpretable category feature [μ N] would be necessary to ensure that syntax combines

this predicate with an Agent with nominal properties, (10). The transitive verb in (8) would need two [uN] features to be specified so that the syntax can combine it with constituents of the relevant class that satisfy all its theta-roles, (11), and the ditransitive in (9) would need three, (12).

(10) cry [V, uN]

(11) watch [V, uN, uN]

(12) send [V, uN, uN, uN]

Note, however, that for a verb such as *send* the Benefactive theta-role can be performed by a PP, as shown in (13).

(13) I sent flowers to my mother.

This means that *send* is stored both as (12), and as (14), thus giving rise to two possible syntactic structures that are fully grammatical.

(14) send [V, uN, uN, uP]

Observe that in Unit 2 we discussed how we should characterise prepositions concerning category features. For the sake of the syntactic argument, we will avoid this debate here and use ‘P’ as shorthand for ‘features that characterise a Preposition (whatever these are and which kind)’.

Given the feature specification for the predicates in (10), (11) and (12), if one intends to form a sentence with them, then other lexical items that meet the feature requirements imposed by these predicates (that is, that carry an interpretable [N] or [P] feature) must be selected from the Lexicon as well. Then these lexical items have to be Merged, two at a time, and as many times as required until all uninterpretable features are checked. In the next section we show how Merge and feature-checking result in the structure of phrases.

3.5. The structure of phrases: Merge and adjunction

Let us show how the sentence in (15) can be derived step by step, starting first by listing the relevant syntactic features for each of the three elements that make up the sentence, (16).

(15) Students read books.

(16) Students [$N, plural$]

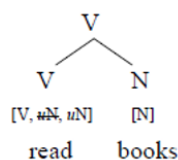
read [V, uN, uN]

books [$N, plural$]

Building our structure bottom up, the first two elements we would select for Merge would be the predicate *read* and the noun *books*. It does not really matter in which order you Merge them (*read books* or *book reads*), but the resulting syntactic object at some point will be linearised according to some word order parameter that requires heads to precede their complements, as in English, or to follow them, as in Basque. For the sake of clarity, though, we will represent Merge of the two different lexical items respecting the surface order of English.

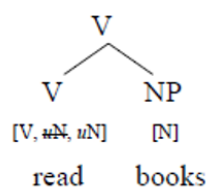
(17) Step 1.

- a. Merge the predicate and one of its arguments.
- b. Checking of [uN] takes place.
- c. The resulting syntactic object is labelled as a projection of the head.



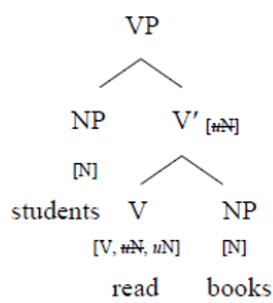
Syntactic objects with no c-selectional features that need to be checked (e.g., *books*) are called *maximal projections* (as they project no further), and are annotated as XP, X corresponding to the category of the head. Thus, in a tree such as that in (17), it would be more accurate to represent *books* as a NP, as in (18).

(18)



Lexical items (e.g., *read* or *books*) are *minimal projections*. So, when *books* merges with the verb *read*, *books* is both a *minimal* (N) and a *maximal* (NP) projection. A tree such as (18) illustrates a *head-complement* structure, and it arises from *First Merge* (that is, our first application of Merge). Note, though, that if we stopped the derivation at this point and tried to send the syntactic object for interpretation at the interfaces, the Principle of Full Interpretation would be violated, as there is still an uninterpretable feature to be checked in (18). This means that *Second Merge* is needed (that is, a second application of Merge, where the resulting syntactic object in (18) merges with a lexical item that carries an interpretable category feature [N]). This operation turns our former root projection V into an *intermediate* projection, V'. To indicate that this second application of Merge checks the [uN] feature that was still unchecked in (18), we recopy it at the intermediate level and indicate that it has been checked with a strikethrough. The position occupied by the Agent *students* is known as *Specifier*. As there are no more uninterpretable features left to check, the root node can now be a maximal projection, VP.

(19)



But could we continue enlarging the structure of (19) somehow? The *Extension Condition*, (20), constrains this issue, and establishes that a syntactic derivation can be continued, but only by applying operations to the root projection of the tree.

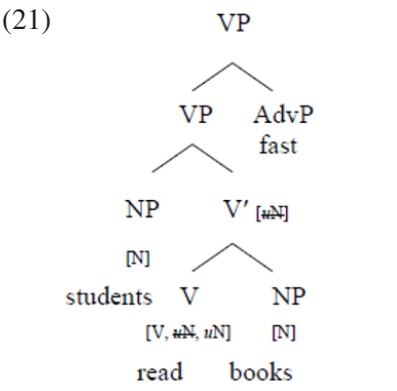
(20) The Extension Condition

A syntactic derivation can only be continued by applying operations to the root projection of the tree.

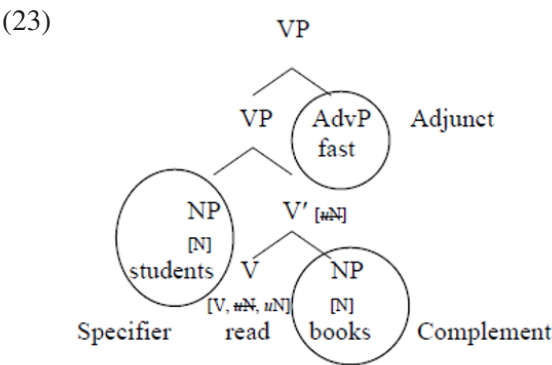
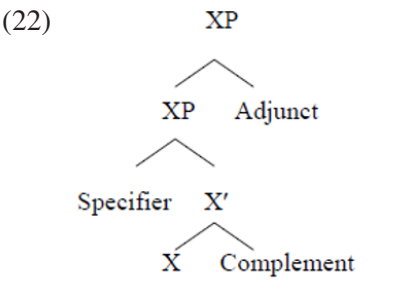
In a sentence such as the one we analysed in (19), we certainly cannot add more arguments because *read* is transitive and it only takes an Agent and a Theme, and we have already Merged those (hence satisfying the Unique Theta Generalisation, (6)). However, we can add material that is modifying rather than argumental.

Modifying material is not required by uninterpretable features. These kinds of constituents are called *adjuncts*, and they are Merged to a maximal projection once all the c-selectional requirements of the predicate have been satisfied.

Let us add an adjunct to (19), then. This could be an adverb such as *fast*, or a Prepositional Phrase such as *on Mondays*, and it would be Merged as shown in (21).



For any head X, then, the structure in (22) can be maximally built. (23) shows how we can talk about the different parts of the tree in (21) by using the terms *Complement*, *Specifier* and *Adjunct*.



3.6. C-command

As we saw in the previous section, Merge builds tree-like structures such as (22). As lexical items are Merged in pairs (two at a time), *sisterhood* is an important relation in syntax. Actually, so far, we have considered it

to be the syntactic configuration that allows uninterpretable features to be checked. Sisterhood is, however, not the only relevant syntactic relation. In this section we introduce *constituent-command* (*c-command* for short), which can be defined as follows:

(24) *C-command*

A node X c-commands its sister and anything dominated by its sister.

According to the definition in (24), sisters c-command each other (e.g., in a tree like (23), the V *read* c-commands the NP *books* and the other way around). The second part of the definition, namely that X c-commands anything dominated by its sister, means that in a tree like (23), the NP *students* c-commands the V *read*, the NP *books*, and the *read books* intermediate position. The NP *students*, by contrast, does not c-command the adjunct *fast*.

Several syntactic phenomena are crucially dependent on c-command being obtained between certain elements. Recall, for example, that in the previous units you learnt that anaphors must agree with their antecedents in person, number and gender in English. Below we will show that such a condition does not suffice to capture all the data.

Consider the pair in (25):

(25) a. I talk to myself.

b. *Myself talk to I.

There is certainly agreement concerning phi-features both in (25a) and (25b). Yet, the latter is ungrammatical. We could of course say that it is a matter of precedence in that the antecedent must precede the anaphor and not the other way around. The problem with this, though, is that we would still not be able to make sense of the ungrammaticality of examples such as (26a).

(26) a. *The woman I met talked to myself.

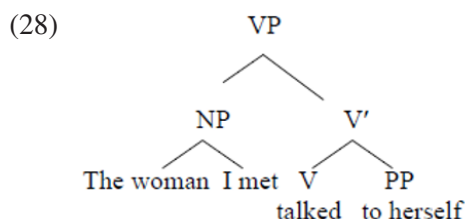
b. The woman I met talked to herself.

Adding c-command to the requirements of anaphors allows us to account for the ungrammaticality of (26a) and the grammaticality of (26b). Let us refine our generalisation on anaphors as follows:

(27) *The Anaphor Generalisation*

Anaphors must agree in phi-features with a c-commanding antecedent.

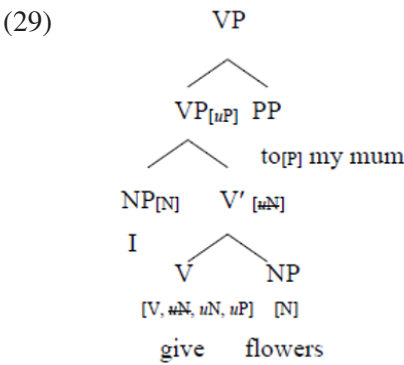
Assume a (simplified) analysis for the examples in (26):



In (28) (which corresponds to (26b)), we see that the entire NP *The woman I met* agrees in phi-features with the anaphor *herself*. Moreover, the NP c-commands the anaphor *herself*. If, by contrast, you replace *herself* with *myself* in (28), you will obtain a tree representation for (26a). The NP that c-commands *myself* does not agree in phi-features with it, and *I*, which does, is too deep in the structure of the NP to c-command *myself*. So anaphor licensing does not reduce phi-feature agreement: c-command is also a requirement. This is also the case for negative polarity items such as *any*, *anything*, etc. in relation to negation. The relevance of c-command for the distribution of negative polarity items has been left for you to work out in the Exercises.

3.7. Ditransitives and little *v*

To draw this unit to a close, let us address one last problem. Thus far, we have seen what the syntactic structure for two-place predicates is. But what is the structure of three-place predicates? These certainly constitute a problem for our model, as it is not obvious where the second complement would go in a structure such as (22). Recall that whatever structure we propose, it must make the right predictions for constituency. So, is it possible that the second complement occupies the position of an adjunct, as in (29)?

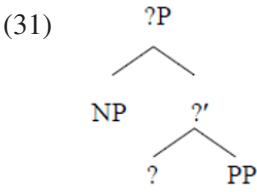


Such an option has several theoretical problems, indeed. A serious one is that we had concluded that adjunction did not involve the checking of features, but in the case of (29), adjunction would be motivated by the need to check the *[uP]* feature that the ditransitive verb carries to ensure that the Benefactive theta-role is assigned to a constituent.

Let us explore another binary-branching alternative. Crucially, the coordination constituency test points in the direction of an analysis where both objects form a constituent independently of the verb.

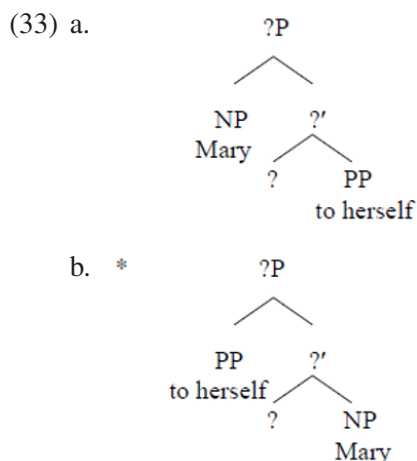
(30) I give [flowers to my mum] and [chocolates to my dad].

But what would a structure where both objects form a constituent independently of the verb look like? It would have to be something like (31). Note that some nodes have been labelled as ‘?’ to capture the right constituency predictions in the structure. We will soon see what category they are a projection of.



The distribution of anaphors in the context of ditransitive verbs, (32), shows that the structure is most likely correct. In the Exercises, you will be able to further investigate how the structure accommodates negative polarity items, thus further strengthening the proposed analysis.

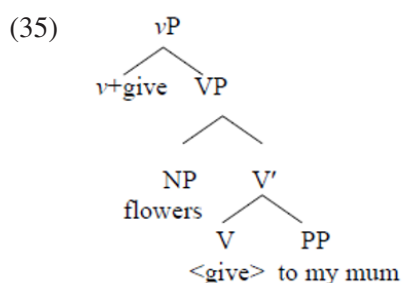
- (32) a. I showed Mary to herself (in the mirror).
b. *I showed herself to Mary (in the mirror).



But what category is ? and ?P, then? Given that it is the predicate that assigns the theta-roles to its PP and NP complements, we would like to keep it close to them. Hence, ? is simply V (and ?' is V', and ?P is VP). But if this is so, then how does it end up being pronounced higher than its two objects? We find inspiration to answer this question in some languages of the world that reveal a semantic relation of *causality* between the verb and some functional category in the verbal domain. In Catalan and Spanish, for example, we get a different word order if the causal verb *fer / hacer* is included in the structure. Consider the data in (34).

- (34) a. En Joan menja. $S_{En\ Joan} V_{menja}$
 b. La Laia fa menjar en Joan. $S_{La\ Laia} fa V_{menjar} S_{en\ Joan}$
 c. *La Laia fa en Joan menjar.

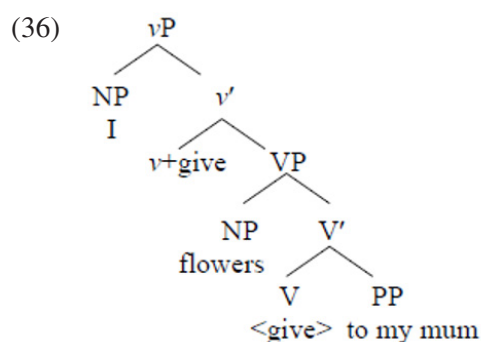
There are just two options to explain the change in word order: either the subject *en Joan* moved to the right in (34b), or else the verb *menjar* moved to the left when combined with the causative *fer*. Data in other languages of the world suggest that it is the verb that moves. But where does it move to? We propose here that English has a functional category, known as little *v*, that encodes causality and that triggers upward movement of the lexical verb to incorporate into it, thus eventually surfacing higher than its two objects. The proposed analysis is shown in (35).



Such an analysis of ditransitive verbs has been referred to in the literature as a VP-shell analysis, as the ditransitive V is somehow enclosed between its two internal arguments. For ditransitive verbs, movement of the lexical verb (big V) to little *v* is essential to account for the observations about constituency.

For transitive verbs, though, we might be tempted to assume that little *v* is not required. However, this is an operation that takes place across the board in English. Hence, we will include it in the syntactic representation of all kinds of predicates. If we do so, then what is the position of the subject?

We had represented it as the Specifier of VP, but if we continue to do so, then we are predicting that the verb will surface higher than the subject in English, and this is not the case. Recall that all heads can take a complement (in the case of little *v*, the complement is the VP), and a Specifier. The latter offers a position for the subject. We can complete our analysis as in (36), then. As V is a head and little *v* as well, the movement of the verb is an instance of *head movement*.



3.8. Suggested exercises and key

3.8.1. Exercises

Exercise 1.

[This exercise has been adapted from Adger (2003: 99)]

The following two sentences look similar in English.

- (1) He looked up the number.
- (2) He walked up the hill.

However, it turns out that they have rather different constituent structures, as shown by the fact that only in the first can *up* occur at the end.

- (3) He looked the number up.
- (4) *He walked the hill up.

Task 1. Show by means of the movement test (tip: use a cleft of the kind *It is...that...*) that (1) and (2) have different constituent structures for the words that are underlined.

Task 2. Show, using replacement, that (1) and (2) have different constituent structures (tip: use the word *there* as a test for a PP constituent).

Exercise 2.

In this exercise you are required to end up formulating a Negative Polarity Item Generalisation in Task 3. You need to work on Tasks 1 and 2 first, though.

Task 1. Words such as *anything* are negative polarity items and seem to have some special requirements. By looking at the data in (1) and (2), work out what is a necessary condition for them to be grammatical in a sentence.

- (1) She didn't see anything.

(2) *She saw anything.

Task 2. Now consider the data in (3) and discuss whether your generalisation accounts for them or not.

- (3) a. Nobody saw anything.
b. *Anybody saw nothing.

Task 3. Recall that for anaphors, precedence of the antecedent was not a determining factor, as there were cases with a preceding antecedent that were ungrammatical. What was the relevant relation there and how does it apply to the data in (3)?

Exercise 3.

In the previous Exercise you should have concluded that c-command plays a role in the distribution of negative polarity items. Bearing this in mind, explain the ungrammaticality of the following example (tip: analyse *Nobody's cat* as an NP made up of two other NPs, *nobody's* and *cat*).

(1) *Nobody's cat likes anybody.

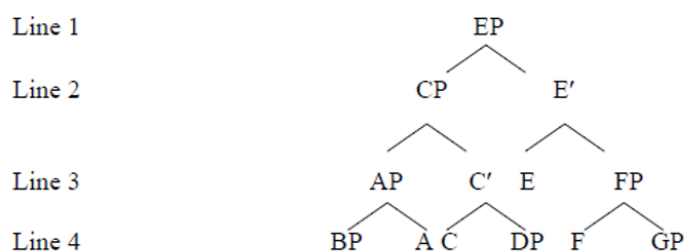
Exercise 4.

Using negative polarity items (e.g., *anything/anyone*) make up examples with three-place predicates, along the lines of the examples with anaphors that we saw in class. Show how these can be used as an argument in favour of the VP-shell analysis for ditransitives.

Exercise 5.

[To be done in the classroom, for fun].

The Fibonacci series is a mathematical curiosity that pervades the natural world. Many plants, for instance, have a quantity of petals that corresponds to a number taken from the series. In the Fibonacci series, each number is the sum of the two that precede it (e.g., 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, etc.). It seems that it emerges in syntactic structure, as well. Just check the following tree. I have built it filling every Specifier and Complement position. If you count the number of XPs in each line, what numbers do you get?



Shall we add another layer and count again?

3.8.2. Key to exercises

Exercise 1.

Task 1.

- (1) a. *It is up the number that he looked.
b. It is up the hill that he walked.

Exercise 5.

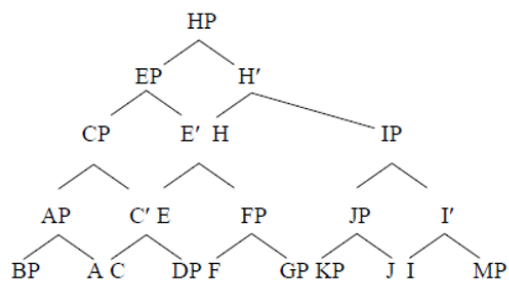
Line 1 (1)

Line 2 (1)

Line 3 (2)

Line 4 (3)

Line 5 (5)



and so on...

3.9. Suggested reading

Radford (2004). Chapter 2. *Structure*.

Radford's chapter complements the contents of Unit 3 very well, but students must be warned about the fact that the tree representations he provides should be used with caution, as they will certainly show significant differences to the ones that we have developed so far during the course (e.g., Radford's trees include T(ense), which is seen in Unit 4, and do not include little *v*).

UNIT 4. Functional categories and movement I (TP)

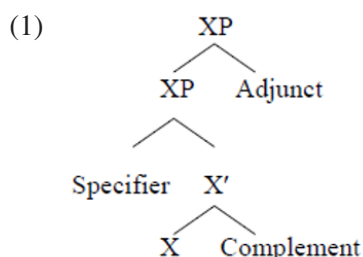
Unit 4 focuses on expressing the need for functional categories other than *v* such as *T(ense)*, *Perf(ective)* and *Prog(ressive)*, and it is shown that they obey a *Hierarchy of Projections* in English. In this unit we also introduce *XP-movement* to account for the relocation of the subject and show that the syntactic relation of *Agree* is another way of checking morphosyntactic features. We also explore the phenomenon of *do-support* in connection to the functional category *Neg(ation)* and the *Pronouncing Tense Rule*. The unit closes with the syntax of *unaccusative* and *unergative* predicates (two kinds of intransitive predicates), and its relation to the *Uniform Theta Assignment Hypothesis*.

4.1. Objectives of the unit

- Introduce and argue for the functional category *T(ense)*.
- Introduce *XP-movement*.
- Introduce other functional categories: *Perf(ective)*, *Prog(ressive)*.
- State the *Hierarchy of Projections*.
- Explore *valuation* of uninterpretable T features and connect it to checking via *Agree*.
- Connect the functional category *Neg(ation)* to *do-support* by means of the *Pronouncing Tense Rule*.
- Explore the syntax of intransitive verbs (*unaccusative* vs. *unergative*) and connect it to the *Uniform Theta Assignment Hypothesis*.

4.2. Sentences as Tense Phrases

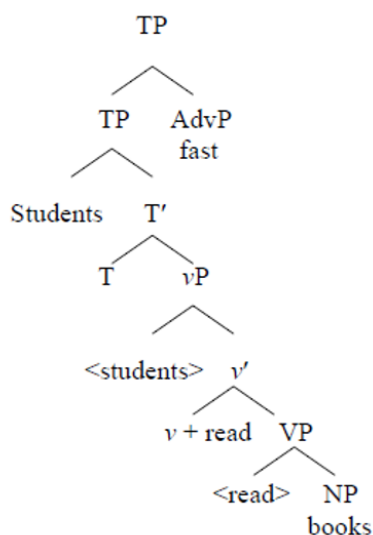
Thus far, our tree analysis has expanded the structure of a simple clause from VP to *vP*. In this unit we will see that we need to conceive sentences as being projections of a functional category head, which we will refer to as *T(ense)*. There is a theoretical advantage in doing so: if a ‘sentence’ is not a special category on its own, but a projection of T (i.e. a TP), then our system is consistent with regards to the fact that all syntactic categories can project into a phrase and have the structure in (1).



That is, our analysis of simple sentences can be now updated by replacing X with T. And what would the complement of T be, then? And its Specifier? The tree representation for the sentence in (2) should allow you to see if you answered this question correctly.

- (2) a. Students read books fast.

b.



What we can observe in (2b) is that the subject *students* seems to have moved from the Specifier of little *v* to the Specifier of TP, as indicated by the angled brackets. We saw that big *V* moved to be combined with little *v* at the end of the previous unit, so subject movement does not really come as a surprise. Note, though, that we had defined V-to *v* movement as *head movement*, as it involved a lexical head (*V*) moving to a functional head (*v*). This is certainly not the case in subject movement, which involves an NP in a Specifier position moving to another Specifier position. We call this movement *XP-movement*.

We will see how these two types of movement are syntactically argued for later in the unit. For the moment, let us return to the claim that sentences should be analysed as TPs, that is as phrasal projections of a *T* syntactic head. Let us see some evidence for the existence of a functional head where Tense features are encoded.

Recall that in Unit 3 we discussed the fact that several tests can be used to determine whether a linguistic string was to be considered a constituent or not. Take the examples in (3) and apply the movement and substitution tests, for instance. What can you observe?

- (3) a. I write letters.
b. I wrote letters.
- (4) a. What I do is write letters.
b. You write letters and I do too.
- (5) a. What I did is write letters.
b. You wrote letters and I did too.

You should have observed that you used *do* when testing (3a) and *did* when testing (3b). This indicates that the tense features in (3) and (4) are different and to a certain extent independent of the verb. For this reason, we will assume that the feature [past] we saw in Unit 2 is actually not a specification of lexical verbs, but of the functional head *T*.

T may not be the only functional head that we need to fully represent the syntax of English clauses. In English, apart from the lexical verb (and little *v*), it is possible to have one or more auxiliaries. These express dedicated semantic notions and determine the morphology of the verb. Consider the following sentences:

- (6) a. I eat.
b. I ate.
c. I may/might eat.
d. I am eating.

- e. I was eating.
- f. I may/might be eating.
- g. I have eaten.
- h. I had eaten.
- i. I may/might have eaten.
- j. I have been eating.
- k. I had been eating.
- l. I may/might have been eating.

Can you identify the different tenses of English?

(6a) is in the simple present, and (6b) in the simple past. (6c) contains the modals *may/might*, and so do (6f), (6i) and (6l). (6d) and (6e) are, respectively in the present and past progressive, while (6g) and (6h) are in the present and past perfect. The examples in (6j) and (6k), furthermore, show that the present, the past, the progressive and the perfect can be combined into the present perfect progressive, (6j), and the past perfect progressive, (6k).

We can thus entertain the idea that these auxiliaries (that is, the perfective, and the progressive) are also functional heads that interact with T (either specified as [past] or not). If this is so, then we would add Perf(ective) and Prog(ressive) to the inventory of functional categories that are relevant for the English clause.

And what about modals?

These are assumed to be T heads, as when they occur in a sentence, the verb is necessarily uninflected for tense, (7).

- (7) a. I ate.
- b. I can eat.
- c. I could eat.
- d. *I can ate.
- e. *I could ate.

The data in (8) further show that when the functional categories T, Perf and Prog occur in the same sentence, they do so in a constrained way. In (8) the order of auxiliaries for some of the sentences in (6) has been shifted, and their morphology updated according to the new position they occupy. The result is clearly ungrammatical.

- (8) a. *I eat may. (vs. I may eat.)
- b. *I might eat being. (vs. I might be eating.)
- c. *I may eat had. (vs. I may have eaten.)
- d. *I eat be had mighting. (vs. I might have been eating.)

This tells us, on the one hand, that we have intuitions about the forms that verbs and auxiliaries take when following other auxiliaries (e.g., *I may eat* vs. *I am eating*; *I have been eating* vs. *I have eaten*), and that these seem to occur in a fixed hierarchy. We will refer to such an established ordering as the *Hierarchy of Projections* and will include *v* and *V* in it as well.

- (9) *The Hierarchy of Projections*
 T > (Perf) > (Prog) > *v* > *V*

The '>' symbol indicates that a given category is higher than the one to its right, and lower than the one to its left in tree representations. The parentheses are used to indicate that a given functional category is optional. In this way, an English clause will minimally contain T>*v*>*V*.

If T carries the feature [past] (or not if the tense is present), it is logical to think that other functional heads may carry features too. But what are they?

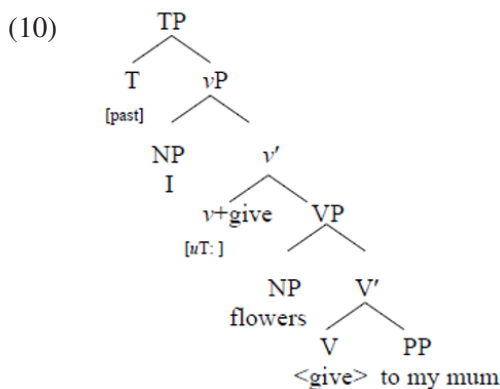
As these functional heads are dedicated to the expression of perfectiveness and progressive aspect, we will be assuming that this is precisely the kind of feature they carry: Perf[perfective] and Prog[progressive]. We will further assume that this feature is interpretable and ultimately related to the expression of tense.

In the next section we address the following issue: if T (and Perf and Prog) and *v*/V are independent categories, how is it that tense is morphologically realized on *v*/V (and on the auxiliary closest to T), and that auxiliaries take different forms depending on whether they occur with other optional elements in the Hierarchy or not?

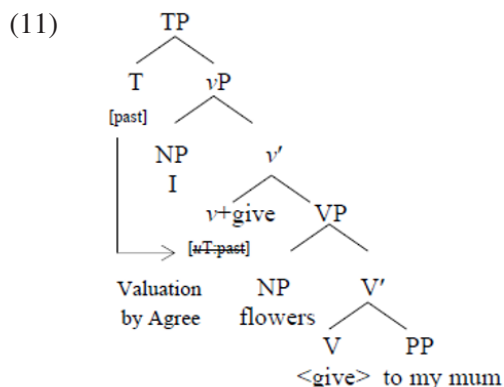
4.3. Checking via Agree

Recall that in Unit 2 our discussion on morphosyntactic features brought us to consider the possibility that features may be open for a value, that is, of the kind [F: (value)]. At that point we had not yet discussed interpretable and uninterpretable features. But now that we have, suppose that *v* carries an uninterpretable T feature with an open value. This would be represented as [*u*T:] and would be consistent with the view of uninterpretable features as instructions for Merge that we have developed thus far. In other words, if *v* carries such a feature, it actually requires Merge with T to occur. So, to a tree representation such as (10) (which we saw up to the *v*P projection in Unit 3), the T layer would be added. Note that the T head carries an interpretable [past] feature, and that the uninterpretable T feature of *v* must be given a value.

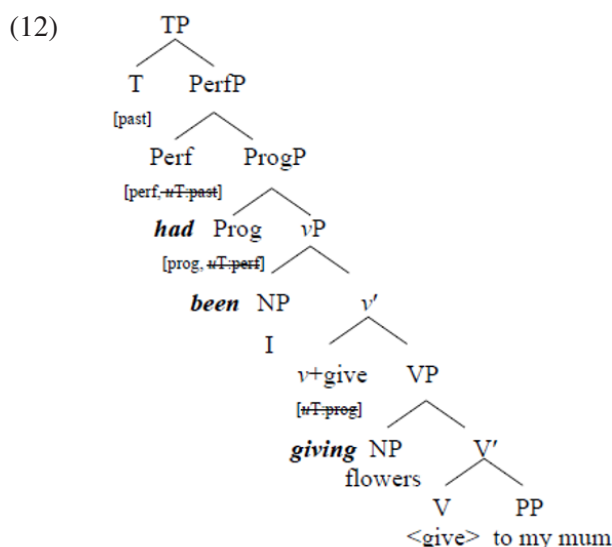
Let us assume that the value for the [*u*T:] feature is supplied by the feature specification of T and results in the verb *give* surfacing as the past form *gave*. This mechanism should be thought of as feature checking but, crucially, it does not involve sisterhood.



Note, nonetheless, that another relevant syntactic relation was mentioned in Unit 3. C-command was shown to be responsible for the correct distribution of anaphors and negative polarity items. Hence, the fact that c-command is also relevant for feature checking does not come as a surprise. Feature checking under c-command is called *Agree*. The tree in (10) shows how [*u*T:] in *v* is valued as [*u*T: past] once the T head is Merged into the structure. After valuation the uninterpretable [*u*T:] feature of *v* is checked and, hence, eliminated from the derivation.



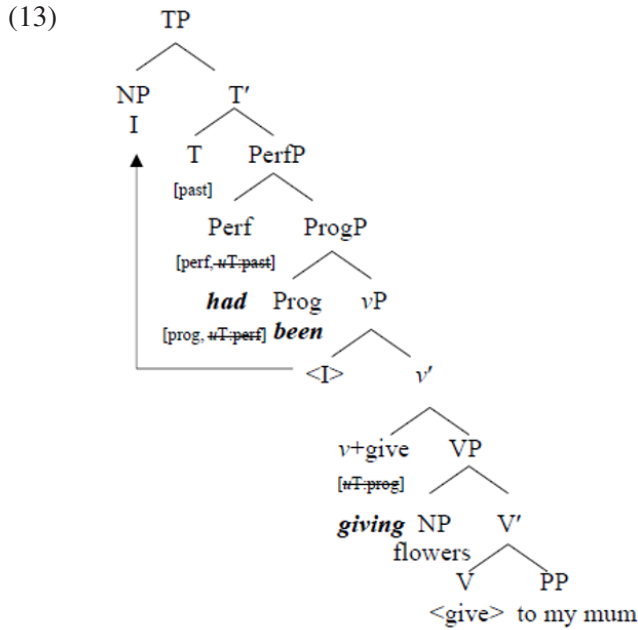
At this point, you may have guessed how the different auxiliaries change in morphological form depending on how close they are to T and to other auxiliaries. In the same way that *v* carries an unvalued uninterpretable Tense feature, so too do auxiliaries. This can be observed in (12) for the past perfect continuous, for example.



Note that the Hierarchy of Projections determines the order in which the functional projections are Merged, and that the Perfective auxiliary will be pronounced as *had* (in being valued as past), and little *v* (to which V adjoined earlier) will be pronounced with an *-ing* progressive morpheme. For the moment, do not worry about the fact that the subject seems to be too low in (12). We will talk about it in the next section.

4.4. Movement: what drives it?

As we observed in connection to the tree in (12), the subject seems to be represented too low, with two auxiliaries surfacing to its left. By now, you probably know enough about syntax to conclude that this can be fixed by resorting to movement. Actually, you may even know where the subject moves to. Consider the tree in (13).



If T is a syntactic head (which takes the vP as its complement), it must also make available a Specifier position that allows the subject to be repositioned. What triggers movement, though? An answer to this question will be given once we have seen yet another instance of head movement (different from V-to- v movement) that concerns auxiliaries.

Let us add negation to a sentence such as (14), which contains different kinds of auxiliaries. Negation is a functional category, Neg(ation), that enjoys some Merge flexibility, as it can occur in several positions of the Hierarchy of Projections. We focus here on *sentential negation*, that is, the negation of the entire clause, where (14) can be paraphrased by (15).

(14) I have not been playing the violin.

(15) It is not the case that I have been playing the violin.

So, it is clear that sentential negation is quite high in the Hierarchy of Projections, but we still need to decide where exactly it fits. In (16) and (17) two possibilities are shown:

(16) $T > \text{Neg} > \text{Perf} > \text{Prog} > v > V$

(17) $\text{Neg} > T > \text{Perf} > \text{Prog} > v > V$

We can test which one is correct by looking at sentences with modals, which are T heads.

(18) I [T may [Neg not [Perf have [Prog been [$v + V$ playing]]]]] the violin.

(19) *I [Neg not [T may [Perf have [Prog been [$v + V$ playing]]]]] the violin.

The position for sentential negation is thus the one in (16), but the question arises if Neg is Merged higher than Perf, but lower than T, how can we ever generate sentences such as (20)?

(20) I have not been playing the violin.

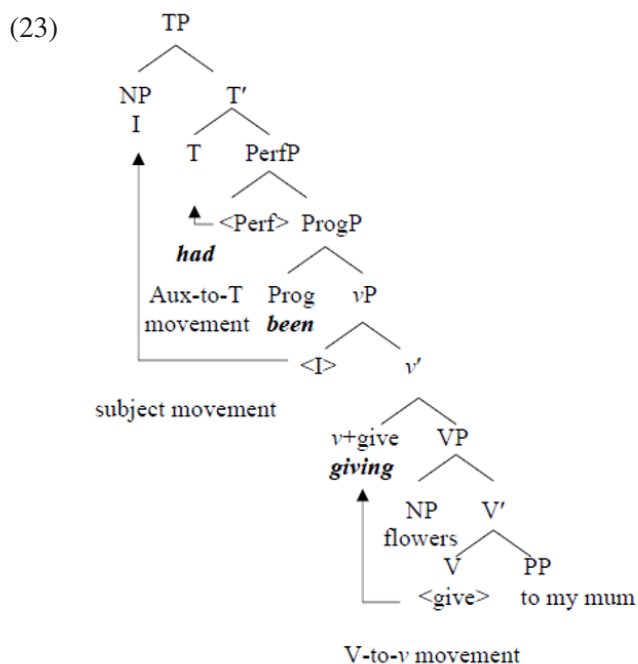
The answer is, again, by means of movement. That is, in (20) Perf moves to T across Neg, (21).

(21) I [T have [Neg not [Perf <have> [Prog been [$v + V$ playing]]]]] the violin.

It seems, though, that this only applies to Perf. In fact, it extends to the highest auxiliary in the sentence. In other words, in a sentence without Perf but with Prog, it is Prog that moves to T across Neg.

(22) I [T am [Neg not [Prog <am> [v + V playing]]]] the violin.

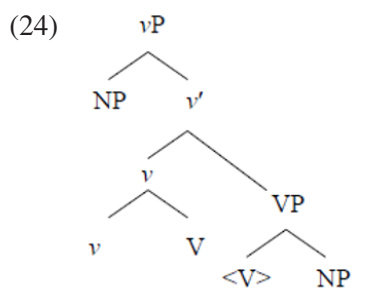
At this point, we have seen two instances of head movement (V-to-*v* and Aux-to- T) and one instance of XP-movement (subject movement), (23).



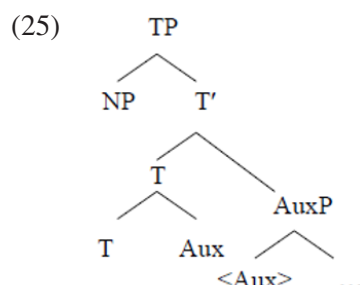
They all have in common that an element Merged in a lower position is moved to a higher position. The question is how is this need for Movement encoded in the syntax?

In the case of the Merge operation, we argued that it was driven by the uninterpretable features of a head. In the case of Movement it is argued that it is driven by the *strength of uninterpretable features*. That is, uninterpretable features can be either *weak* or *strong*. If they are weak, they can be checked either by means of sisterhood, or by means of Agree. But if they are strong, they can only be checked by means of sisterhood, and how is this to happen if they are not Merged as sisters of the checking head in the first place?

We are arguing here that Movement exists, like Merge, to solve the 'problem' of uninterpretable features. Take for instance the case of V-to-*v* movement. In a VP with two objects (e.g., the one in (23)), *v* and V are clearly not sisters. But the same happens in a VP with just one object: *v* is sister to the VP projection, but not to the V head. Little *v* needs to Merge with V. Hence, in the system we have developed, this amounts to assuming that *v* carries an uninterpretable [*u*V] feature. If, in addition, such feature is strong (which we indicate by means of an asterisk –not to be confused with ungrammaticality!), [*u*V*], then *v* can only check its uninterpretable feature by dragging V upwards and Merging under sisterhood with it. The result is a structure such as (24).



Such an analysis can be extended to Aux-to-T movement straightforwardly. The highest auxiliary (if the sentence contains more than one) carries an uninterpretable T feature with an open value, which is to be specified as [past] or not in the checking process. Such an unvalued uninterpretable feature is strong, [μ T*:] and, therefore, requires movement of the auxiliary that carries it to T for checking to occur in a structural relation of sisterhood. The resulting structure is similar to the one you can see in (24).



Note that in (25) the labels ‘Aux’ and ‘AuxP’ have been used to indicate that this applies to Perf and Prog if they happen to be the highest in a given sentence.

In the case of subject movement, English requires the subject to be overt and to occur before the inflected verb/auxiliary. Thus, it is plausible to assume that T carries an uninterpretable noun feature that is strong, [μ N*]. That is, T requires a lexical item with noun features in its domain, and that is why it drags it from the Specifier of ν P to the Specifier of TP. In Unit 5, when we examine another instance of XP-movement, you will see that unlike the case of head movement (which targets a head to merge it with another head), XP movement targets an XP from either a complement or a Specifier position and moves it to a higher Specifier position.

4.5. The Pronouncing Tense Rule, Neg(ation) and *do*-support

Note that the weak vs. strong property of uninterpretable features allows us to capture yet another observation about English. If we compare English to French, we see that adverbs cannot come between the verb and the object, (26b), whereas in French they can, (27b).

- (26) a. John often eats apples.
 b. *John eats often apples.
- (27) a. Jean souvent mange des pommes.
 b. Jean mange souvent des pommes.

In a similar vein, English allows auxiliaries to appear before negation, but not verbs, (28), whereas in French both auxiliaries and verbs can appear before the negative marker *pas*, (29).

- (28) a. John has not eaten apples.
 b. *John eats not apples.
- (29) a. Jean (n’)a pas mangé des pommes.
 b. Jean (ne) mange pas des pommes.

Assuming the syntax of English and French to be the same, we can devise the following parametric difference between these two languages in terms of the strength of features of Aux(iliaries) and little ν :

(30)

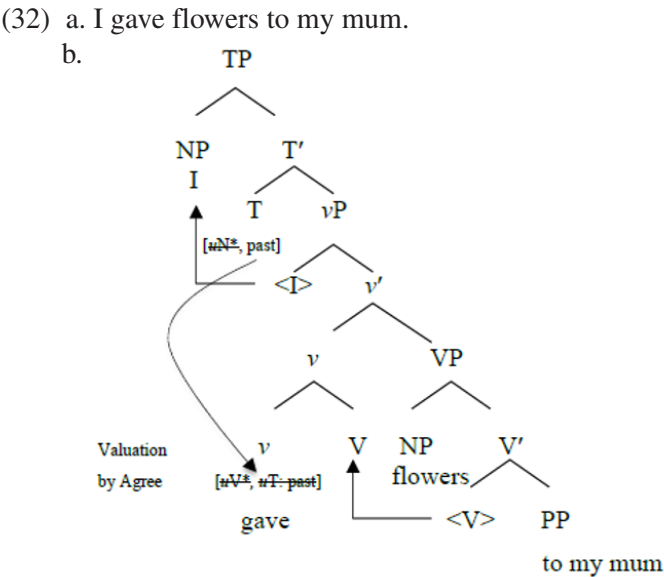
	[<i>u</i> T:] on Aux	[<i>u</i> T:] on <i>v</i>
English	Strong, [<i>u</i> T*:]	Weak, [<i>u</i> T:]
French	Strong, [<i>u</i> T*:]	Strong, [<i>u</i> T*:]

The fact that *v* carries a weak [*u*T:] feature has a very important consequence for the syntax of English negation. As we will now discuss, the impossibility of *v* moving to T to get inflection across Neg (as the [*u*T:] feature of *v* is weak), requires the use of *do*-support. Although there are different alternative accounts of *do*-support, the one we will see here relies on the structural relation of c-command, which has been seen to be relevant to the syntax of other phenomena in earlier units. In our account of *do*-support, we treat it as a Last Resort operation. That is, as an operation that rescues the syntactic object from a problem that cannot be fixed in any other most optimal way given the grammatical constraints of English.

Let us assume that English enforces a *Pronouncing Tense Rule* (PTR for short), which can be formalised as in (31).

(31) *The Pronouncing Tense Rule*
Pronounce the tense features of T on *v* only if *v* is the head of T’s sister.

As you can see, if ‘*v* is the head of T’s sister’, then T is sister of *v*P and T c-commands *v*P. Let us see how Neg disrupts such a necessary configuration for the verb (*v*/V) to be inflected for tense. Consider a sentence such as (32a), which can be analysed as in (32b).



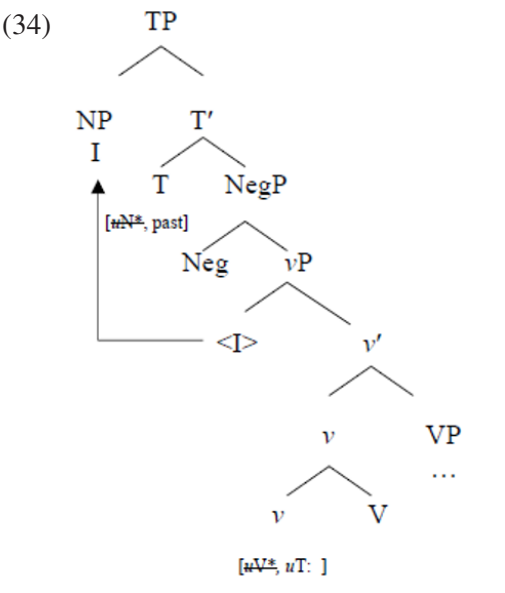
Now make it negative, as in (33):

(33) I didn’t give flowers to my mum.

As you can see, the past tense can no longer be expressed on *v*. Rather, it needs to be expressed on the auxiliary *did*.

Let us syntactically analyse the sentence in (33). Recall that Neg is Merged between T and *v*. This is the position that we argued for earlier (T > Neg > (Perf) > (Prog) > *v* > V) by looking at sentences with modals and auxiliaries. As you can see, NegP is now T’s sister, not *v*P. This means that the past tense cannot be pro-

nounced on *v* (as the PTR is not satisfied). In addition, in having a weak [*uT*:] the feature cannot Move to solve this problem. Therefore, as a Last Resort (that is, when nothing else is possible in the grammar of English), a dummy element *do* is inserted as a place-holder for the tense feature of T. As the tense feature of T is [past], *do* surfaces as *did* and the [*uT*:] on *v* is pronounced as a bare infinitive, the default value.



4.6. The syntax of intransitive verbs

One last issue that we can explore as part of this unit is the syntax of two different kinds of intransitive verbs. Recall that these are one-place predicates (as they only assign one argument). Intransitive predicates, though, can be of two different kinds, *unergative* and *unaccusative*, depending on which theta-role is assigned to their only argument. This can be the Agent theta-role (unergatives), or the Theme theta-role (unaccusatives).

As it has been suggested in the linguistic literature that there is a one-to-one correspondence between argument structure and phrase structure, before continuing with our discussion we will introduce the *Uniformity of Theta-Assignment Hypothesis* (UTAH for short), which can be formalised as in (35).

(35) *The Uniformity of Theta-Assignment Hypothesis*

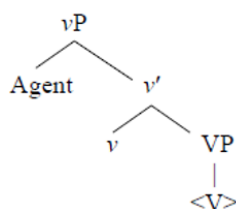
Identical thematic relationships between predicates and their arguments are represented syntactically by identical structural relationships.

In other words, what the UTAH says is that Agents and Themes correspond to fixed syntactic positions. These are the following:

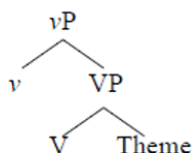
- (36) a. Agent – Specifier of vP
b. Theme – daughter of VP

So, if unergative predicates assign an Agent theta-role and unaccusative predicates assign a Theme, and these correspond to different syntactic positions, then the syntax of unergatives and unaccusatives is essentially different, as shown in (37).

(37) a. Unergative



b. Unaccusative



But how can one determine whether an intransitive predicate is unergative or unaccusative? Consider the examples in (38):

- (38) a. The cat jumped.
b. The ice melted.

At least two tests can be used to distinguish these two apart (although they must be used with caution, as they are not infallible). First, unergatives and unaccusatives correspond to different paraphrases: while unergatives can be roughly paraphrased as ‘X (the subject) is the cause of an event of V (the predicate)’, unaccusatives roughly correspond to ‘X undergoes an uncaused event of V’. Second, in assigning a Theme, unaccusatives can have a transitive equivalent where the Theme becomes obvious.

Let us see how the first of our two tests applies to the examples in (38):

(39) *Test 1: Paraphrasing*

- | | |
|--|--------|
| a. The cat is the cause of an event of jumping. | OK |
| b. The cat undergoes an uncaused event of jumping. | NOT OK |
| c. The ice is the cause of an event of melting. | NOT OK |
| d. The ice undergoes an uncaused event of melting. | OK |

Based on the first of the two tests, then, (38a) is unergative, and (38b) is unaccusative. In (40) you can further see that *melt* has a transitive counterpart where *the ice* is clearly a Theme. *Jump* also has a transitive counterpart, (41), where *the cat* is clearly an Agent.

(40) *Test 2: Transitive counterpart*

- | |
|---|
| a. The sun melted the ice. [The sun _{Agent} ; the ice _{Theme}] |
| b. The cat jumped the fence. [The cat _{Agent} ; the fence _{Theme}] |

One last question is how a Theme Merged in the VP manages to occupy the position of a subject in the Specifier of vP. The answer is straightforward: as unaccusative predicates do not have an Agent, the Specifier of vP is an empty position. Given that T continues to require a nominal element in its domain as a result of carrying a [*uN**] feature, dragging the Theme through the (empty) Specifier of vP allows the checking of [*uN**] to take place.

Notice, by contrast, that in the transitive counterpart of an unaccusative predicate (e.g., *The snow melted the ice*), there is an Agent, *the snow*, that is Merged in the Specifier of vP and later moves to the Specifier of TP. Thus, there is no motivation for the Theme *the ice* to move in (40a).

4.7. Suggested exercises and key

4.7.1. Exercises

Exercise 1.

[This exercise is adapted from Adger (2003: 196)]

As can be seen in (1), non-finite clauses are headed by *to*. This has been considered to be a T head (albeit with no tense features). Once you have considered the data, do each of the tasks.

- (1) a. John decided [not to eat apples].
- b. *John decided [to not eat apples].

Task 1. What problem does (1) raise for the Hierarchy of Projections?

Task 2. Hypothesise a possible solution.

Exercise 2.

[This exercise is adapted from Adger (2003: 197-198)]

As was seen in constituency tests, it is possible to elide the VP. One of the constraints of VP-ellipsis is that the material that is elided is recoverable from the context of the utterance. For example, (3) is an acceptable continuation of (1) but not of (2).

- (1) John's car fell over a cliff last year...
- (2) The price of petrol increased considerably last year...
- (3) and Mary's truck did [] too.

(3) is an odd continuation of (2) because the elided VP is naturally interpreted to be the same as the VP in (2) (that is, *Mary's truck increased considerably last year*). One way of accounting for this is by assuming an identity condition on VP-ellipsis, which can be formalised as in (4).

- (4) *Identity Condition on VP-ellipsis (ICVE)*

A vP can only be elided if its content can be recovered from an *identical* vP in the preceding discourse.

Assume ICVE in the tasks that follow.

Task 1. Say, for each of the following sentences, what the elided material is. Does this support the ICVE or not?

- (5) I will eat a mango, and John will [] too.
- (6) John has made pasta, and Mary has [] too.
- (7) John is making scones, and Mary is [] too.

Task 2. Now consider the following sentences. If you fill in the blanks using literal identity of form with the preceding vP, what is the result? Why is *do*-support needed? (Tip: If ellipsis applies to a constituent, what happens to this constituent? How is this relevant for the Pronouncing Tense Rule?)

- (8) I ate a mango and John did [] too.
- (9) Mary makes pesto pasta, and John does [] too.

Task 3. Consider the following sentences and explain why *do*-support is not needed.

- (10) Mary slept well, and John will [] too.
 (11) John ate mangoes and Mary should [] too.

Exercise 3.

Consider the data in (1). As you can see, the presence of *never* does not trigger *do*-support. What does this tell you about the position of this adverb in the clause?

- (1) a. Mary doesn't read international newspapers.
 b. Mary never reads international newspapers.

Exercise 4.

[This exercise is adapted from Adger (2003: 196-197)]

In this exercise, you are requested to test the nature of the copula *be* in English.

As shown in the data in (1), *be* has the function of relating a subject to a complement (an NP, an AP or a PP).

- (1) a. Mary is a student.
 b. Mary is intelligent.
 c. Mary is at the university.

Task 1. Assume that *be* is a version of little *v*, with the subject in its Specifier and the NP, AP or PP as its complement. Provide a complete analysis including T and showing all the movements.

Task 2. Now construct negative versions of the sentences in (1). What do your examples tell you about the position of *be*? Revise your analysis in Task 1 if you need to.

Task 3. Does (2) support your revision of the analysis of the data in (1)?

- (2) *Mary doesn't be at the university.

Exercise 5.

Some food for thought! Some time ago in a Facebook post, the linguist Hedde Zeijlstra observed that a phrase such as *sleep often* behaves verbally. That is, you can put it in the past and combine it with a subject, (1).

- (1) I slept often.

He also observed that verbs do not need adverbs, but that it is adverbs that rather need verbs. The puzzle is, then, why would the union of an adverb and a verb behave like a verb? And he solved it by claiming that an adverb is really a verb that needs a verb.

Task 1. How would this idea be formalised in the system that we have developed thus far?

Task 2. How would this characterisation fit into the feature specifications for the major syntactic categories that we defined in Unit 2?

4.7.2. Key to exercises

Exercise 1.

Task 1. If *to* is a T head, the Hierarchy of Projections would predict negation to occur after it, not before.

- (1) T (*to*) > Neg > *v* > V

Task 2. Given that *to* has no T features but when occurring in finite clauses T does have tense features, we could hypothesise that two different T heads exist, a T[finite] and a T[nonfinite] which occupy different positions in the Hierarchy of Projections.

- (2) T[finite] > Neg > T[non-finite] > *v* > V

Exercise 2.

Task 1.

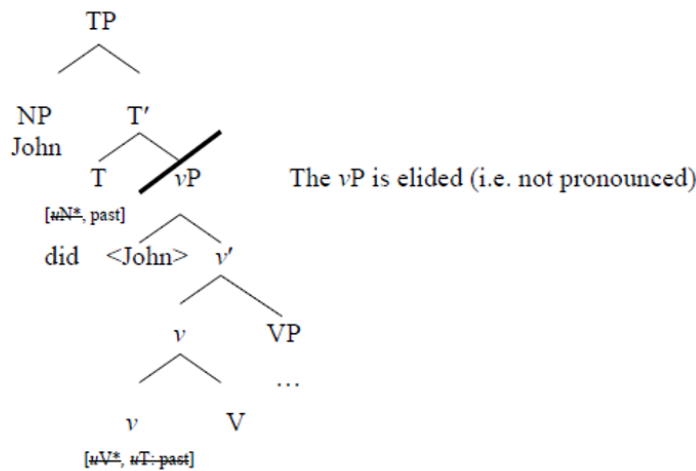
- (5) ...eat a mango
(6) ...made pasta
(7) ...making scones

The data support the ICVE, as the elided material is identical to that in previous discourse.

Task 2.

- (8) ...eat a mango
(9) ...make pesto pasta

The ellipsis of a constituent results in it not being pronounced. The tree below represents *John did* [eat a mango]. Given that the *vP* is not pronounced, even if it is T's sister, the features of T cannot be pronounced on *v* (for it is itself unpronounced). Thus, *do*-support is required.

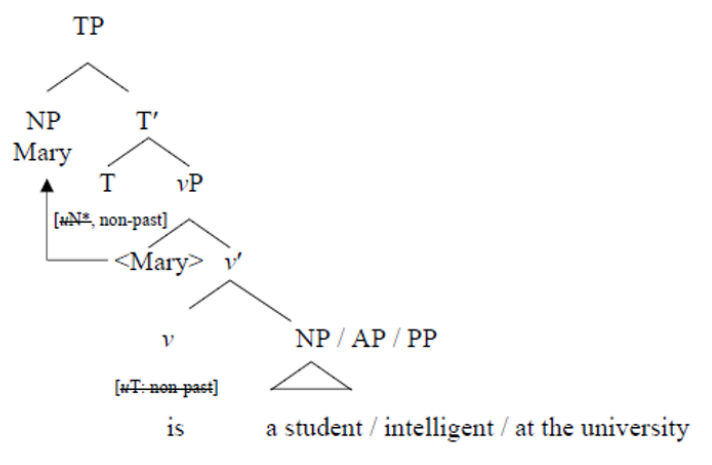


Task 3. The elided *vP* is *sleep well* in (10) and *eat mangoes* in (11). *Do*-support is not needed because the T head is occupied by the modals *will* and *should*, which express T's [non-past] and [past] features respectively.

Exercise 3.
The data in (1) indicate that *never* is in a position that does not make it impossible for T and *v*P to be sisters. *Never* is a *v*P adjunct, not a NegP like *–n’t*.

Exercise 4.

Task 1.



Task 2.

- (1) a. Mary isn’t a student.
b. Mary isn’t intelligent.
c. Mary isn’t at the university.

Do-support is not required and *is* occurs before negation. This indicates that *is* gets to T. However, it cannot get there by means of Movement if its [*u*T:] feature is weak. Hence, if we still consider it a *v* head, then it must carry a [*u*T*:] feature. Alternatively, we can argue that *be* is a T head (and that *v* encodes causality and hosts no lower V in lacking a [*u*V*] feature, as it selects a non-verbal complement).

Task 3. The data in (2) support both alternative analyses, as they show that *be* is incompatible with *do*-support. This would be the case both if it had a strong [*u*T*:] feature (thus behaving like an auxiliary), or if it were a T-head. What we know for sure is that it is not a V head.

Exercise 5.

Task 1. In the system that we have developed thus far, if an adverb is ‘a verb that needs a verb’, then it would be specified with the category feature [V] and an uninterpretable V feature, i.e. [V, *u*V]. The presence of the [*u*V] feature is what would distinguish an adverb from a verb.

Task 2. The characterisation could be completed as follows:

Nouns	[N]	or	[+N, -V]
Adjectives	[N, V]	or	[+N, +V]
Prepositions	[]	or	[-N, -V]
Verbs	[V]	or	[-N, +V]
Adverbs	[V, <i>u</i> V]	or	[-N, +V, <i>u</i> V]

4.8. Suggested reading

Edelstein (2020). Chapter 6, section 6.2. *Do-support with negation*.

This is a more sophisticated alternative to the account of *do*-support given in Unit 4, although it uses machinery that the students are already familiar with. Students should be aware that alternative accounts for the same phenomenon often co-exist within the same theoretical model.

UNIT 5. Functional categories and movement II (CP)

In this unit we introduce the functional category C(omplementiser), which adds a further clausal layer on top of TP and relates to the semantic status of the proposition (that is declarative vs. interrogative). We will explore the syntax of English *yes/no*-questions by exploring the phenomenon of Subject-Auxiliary Inversion in interrogatives.

5.1. Objectives of the unit

- Introduce and argue for the functional category *C(omplementiser)*.
- Discuss the phenomenon of *Subject-Auxiliary Inversion* in interrogatives (direct and embedded *yes/no*-questions).

5.2. Complementisers (*that, whether, if*)

The English clause can be further expanded beyond TP. This can be done by means of a functional layer that has to do with the semantic status of the whole proposition (that is, the complete set of thoughts that a sentence expresses). Such a functional layer is headed by the functional head C(omplementiser) and it is thus known as the C(omplementiser) P(hrase). Evidence for CP comes from the way clauses are embedded in English. Consider the two sentences in (1) and (2):

- (1) Mary said something.
- (2) John is handsome.

Now try to replace *something* in (1) with the clause in (2). The result is (3), which may contain the word *that*.

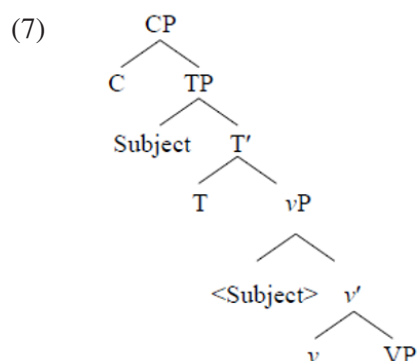
- (3) Mary says (that) John is handsome.

In English, as indicated by the parentheses, *that* is optional in (3), but note that in the Catalan and Spanish counterparts of (3) the word for *that*, 'que', is obligatory.

Constituency tests can be used to show that the string *that John is handsome* is a constituent, as shown in (4), (5) and (6).

- (4) a. It is that John is handsome that Mary says. [Transposition]
b. *It is John is handsome that Mary says that.
- (5) Mary says it. [Substitution]
- (6) Mary says [that John is handsome] and [that Peter is handsome] too. [Coordination]

That is one possible expression of C, a functional head that introduces subordinate clauses. The complement of C is the TP, with English clauses thus having the structure in (7), and the Hierarchy of Projections now being updated to (8).



- (8) C > T > (Neg) > (Perf) > (Prog) > v > V

The C head does not assign a theta-role. Its semantic contribution is related to how the hearer should think of the proposition expressed by the clause. This is reduced to two possibilities: the clause either states the truth of a given state of affairs or questions the truth of such a state of affairs. While the former kind of proposition is expressed by means of a declarative clause, the latter is usually expressed by means of a question. An example of each kind is given in (9). Note that the word order is radically different.

- (9) a. John is playing cards. Declarative clause
 b. Is John playing cards? Question

We will come back to the syntax of (9b) but let us first concentrate on analysing the syntax of embedded clauses, which is where evidence for the C head is most obvious. Consider the data in (10), where Mary is the speaker and the material in inverted commas is what she utters.

- (10) a. Mary: 'John lives here.'
 b. Mary: 'Does John live here?'

While (10a) can be paraphrased as (11a), (10b) must be paraphrased as (11b) or (11c).

- (11) a. Mary said *that* John lives here.
 b. Mary asked *whether* John lives here.
 c. Mary wondered *if* John lives here.

This means that the choice of C head (*that*, *whether* or *if*) is different depending on whether the proposition expressed by the embedded clause is a declarative clause or a question. Crucially, though, note that the word order is the same for all three sentences in (11). That is, the sentences in (11b) and (11c) display the same word order as (9a) (and not the word order in (9b)). This shows that embedded questions in English syntactically behave as declarative clauses. However, the examples in (12) show that *that* on the one hand, and *whether* and *if* on the other are in complementary distribution. That is, where *that* is grammatical, *whether/if* are not, and the other way around.

- (12) a. *Mary said *whether* John lives here.
 b. *Mary said *if* John lives here.
 c. *Mary asked *that* John lives here.
 d. *Mary wondered *that* John lives here.

It seems, therefore, that the choice of *that* vs. *whether/if* is the result of both syntactic and semantic features. In particular, we assume that these three different C heads all encode an interpretable syntactic feature [clause-type:], which is valued, in the three cases, as ‘declarative’ (that is, as [clause-type:decl]). We will leave the discussion of non-embedded *yes/no*-questions for later, but, as we will show, a sentence such as Mary’s utterance in (10b) will be argued to contain a C head with a different syntactic feature specification, namely an interpretable clause type feature valued as Q(uestion) (that is, [clause-type:Q]). This will result in the word order differences observed between embedded questions (which behave like declarative clauses) and non-embedded questions, which we assume to carry an interpretable clause-type feature valued as ‘question’ (that is, [clause-type:Q]).

In short, *that*, *whether* and *if* are all C heads specified with an interpretable [clause-type: decl]. And yet, as was seen in (11) and (12), they are in complementary distribution. We can account for this fact by showing that the actual choice of a C head in a declarative clause is also a semantic matter. Let us see how.

Recall that in the syntactic model we are developing, a lexical item Merges with another one because the former carries an uninterpretable feature that instructs it to do so. Hence, if a predicate Merges with a full clause (e.g., a CP, as in (3)), it must be because it bears an uninterpretable feature that instructs it to do so. This allows us to assume that the predicates *say*, *ask* and *wonder* not only carry an interpretable category feature [V], but also an uninterpretable c-selectional feature [*uC*], (13). Recall that c-selectional features are those that determine which category a given lexical item requires to Merge. The feature [*uC*] determines that the predicates in (13) Merge with a CP headed by a C and, crucially, with the word order of a declarative clause (see (14), which are the sentences in (11) but with a *yes/no*-question word order).

- (13) a. *say* [V, *uC*]
 b. *ask* [V, *uC*]
 c. *wonder* [V, *uC*]
 (14) a. *Mary said that does John live here.
 b. *Mary asked whether does John live here.
 c. *Mary wondered whether does John live here.

It seems, therefore, that the choice between *that* and *whether/if* is a matter of semantic selection imposed by the predicate (rather than some further syntactic specification on the C head). This idea is confirmed by the behaviour of certain predicates such as *maintain*. As can be seen in (15), *maintain* is only compatible with C *that*, but the data in (16) show that such incompatibility extends to nominal complements with question semantics.

- (15) a. Mary maintained that John is handsome.
 b. *Mary maintained whether/if John is handsome.
 (16) a. Mary maintained her opinion throughout the debate.
 b. *Mary maintained her question throughout the debate.

To summarise, predicates such as *say*, *ask*, *wonder*, *question*, *investigate*, *claim*, etc. may carry a [*uC*] feature that requires them to Merge with a clausal (i.e. CP) complement headed by a C head specified with an interpretable clause-type feature valued as declarative (i.e. [clause-type:decl]). Whether the C head is *that*, *whether* or *if* is related to the semantic selection properties of the predicate.

Finally, we saw earlier that in English the C *that* may only be optionally used (e.g., (3), where we used *that* in parentheses precisely to indicate that it may or may not be there). The question is, then, does a sentence such as (17) have a C or not?

- (17) Mary said John is handsome.

Trying to apply the transposition constituency test to the chunk *John is handsome* should give you the answer. As shown in (18), an otherwise silent *that* needs to be overt in a pseudo-cleft clause. Hence, in the

same way we assume that both *whether* and *if* introduce subordinate declarative clauses that question a given state of affairs, we will assume that both *that* and a silent C can introduce subordinate declarative clauses that state a given state of affairs.

- (18) a. What Mary said is that John is handsome.
b. *What Mary said is John is handsome.

In the next section, we address the syntax of *yes/no*-questions in English, which are different from declaratives and embedded questions.

5.3. Interrogatives: *yes/no*-questions

In this section we address the syntax of *yes/no*-questions. As clearly indicated by the terminology used, *yes/no*-questions are those that can be answered with a *yes*-answer or a *no*-answer. This is the case, for instance, in (19a).

- (19) a. Has John called?
b. John has called.

If the word order of (19a) is compared to its declarative counterpart in (19b), what can be observed?

Your answer to the previous question should have been that in (19a) the auxiliary *has* precedes the subject *John*, while in (19b) the subject *John* precedes the auxiliary *has*. Continuing with self-explanatory terminology, the phenomenon that we observe in (19a) in connection to the syntax of questions is known as *Subject-Auxiliary Inversion* (SAI for short). As you can see in (20), this operation applies to any auxiliary in the clause, and, as shown in (21), to the highest in the syntactic structure when there is more than one auxiliary. Note, as well, that modals, which we considered T elements in earlier units, also display the same syntax.

- | | |
|--|-----------------|
| (20) a. Is/Was Mary singing? | Prog |
| b. Has/Had John played cards? | Perf |
| (21) Has/Had Mary been singing? | Perf > Prog |
| (22) a. Can Mary sing? | T |
| b. Could John be playing cards? | T > Prog |
| c. Could Mary have sung? | T > Perf |
| d. Could John have been playing cards? | T > Perf > Prog |

Adding the CP-layer to English clauses allows us to maintain the syntax of clauses that we have previously developed while accommodating the data in (20)-(22). Recall that in the previous unit we concluded that subjects sit in the Specifier of TP (after having moved from the Specifier of vP to check the [*uN**] feature of T), and that the highest auxiliary in the clause moves from its dedicated functional head (Perf or Prog) to T due to its strong [*uT**:] feature. Where does such an auxiliary Move (further) in *yes/no*-questions to end up preceding the subject (which sits in the Specifier of TP)?

You may have guessed that as the sentences in (20)-(22) are not embedded but main clauses, they do not contain a *that* C. Hence, if the CP-layer is part of every clause (which makes sense, as every clause is, potentially, embeddable in a superordinate one) but the C head is not occupied by *that* in questions, then the empty slot can serve as the landing site for the Moved auxiliary. The diagram in (23) schematically represents this idea.

As usual, the question that comes next is what triggers the movement of the auxiliary from T to C. And, as usual by now, the answer is that it is the strength of a feature. But the strength of which feature?

We argued earlier that C carries an interpretable [clause-type] feature that is valued as [clause-type:decl] in declarative clauses (and embedded questions). For non-embedded questions, let us assume that such a

feature can be valued as ‘question’ (i.e. [clause-type:Q]). Let us further assume, as well, that T carries the uninterpretable counterpart of this feature (i.e. [*u*clause-type:]), which triggers Merge of TP and C, and that when such a feature is valued as Q, it is strong (i.e. [*u*clause-type: Q*]). The strength of such a feature then forces movement of T-to-C. By contrast, if the [*u*clause-type] feature is [*u*clause-type:decl], no movement of T-to-C occurs.

If T contains an auxiliary (e.g., a modal, Perf, or Prog), then SAI takes place. If T only contains a tense feature (as in (24)), T-to-C movement requires *do*-support, thus also resulting in the subject and an auxiliary (a form of *do*) inverting.

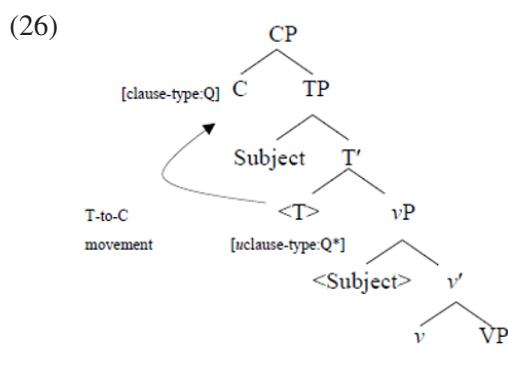
- (24) a. I bought biscuits.
b. Did I buy biscuits?

Notice that the reason why movement of T to C requires *do*-support follows straightforwardly from the Pronouncing Tense Rule that we discussed in Unit 4, repeated here for convenience.

(25) *The Pronouncing Tense Rule*

Pronounce the tense features of T on *v* only if *v* is the head of T’s sister.

If T is displaced to incorporate into the head C (i.e. T-to-C is an instance of head movement), T and *v*P can no longer be sisters and, therefore, the tense feature of T cannot be pronounced on *v*. This is schematically shown in (26).



5.4. Suggested exercises and key

5.4.1. Exercises

Exercise 1.

[This exercise has been adapted from Adger (2003: 335)]

Assuming the analysis of SAI given in the unit consider the data in each task and answer the questions.

Task 1. Explain what problems arise for this analysis given the grammaticality of examples such as (1) and (2).

- (1) Hasn’t Mary played cards?
(2) Doesn’t John watch TV?

Task 2. Hypothesise a solution for this problem by concentrating on the different behaviour of the two expressions of negation *–n’t* and *not*. You might find the data in (3) and (4) useful.

- (3) Has Mary not played cards?
- (4) *Has not Mary played cards?

Exercise 2.

[This exercise is adapted from Adger (2003: 335)]

In some Irish varieties of English, we find so-called *embedded inversion*. These are structures where SAI occurs in embedded questions.

- (1) I asked did John live there.
- (2) I wondered had Mary left yet.

What do you think is going on here? What is the difference between Standard English and Irish varieties of English when it comes to the syntax of embedded questions?

Exercise 3.

[This exercise is adapted from Adger (2003: 372)]

The sentence in (1) involves movement of a negative phrase.

- (1) Never will they forget these lectures.

Compare the example in (1) to (2) and (3), and then hypothesise a possible analysis for (1) and enumerate what theoretical questions it raises.

- (2) They will never forget these lectures.
- (3) What will they never forget?

5.4.2. Key to exercises

Exercise 1.

Task 1. According to the Hierarchy of Projections, Negation is sandwiched between TP and PerfP ($C > T > \text{Neg} > \text{Perf} > v > V$). As Perf is the highest auxiliary in (1), it moves to T to check its [μT^* :] feature. As (1) is a *yes/no*-question, T moves further to C to check its [μQ^*] feature. Neg has not been argued to undergo any movement, so the predicted word order for a question such as (1) would be:

- (i) Has Mary not played cards?

The sentence in (2) shows a similar problem: T (realised as *does*) moves to C, so the predicted word order is (ii):

- (ii) Does John not watch TV?

It seems, therefore, that Neg moves with T to C.

Task 2. We can assume that Neg undergoes head movement to T before T moves to C. When it does, Neg is affixed to T (as in (1) and (2)). If Neg does not undergo head movement to T before T-to-C, we get (3). The example in (4) shows that Neg can only move to C if it previously head-adjoins to T.

Exercise 2.

These sentences contain embedded questions, which in Standard English have the word order of a declarative clause. In Irish varieties of English, though, embedded questions display the word order of direct questions.

The syntactic difference can be hypothesised to be that in (1) and (2) C is specified as [clause-type:Q] rather than as [clause-type:decl]. This triggers SAI in embedded questions.

Exercise 3.

If (1) is compared to (2) and (3), we can see that both (1) and (3) involve SAI.

Our analysis of (1), then, would possibly involve the negative phrase moving from a *v*P-adjunct position to the Specifier of CP. The question is, as usual, what feature is triggering such movement.

5.5. Suggested reading

Bruening (2017). *Subject-Auxiliary Inversion*.

In this reading students discover that Subject-Auxiliary Inversion may have to be assumed to come in two distinct types. In addition, they can also learn about alternative accounts of SAI that involve no movement at all. Throughout the article, the focus is on the problem posed by the possibility of inverting multiple auxiliaries.

6. Bibliography

- ADGER, David. 2003. *Core Syntax. A Minimalist Approach*. Oxford: Oxford University Press.
- BRUENING, Benjamin. 2017. Subject-Auxiliary Inversion. In Martin Everaert and Henk van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax*, second edition, volume VII, 4250-4277, Somerset, NJ: John Wiley and Sons.
- CHOMSKY, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- 1986. *Knowledge of Language*. New York: Praeger Publications.
- 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- EDELSTEIN, Elspeth. 2020. *English Syntax: A Minimalist Account of Structure and Variation*. Edinburgh: Edinburgh University Press.
- KOENEMAN, Olaf and Hedde ZEIJLSTRA. 2017. *Introducing Syntax*. Cambridge: Cambridge University Press.
- LLINÀS-GRAU, Mireia and Mercè COLL-ALFONSO. 2001. Telic verbs in early Catalan. *Probus* 13(1): 69-79.
<https://doi.org/10.1515/prbs.13.1.69>
- RADFORD, Andrew. 2004. *English Syntax: An Introduction*. Cambridge University Press.



SUSAGNA TUBAU (Manlleu, 1978) holds a PhD from the Autonomous University of Barcelona and the University of Amsterdam since 2008 and is Associate Professor at the Department of English Philology and German Studies of the Autonomous University of Barcelona since 2023, where she teaches courses related to English syntax and linguistic variation. Her research on the syntax and semantics of phenomena related to negation in natural languages has been published in the form of articles and book chapters in various journals and renowned publishing houses and has been debated with the international linguistics community in numerous national and international conferences and seminars. Tubau is currently leading a research project on the syntax and semantics of expletivity and expressive meaning.

The goal of this manual is to introduce students to the syntactic analysis of English by offering a comprehensive discussion of the core concepts and theoretical assumptions that underlie minimalist syntax. In this book, syntactic operations such as Merge and Agree are systematically linked to the morphosyntactic features that are relevant for English. In addition, they are shown to result in the phrase structure predicted by constituency tests. By exploring the properties and syntactic behaviour of functional categories such as little *v*, Tense and Complementiser, the manual offers students analytical tools to understand and discuss the grammar underlying basic English sentences. Each unit contains exercises with a key to practise and widen the content and skills covered.