

CHAPTER 2

Economy

In the previous chapter, I have briefly mentioned two Economy Principles that will be used in the remainder of this book as mechanisms of grammaticalization. One is the 'Spec to Head Principle', and the other is the 'Late Merge Principle'. Spec to Head has not been discussed much in language change. It provides an explanation of the origins of complementizers and relative markers. Late Merge is mainly concerned with why heads go to higher positions in the tree (e.g. $V > I$), when they grammaticalize. This kind of grammaticalization has frequently been examined (Haspelmath 1989; van Gelderen 1993; Jbema 2002; Roberts & Roussou 2003), but not from the Late Merge point of view.

As for other Economy Principles, Chomsky (1995: 265) argues that movement carries along as much as is absolutely necessary. In English, the most well-known case of this is preposition stranding. There is tremendous prescriptive pressure 'not to end a sentence in a preposition',¹ and therefore, pied-piping, as in (1), is found in formal writing, rather than the expected stranding, as in (2):

- (1) About who(m) were you talking?
- (2) Who were you talking about?

Variation between (1) and (2) is noticed (but not condemned) by early grammarians such as Bullokar in his 1586 grammar. Two centuries later, only (1) is allowed (e.g. Coar 1796).² Preposition stranding is a good example of the tension between internal and external forces. Chomsky does not go into the question of head versus phrasal movement but the conclusion seems obvious. In cases where it is possible to strand, stranding is obligatory. Thus, it is cheaper to move a pronoun as a head than as a phrase, and checking will proceed through Head to Head rather than through Spec to Head.

The outline is as follows. In Section 1, three cases of Spec to Head are discussed, namely, pronouns, negative markers, and demonstrative/possessive elements. In Section 2, Late Merge is outlined with examples of prepositions becoming complementizers.

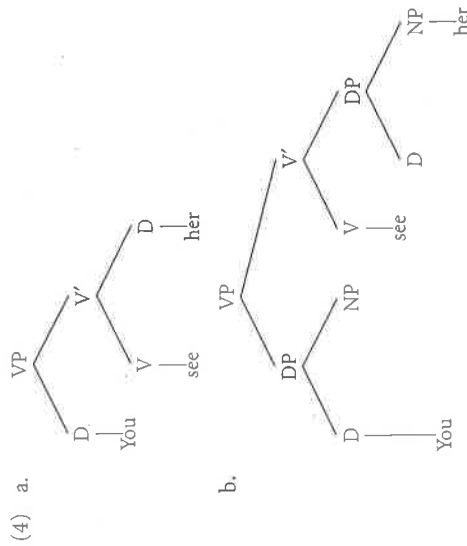
1. Spec to Head

The Principle formulated in the previous chapter is repeated here as (3):

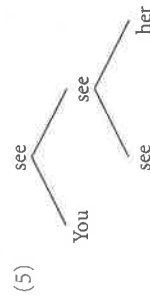
- (3) *Spec to Head or Head Preference Principle*
Be a Head, rather than a Phrase.

Stated as in (3), the principle holds for merge (projection) as well as move (checking). I will not in this book be concerned with objects, and I have therefore formulated (3) more narrowly, namely, in terms of a Spec rather than a phrase, but in principle this should hold for the merging of objects as well. I sometimes refer to the Principle in (3) as 'Spec to Head' and occasionally as the 'Head Preference Principle' and 'Head-over-Spec'.

Principle (3) predicts that structures will be built as in (4a) rather than as in (4b):



In fact, in Chomsky (1995: 246), it is suggested that labels, such as N, V, NP and VP, are superfluous and that lexical items, such as *see*, need not be labelled as heads, as in (4), but that instead we have (5):



Both (4a) and (5) are in accordance with (3). One of the possible problems is the Head Movement Constraint: if *her* in (5) needed to move, it would have to cross another head position (not allowed according to Rizzi's 1990 Relativized Minimality). This problem is a broad one and I will not attempt a solution here. Clitics in a non-bare phrase structure face the same problem, so whatever works for them will work for (5) as well.

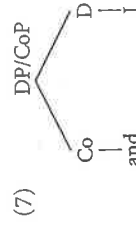
To provide evidence that a principle such as (3) is indeed operative, I'll examine pronouns in ModE, negation, and D-words across time.

1.1 Pronouns

I will argue that, unlike pronouns, nouns do not have the option to be constructed as the pronoun in (4a) above. Nouns, if they are to be argumental, have to have a D(eterminer). Pure NPs occur as predicates, vocatives, and adverbials, but not as subjects or objects, see e.g. Higginbotham (1985), Rothstein (1983), and Longobardi (1994). This is the reason why most researchers assume a D even in languages that do not have an overt one. Comparing pronouns and nouns gives (6ab):



When pronouns are coordinated or modified, they lose the ability to be heads. A possible structure for a coordinate pronoun is as in (7). This is the stage in the derivation when the pronoun has just been combined with the coordinator *and*, and with the pronoun or the coordinator can project, i.e. the result can be a Coordinator Phrase (as in Munn 1992; van Gelderen 1997a) or a DP, but never just a head:



The phrase marker in (7) will combine with a D or DP but this won't make any difference for the final result which will be a phrase. This structural difference

between nouns and pronouns has implications for Case checking. (Uncoordinated) pronouns can check via head to head, whereas nouns will check via Spec to head.

In contrast to (6ab), Cardinaletti & Starke (1995:36) follow an older literature and analyze pronouns as being of three kinds: clitics are 'deficient heads', weak pronouns are 'deficient XPs', and strong pronouns are 'non-deficient XPs' (XPs being full phrases). In their discussion of, for instance, French, they argue that "the strong variant can be used only if the deficient variant is not accessible" (p. 33 bold type omitted), e.g. if an adverb separates it from a verb or when coordinated. The weak pronoun "remains an XP on the surface ..., while ... resisting coordination or modification" (p. 36). Being an XP while resisting modification seems incompatible and I will therefore reformulate Cardinaletti & Starke's three-fold distinction as a two-fold one: elements are either X or XP, but whenever possible X.

The first piece of evidence that pronouns typically function as heads, i.e. incorporate, involves modification. For instance, they are not preceded by adjectives or modified by (restrictive) relative clauses. The reason is that speakers prefer to construct them as heads. So, *nice they* is not possible,³ and relative clauses occur very infrequently.⁴ From the Corpus of Spoken Professional American English (hence CSE), I have selected the 161,000-word Faculty Meetings' section, but the other parts are very similar. These transcripts show a split between nouns and pronouns. Thus, nouns such as *faculty, departments* and *school(s)* are coordinated over 10% of the cases, e.g. *faculty* occurs 353 times and is coordinated with *and*,⁵ as in (8), 62 times, i.e. 17.6%:

- (8) to try to tap into what students and faculty have an interest in doing.
(CSE-FAC97)

Faculty occurs much more often than 353 but I have disregarded the modifying uses of *faculty*, as in (9):

- (9) but for the grieving faculty member who feels that he or she was dismissed ... (CSE-FAC95)

For the noun *school(s)* in the same part of the CSE, the percentage coordinated is 16.4, namely 55 instances of *school(s)* with nine coordinated. In the same part, *students* are coordinated 51 times out of 367, which is 13.9%. *Departments* occurs 52 times of which 12 are coordinated, i.e. 23%. This use is very different from that of pronouns, as Table 2.1 shows. Thus, first person singular pronouns are coordinated less than 1% of the time.

Table 2.1. First and third person singular pronouns versus nouns in CSE-FAC, X-square is 418.061 $p < .001$ for first persons against nouns and 33.340 $p < 0.001$ for third persons against nouns

	uncoordinated	coordinated	total
<i>I</i> and <i>me</i>	3024	20 (= .66%)	3044
<i>s/he</i> and <i>him/her</i>	227	4 (= 1.73%)	231
<i>faculty, students, school(s), and departments</i>	693	134 (=16.2%)	827

Personal pronouns can be 'pushed into' being full phrases through coordination, as in (10).⁶ Then, the Case is often a default Case such as the accusative in (10):

- (10) Huck and me ain't cry-babies. (Mark Twain, *Tom Sawyer*, Chapter 16)

This is true in most colloquial usage, as (10) represents (through an author's filter), and (13) below, and provides more evidence that checking in (10) is different from that with non-coordinated pronouns.

Other, less direct, evidence that pronouns are heads is found in varieties of English, e.g. North East Yorkshire (Cowling 1915, cited in Chapman 1995) where agreement on the verb disappears if the pronoun is adjacent, indicating that the pronoun has 'become the agreement', and that means it is a head in T(ense) or I(nflection). In these dialects, there would be a choice of either a verbal ending or an adjacent pronoun. Chapman (1995) divides English dialects of the British Isles into three groups: those in southwest England and East Anglia where no inflection is left; those in southeast England where the -s ending is used throughout the paradigm; and those of the 'North Country', where "the use of -s as a marker of all persons singular and plural in the present is determined by the type of subject with which the verb agrees" (Chapman 1995:36). If a subject pronoun is immediately adjacent to the verb, as in (11), there is no -s; if it isn't, the -s occurs, as in (12):

- (11) I tell him not to.
(12) I often tells him.
(both from Chapman 1995)

This suggests that the pronoun is the agreement.⁷ The variety also shows that pronouns need not incorporate if, for instance, they are coordinated. Therefore, the agreement in (13) is as it would be if the pronoun was non-adjacent:

- (13) him and me drinks nought but water.
(also from Chapman 1995)

So, pronouns are heads in (11), but phrases in (12) and (13). A related tendency in this variety is that only the pronouns that are adjacent to the verb get nominative Case, as in (11), indicating that the nominative is for heads that check via adjunction to the verbal head. This tendency is not absolute, as (12) shows. Wolfram et al. (1999:70) demonstrate how these same tendencies are reflected in American dialects. For instance, agreement tends to be lacking with pronominal subjects.

Evidence from code switching that pronouns are heads involves subjects and finite verbs. In many languages (cf. Jake 1994; Nortier 1990), a switch between a full DP subject and a verb, as in (14) between English and Dutch, results in an acceptable switch, but in the case of a pronoun and verb, as in (15) again between English and Dutch, it does not:

- (14) Those awful neighbors *schijnen altijd herrie te moeten maken*⁸
those awful neighbors seem always noise to have to make
'Those awful neighbors seem to always have to make noise.'
(15) **They schijnen altijd herrie te moeten maken*
they seem always noise to have to make
'They seem to always have to make noise.'

This can be explained, as in e.g. MacSwan (1999), in terms of a prohibition against language mixing at the word level in the case of a pronoun and a verb. In the languages where pronouns and verbs can code switch, the pronouns are emphatic (see e.g. Jake 1994), i.e. phrasal. Examples of lexical (or emphatic) pronouns that can switch between French and Moroccan Arabic are French *moi* 'me' in (16) and Moroccan Arabic *nta* 'you' in (17):

- (16) *moi dixit*
I went-in-1S
'I went in.'
(17) *nta tu vas travailler*
you you go work
'You go to work' (from Bentahila and Davies 1983:313)

In (17), the non-lexical French pronoun *tu* is included as well, reinforcing the idea that the pronoun *nta* is really lexical and is like a full DP.

The same occurs in Dutch/Moroccan Arabic Code Switching, as Nortier (1990) has shown using naturalistic data. Code Switching occurs between the emphatic subject pronouns and a Dutch verb, as in (18) and (19):

- (18) *humaya vergeijken de mentaliteit met de Islam*
they compare the mentality with the Islam
'They compare the mentality with Islam.'
(19) *ana ik vind 't zo'n knuffeldiertje*
I I find it such a cuddly-little-animal
'As for me, I think it is so cuddly' (from Nortier 1991:165-6)

In (19), as in (17) above, the 'real' pronoun is repeated in Dutch. The reverse switch never occurs, i.e. a Dutch pronoun with a Moroccan Arabic verb, since Dutch does not have emphatic pronouns (in the framework of van Gelderen 2000, pronouns would be checking uninterpretable features whereas DPs and emphatic pronouns have interpretable features).

Another piece of evidence that pronouns incorporate (i.e. move as heads) is that when pronouns are separated from the verb by an adverb as in (20) or interjection, as in (21), they are often repeated. This is reminiscent of subject doubling in e.g. Northern Italian and colloquial French:

- (20) they apparently — they're involved in this. (CSE-WH96B)
(21) We find they, you know — they both work comfortably, to be honest with you. (CSE-COMR6B97)

In coordination of verbs, as in (22), the pronoun is often doubled as well, unless the two verbs form a semantic unit, as in (23a), or a special subordinate, as in (23b):

- (22) a. but I do. And I bet Judith does. (CSE-COMR797).
b. Yes I do, and I am absolutely convinced ... (CSE-FACMT97)
(23) a. And I hope and expect that we could continue to improve...
(CSE-WH94)
b. when I try and work with elementary teachers and try to map stories ... (CSE-COMR6A97)

Without formulating a theory of ellipsis, it is hard to arrive at relative numbers for the two constructions. I will leave that for further research. It is my impression that first person doubles more often than third person, however.

If pronouns can choose to be heads or full phrases, depending on the syntactic environment, one would expect languages with different phonological

forms for these. Even though English has no special phrasal forms, there is a form which is only a head, namely *ye*, the reduced form of *you* where the vowel is a schwa. In English, *ye* is a clitic head⁹ and (24) with the vowel reduced to schwa is grammatical but the coordinated (25a) is not. Instead, the full form is required, as in (25b):

- (24) *Ye can't go there.*
 (25) a. **Mary and ye can't go there.*
 b. *Mary and you can't go there.*

In Swedish, Josefsson (2000:738) argues that certain pronouns occur in the specifier of DP, as in *Han den gamle vaktmästaren* 'he the old janitor', even though regular Swedish pronouns are heads. This means there are two pronominal forms in Swedish, heads and specifiers. Languages such as Arabic and French have special phrasal (emphatic) forms and special head (clitic) forms, see (16) and (17) above. The French phrasal first and second person forms are *moi* 'me' and *toi* 'you'; the heads are *je* 'I' and *tu* 'you', as will become clear below. In French and Arabic, to name but a few, the argument can be made that the head form is an agreement marking. Lambrecht (1981) makes a strong case for Non Standard French as a pronominal argument language (to borrow Jelinek's term, or PAL), since (26a) is ungrammatical and needs to have the subject clitic repeated, as in (26b). Even in Standard French, (26a) is marginal and (26b) is preferred:

- (26) a. **Je lis et écris*
 'I read and write.'
 b. *Je lis et j'écris*
 I read and I-write
 'I read and write.'

So, (26) in French shows that pronouns are further on their way to becoming agreement than the pronouns in (22) and (23) are in English.

A last argument is that if pronouns check as heads, they might be more often cliticized with an auxiliary. This is indeed the case as Table 2.2 shows for cliticization of *am*, *will*, *would*, *has* etc. It occurs quite frequently with first person pronouns, as in (27), but it never occurs with *faculty*, *student*, *department*, and *school*.

- (27) I'm concerned that this perception came across. (CSE-FAC95)

Table 2.2. Cliticization to pronouns in CSE-FAC (significant between first and second and first and third at $p < .001$)

	uncliticized	cliticized	total
<i>I</i>	2037	685 (=25.%)	2722
<i>you</i>	1176	162 (=12.1%)	1338
<i>he</i>	128	19 (=12.9%)	147

Note that, in Table 2.2, the first person is more often a head. I won't go into that further here, but this observation corresponds to similar differences in coordination, mentioned in Chapter 13.

If pronouns check features in a Head-Head configuration, they can start to be seen as agreement. Standard English is not that far (yet), but further along perhaps with first than with third person. Pierce's (1994) data comparing French and English child language shows that French subject pronouns, in accordance with Lambrecht's (1981) observations, are much further along to being agreement markers than English ones. French children only use pronouns with finite verbs (which in the singular have the same null-form) but DPs occur with non-finite verbs. Coordinated pronouns are not used and non-emphatic pronouns are also repeated in (28), showing they are needed for proper checking:

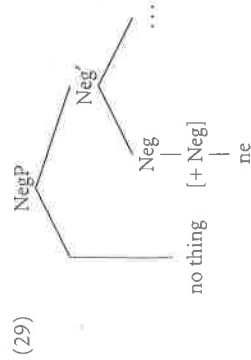
- (28) *Moi, je sautes et je descends*
 me I jump and I go-down
 'As for me, I jump and go down.' (Pierce 1994:329)

English children, in contrast, use pronouns with infinitives as well, coordinate pronouns, and need not repeat them in sentences such as (28).¹⁰

The view that agreement represents an argument, as Jelinek (1984) and Willie (1991) have argued for e.g. Navajo, is similar to the view that the agreement on the verb in Spanish and Italian licenses the (subject) argument (e.g. Safir 1985). Recent work by Taraldsen (1992) and Ordóñez & Trevino (1999) argues that the (subject) DP in languages such as Spanish is an adjunct, not an argument. However, as Jelinek (2001) makes clear, all nominals are adjuncts in pronominal argument languages, and that is certainly not the case in (English and) French. In the latter, only definite nominals can be adjuncts. So, even though English pronouns may be developing towards agreement markers, they are not at that point yet. They do, however, act like heads rather than full phrases.

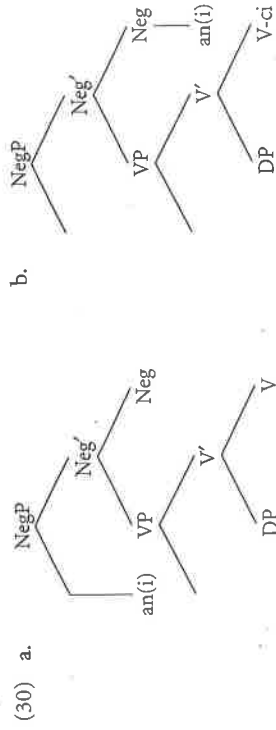
1.2 Negation

Jespersen (1917) describes what has come to be known as Jespersen's Cycle where historically the negative element attached to the verb disappears and the other negative element takes over. Latin *non dico* goes to *jeo ne di* to *je ne dis pas* and *je dis pas*. These changes are described as cases of specifiers becoming heads. Using (3), Jespersen's Cycle can be accounted for straightforwardly. See van Gelderen (2004) for more details. Assume negation is checked in a NegP, as in (29). The change then involves grammaticalization whereby a head *ne* becomes a clitic *n-* (prefixed to a verb). This change is followed by the introduction of a new XP, such as *no thing*, in specifier position:



Subsequently, the XP in the specifier becomes a head *not* and the above changes could start over again. What stops the cycle in English is the prescriptive rule (perhaps the most prevalent of all prescriptive rules) against multiple negation, since for the specifier to be filled again, there would have to be a second negative. For some of the admonitions against the use of multiple negation, see Kirszner & Mandell (1992:A37); Quirk & Greenbaum (1973:186); Swan (1980:182). These rules have existed at least since the 18th century, and are very strong. Nevertheless, speakers continue to produce violations. For instance, Anderwald (2002:104 ff.) examines the BNC and finds considerable numbers of multiple negation in spoken British English (namely 14.3% of negatives).

Other languages have a choice between placing negation in the head or the specifier. For instance, a typologically very different language, Korean, has short and long negation. The literature doesn't talk about this variation in terms of Spec-Head variation (see e.g. Park 1992 who considers the preverbal negative adverbial), but it is easily put in that framework. Both kinds involve the negative element *an(i)*, but short negation involves *an(i)* in preverbal Spec position, as in (30a), and long negation involves *an(i)* in postverbal head position, as in (30b):



In the case of (30a), both object and subject move out of the VP (if both are there), and the verb moves to T and C (not shown here) to check other features. This results in (31a). In (30b), the V is blocked by the head *an(i)* from moving to T and C, and an auxiliary/light verb *-h-* is used. The result is (31b):

- (31) a. *apenim-un an ka-sy-e*
 father-NOM not go-SH-INT
 'Father is not going.'
 b. *apenim-un ka-ci an-h-usy-e*
 father-NOM go-INF not-do-SH-INT
 'Father is not going.' (Sohn 1999:390, glosses adapted)

Under (3), (30b) and (31b) should be more economical. The additional insertion of the light V may make it less economical, however. Korean speakers (Min Young Kim p.c.) find (31b) formal and awkward in speech; Sohn (1999:390) says "the former being slightly less formal than the latter".¹¹

1.3 D-words

As argued in Wood (2003), but see also Lyons (1999), there are a number of instances where elements in the DP change from Spec to Head. Wood shows that, as the demonstrative becomes an article in OE, a shift from specifier to head occurs. An additional shift of the possessive from ME specifier to ModE head also occurs. Two orders predominate in the OE DP: a demonstrative followed by a possessive followed by a noun, as in (32), and possessive followed by a demonstrative and an adjective and a noun, as in (33):

- (32) *sio hira lufu*
 that their love
 'their love' (Pastoral Care 147.10, from Wood 2003:115)

- (33) *Mijn se leofesta freond*
my that dearest friend
'my dearest friend' (Apollo 14.9.14, from Wood 2003: 113)

(32) is less common and Wood argues, basing herself on Mitchell and Traugott, that the construction is appositive. In (33), the possessive is in the specifier, and the demonstrative has become the head. Later on, in ModE, both the (now) article and possessive pronoun are in the head position and hence (33) is no longer grammatical, even though (32) is, with appositional structure.

The numeral *one* undergoes a similar change, from Spec to Head, according to Wood (2003), and so do heads like *such* since they have their origin in the combination of a Spec *so* and a head *lik*.

Concluding Section 1, there is quite some evidence that the Spec to Head Principle is relevant to merge and checking. In Chapter 4, I will show how it is relevant to syntactic change, in particular in cross-clausal grammaticalization.

2. Late Merge

Within recent Minimalism, there is a second economy principle, namely (34), see Chomsky (1995: 316; 378) and Chapter 1 above:

- (34) *Late Merge*
Merge as late as possible.

The intuition behind this principle is that fully lexical words such as Ns and Vs are merged first since they are relevant to theta-marking (Chomsky 1995: 314–5). Grammatical words such as auxiliaries and prepositions are 'needed' later in the derivation and therefore either moved there or merged late. Child language as characterized by the Root Infinitive Stage (see Rizzi 1994) typically uses only lexical categories since functional categories are not yet part of the child's (working) lexicon. So, the child uses early merge without violating Late Merge.

Using (34), I will argue that if, for instance, a verb is less relevant to the argument structure, it will tend to merge higher (e.g. IP or CP) rather than merge early (e.g. in VP) and then move. This can happen when a V merges with another V. Since all the head to head changes are to higher positions, this principle can be considered the driving force behind this form of grammaticalization. In work on grammaticalization and reanalysis, there has been a lot of emphasis on the change from lexical to auxiliary verbs (e.g. *have* and *will*), from

prepositions to complementizers (e.g. *for* and *like*), and from verbs to complementizers. As a lexical head is reanalyzed as a grammatical (or functional) head, it 'climbs' up the tree, as it were (see also Ijbenma 2002). Modals (Traugott 1972, Lightfoot 1979), perfect marking, progressive (Heine & Reh 1984), infinitive markers and complementizers (van Gelderen 1993; 1998) fit into this, and I will come back to these in detail in Chapters 6 and 11. In Chapter 6, which deals specifically with Late Merge as it applies to clausal grammaticalization, I also show that there are cases of a Spec of a lexical category changing to the Spec of a higher functional category. This too is a result of Late Merge.

Chinese aspect markers such as *le* have grammaticalized from verbs. In Chapters 8 and 9, I will argue the same pattern can be seen in English modals and perception verbs. The perfective marker *le* in Chinese has been discussed a lot (e.g. Li & Thompson 1981; Bisang 1992; Sybesma 1999). It derives from two different verbs: *liao* meaning 'to complete' among other meanings, and *lai* 'to come' (Sun 1996: 85; 178; Shi 2002). These are typically elements in the light verb position that come to be generated higher. Both Wu (2000) and Lord (1993; Chapter 8) provide instances of serial verbs becoming auxiliaries.

Chomsky's (1995; 2001a,b) preference of 'merge over move', reformulatable as 'Merge late so that you don't have to merge as well as move', as in (34) above, provides a ready explanation. At the time of the change of main verbs to auxiliaries, verbs move to I and/or C to check features of tense and agreement, and having the auxiliary merge rather than merge and move in e.g. (35) is more economical:

- (35) *ne mahte he wið leasse gref habben arud us?*
not could he with less grief have saved us
'Couldn't he have saved us with less pain?' (Ancrone Wisse, 106r)

Roberts & Roussou (2003) use an insight similar to Late Merge to explain grammaticalization affecting C, T, and D. They couch it in a theory of features and markedness. Following Borer (1984) and Chomsky (1995), they assume all variation between languages is due to the lexicon. Functional categories can be specified as to whether or not they have a PF representation (F* for Roberts & Roussou 2003: 29) which can be achieved either by Merge or by Move. As an example, in English *wh*-questions, Q* is specified as MOVE, since the *wh*-word has to move, but as MERGE in languages with a question particle.¹² The reason for historical changes is through a 'simplicity metric' (Roberts & Roussou 2003: 201): a moved structure will not be as simple (not their words) and the child learning the language will reanalyze it. I will use a less formal approach,

i.e. without using features, and will provide a few examples here of how Late Merge accounts for change.

As is well-known, *for*, the preposition of location and purpose, as in OE (36) to (38), comes to be used as a complementizer by early ME, as in (39) and (40):

- (36) *þæt he for eaxlum gestod*
that he before shoulders stepped
'that he stood in front of ...' (Beowulf 358)
- (37) *for weryfhtum ... ond for arstafum usic sohtest*
for fighting and for support (you) us sought
'You wanted us to help fight.' (Beowulf 457–8)
- (38) *forþan ic hine sweorde swebban nelle*
therefore I him sword-DAT kill not-want
'Therefore I don't want to kill him by sword.' (Beowulf 679)
- (39) *Loctrin 7 Camber to þon scipen comen. for to habben al þa æhte*
Loctrin and Camber to the ships came for to have all the goods
'Loctrin and Camber came to the ships to take all the goods.'
(Layamon, Caligula 1113–4)
- (40) *moche he lofde ehtm(e) criht. þat lofde for to segg(e) riht*
much he loved every knight that loved for to say truth
'Much he loved every knight who loved to say the truth.'
(Layamon, Otho, 5523)

So, between OE (36) and ME (40), *for* has been reanalyzed from P to C, i.e. is located higher in the tree. This means that, out of the original meaning of *for*, the benefit and future features become grammaticalized. At first sight, these constructions don't seem to be cases where the P actually moves from the lower PP to C, as with the verb from VP to I in (35) above, and where Late Merge would be simpler than merge and move. I will argue they are nevertheless instances of Late Merge. My explanation follows Kayne's (2001; 2002) account of certain prepositions. He argues that prepositions such as 'of' are merged outside the VP. Chomsky (1995: 232) states the same insight as follows: "overt movement of β targeting α [...] is possible only when α is nonsubstantive [read functional]". Adapting that to (36) to (40), I argue that once a preposition is no longer fully lexical but is used to express purposive, as in (37), and dative Case, the PP of which it is the head moves outside the VP (to the Spec of CP). Once it moves outside VP, it can be reanalyzed as a complementizer indicating purpose, e.g. in (40). This accounts for the change from P to C. Prepositions are notoriously 'in-between' categories, i.e. can be either lexical or functional.

In many early English sentences, *for* as head of the PP is indeed at the beginning of the sentence, e.g. as in (41) and in (42):¹³

- (41) *þa wile þe ich hæwede mi kinelond. lueden me mine leoden.*
then while that I had my lands loved me my people
for mine londes 7 for mine feo. mine eorles fulle to mine cneo
for my land and for my property my earls full to my knees
'While I had my lands my people loved me. All my barons knelt at my knees for my land and property.' (Layamon, Caligula 1733–4)
- (42) *þu 3ef þeseluhen for me to lese me fra pine*
'you gave yourself to me to release me from pain.' (Wohunge 88–9)

According to van Dam (1957: 6), this fronting occurs regularly in OE. In (42), *for* is ambiguous between P and C, and hence the language learner ends up reanalyzing the P as C, and the DP as a topicalized element. In this connection, it is remarkable that the first instances of *that*-deletion listed in the OED (entry for *that* II 10) are as in (43) and (44), from the 14th century, i.e. where a *for*-phrase has been fronted and can serve as C:

- (43) *I dred me sare, for benison He sal me giue his malison*
I dread me sore for blessing he will me give his curse
'I am afraid, instead of a blessing he will give me cursing.'
(Cursor Mundi, Cotton, 3665)
- (44) *Joab ... slowh Abner, for drede he scholde be ...*
'Joab killed Abner, out of fear that he should be ...'
(Gower, Confessio I. 263)

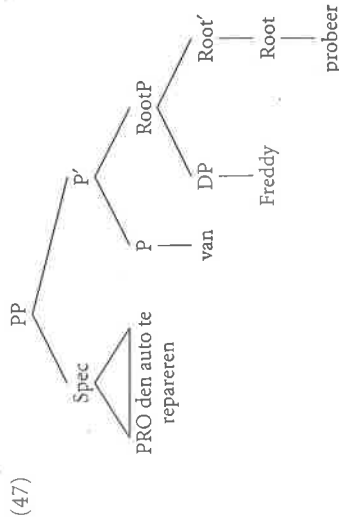
I think these are not cases of *that*-deletion at all but cases of a PP moving to the CP, resulting in *for* being reanalyzed in a higher position. This means that it doesn't have to merge and move, but can just merge late. So these cases are indeed similar to (35) above, with the change going from merge and move to Late Merge. Other evidence is that, in some sentences, *for* clearly 'belongs' in the lower VP, e.g. in (45):

- (45) *Seoð seide Cordoille. for cuð hit is me noupe*
true said Cordoille for known it is me now
'Cordoille spoke true: because it is now known (to) me.'
(Layamon, Caligula 1727)

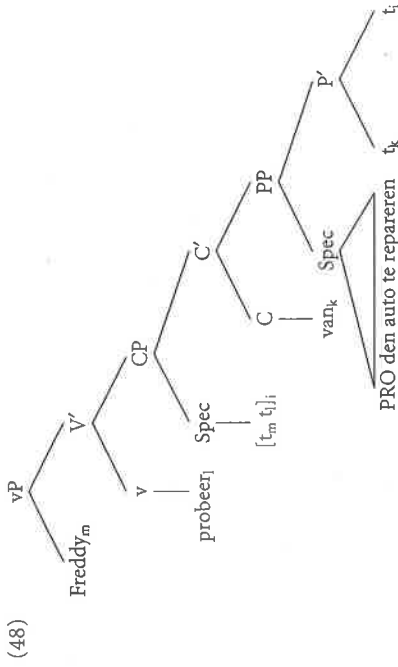
In Chapter 5, I come back to *for* in particular where it introduces a finite clause.

Van Craenenbroeck (2002) discusses a similar situation in South Brabant Dutch where the preposition *van* 'of' becomes a complementizer indicating aspectual/evidential dissociation. There are three stages, all still present in the dialect: one where *van* is a clear P that is merged as P and doesn't move, a second where it is merged as P but then moves, and a third stage where *van* is merged high in the tree. To show 'Late Merge' (not mentioned in van Craenenbroeck), the crucial stage is the second one. Basically following Kayne (1999; 2001), the structure for a sentence such as (46) starts out as in (47) (with a RootP rather than a VP-shell):

- (46) *Freddy probeert van den auto te repareren*
 Freddy tries of the car to repair
 'Freddy tries to repair the car [but won't succeed]'. (p.49)



The subsequent stage in the derivation is one where *van* moves to C (right above PP) and the RootP to the specifier of CP. After that stage, a VP is added to the CP and the subject *Freddy* and the verb *probeert* move to their ultimate positions. Van Craenenbroeck's final tree is the somewhat unorthodox (48):



Hoekstra (1995) has a similar analysis of *van* being outside the VP and having a predicate move to its specifier. So, like English *for*, Dutch *van* ends up in a higher position.

There are many similar cases of P becoming C that will be discussed in Chapter 6. Some of the formulations, in other frameworks, come quite close to Late Merge. For instance, Heine & Claudi (1986: 105) describe a stage in the change from P to C, quite like (41) and (42), as "[f]or some time, the emerging embedded clause is confined to that position in the sentence which is occupied by the adpositional phrase from which it emanated. Later on, it tends to gain more syntactic independence...". Psycholinguistic evidence that the PP and CP are very similar is provided in e.g. Josefsson & Håkansson (2000).

Emonds (2001) uses a different version of Late Merge to account for Romance restructuring verbs. He argues that, due to the semi-lexical nature of verbs such as Italian *cominciare* 'begin', and French *laisser* 'let', the latter can merge late (but don't need to) and are then allowed to be ternary branched structures. In answer to the question why these semi-lexical verbs do not always merge late, he appeals to pragmatic effects (2001:58).

3. Conclusion

In this chapter, I have provided evidence for two Economy Principles, namely Spec to Head and Late Merge in both synchronic grammars and language change. Before looking at historical changes involving the CP in the light of these principles, I will first discuss the structure of the CP in Chapter 3. In

Chapter 4, I will use the first Economy Principle, Head-over-Spec, and in Chapters 5 and 6, the second, Late Merge. Part III is organized the same way. After first looking at the structure of IP and VP in Chapter 7, I apply the Late Merge principle to account for modals in Chapter 8, perception verb complements in Chapter 9, and infinitival *to* in Chapter 11.

PART II