

THE CONJUNCTION *THAT*:
A SEMANTICALLY EMPTY PARTICLE?

RICHARD J. WATTS

University of Zurich

1. *Introduction*

Ferdinand de Saussure's well-known axiom that a language may be studied *diachronically* or *synchronically* is certainly one of the major factors which has contributed towards the overriding interest in synchronic linguistics during this century. The student's approach to theoretical linguistics these days normally lies along a path mapped out by structuralists and functionalists, by American descriptivists and functionalists, and by various schools of generative linguists. It tends to be strewn with the empty shells of obsolete theoretical models and to present the unsuspecting student with unscalable walls of logical and mathematical formalism. So it is hardly surprising that in the effort to overcome all the obstacles our student (and possible future linguist) shows a tendency to ignore the historical aspects of language study.

Henry M. Hoenigswald, however, warns us that it would be radically wrong to assume that "historical and 'comparative' linguistics... was overtaken by the new synchronic wave a few decades ago" (1977:168). And the growing number of international conferences and colloquia on historical linguistics during the last ten years or so certainly adds weight to his timely warning. Indeed, since the days of de Saussure interest in the diachronic aspects of language study has clearly not abated at all. We might then justifiably state that the mere fact that historical research has been carried on doggedly and purposefully all this time is an indication of the mature state it has reached as well as a sign that it has been largely successful.

Some generativists, notably Paul Kiparsky, Elizabeth Closs Traugott and John Anderson, are aware that it was not de Saussure's purpose to split linguistics into two camps, one synchronic and the other diachronic. The axiom

should obviously be interpreted as a call to linguists to concentrate their efforts more on the synchronic study of language, but de Saussure was above all a historical linguist, albeit a rather disenchanted one at the time he held his famous Geneva lectures. He was convinced that an atomistic study of isolated linguistic elements without considering how these elements are structured in a coherent communicative system is rather pointless. But this does not mean that some other, more structured approach to the historical development of language and the principles underlying and explaining language change is necessarily pointless. The dichotomy was not meant to be understood as an exclusive disjunction (i.e. either we study language diachronically or we study it synchronically, but not both), but as an inclusive disjunction (i.e. we may study it diachronically, synchronically, or both). It is the area of overlap in the inclusive disjunction which is of most interest to modern historical linguists. Much of the formalism which has been developed in the synchronic study of language has been adapted to diachronic purposes, and much present-day work in historical linguistics is aimed at demonstrating that the two approaches are essentially interlocking and mutually interdependent. I feel sure that this is what de Saussure meant to express with the synchrony-diachrony axiom, and the following discussion should be viewed in this light.

A synchronic linguist is likely to claim that it is not essential to have a knowledge of the grammatical structures of a language in earlier stages of its development in order to set up a contemporary grammar of that language. Clearly he is justified in making such a claim. If he were not, it would, for example, be impossible to teach or learn the language as an L2 without constantly making reference to those earlier stages. But there are cases in which the synchronic linguist might well be advised to dip into the history of the language in order to avoid certain false assumptions concerning its contemporary state. One such area in the grammatical description of English is the conjunction system, in particular that part of it which deals with the relative pronouns, the temporal conjunctions and the nominal conjunction *that*. It is my purpose in the present paper to demonstrate that one of the explanations for the conjunction *that* which is commonly propagated in transformational-generative grammar rests upon a false assumption. I shall demonstrate this by considering the development of the conjunction from Old into early Middle English. In doing so I shall be working in the overlap area of the inclusive disjunction between synchronic and diachronic linguistics.

2. Nominal *that* in generative analyses

Early transformational analyses of contemporary English usually only included a rule for the generation of *that* in nominal embeddings, such as those contained in the following sentences:

- (1) The Prime Minister stated *that* the unemployment rate had decreased by ten per cent since last January.

In (1) *that* introduces a so-called object clause.

- (2) *That* rhododendra grow well in woodland areas is not surprising, since they thrive on peaty soils.

In (2) *that* introduces a so-called subject clause.

- (3) It's amazing *that* many people still believe in ghosts.

In (3) *that* introduces a subject clause, which has been moved to the right of the sentence structure by means of the *extraposition* transformation.

In those studies nominal *that* was generated by one of the phrase structure rules. The NP-rule in Chomsky (1958) contains the following sub-rule:

- (4) NP → *that* + S

Looking back over the history of transformational-generative grammar, one cannot help wondering whether it might not have been the simplest solution, and therefore the most elegant. The generation of *that* would have remained purely syntactic, and Chomsky would have been forced in the *Aspects* model to explain the appearance of *that* as a lexical insertion replacing a complex symbol of syntactic features, one of which would have had something to do with coreference. The rewriting of NP as *that*+S creates a mixture of syntactic category and lexical item on the same level of the derivation. Ways in which complex sentences with *that* should be derived were never explicitly discussed in *Aspects*. If they had been, the generative semanticists might well have made an attempt to explain *that* semantically. On this point, however, we can only speculate.¹

Towards 1970 the conjunction *that* in nominal embeddings was introduced into the sentence structure by a rule of Chomsky-adjunction, to which Chomsky gave the name "*that*-insertion". This was, however, a purely syntactic reflex

¹ One possibility of explaining *that* in generative semantic theory might be to take the conjunction as the surface structure realization of a binding variable in the semantic representation linking the embedded sentence with the noun phrase object or subject of the verb. It would thus have the function of a cataphoric pronoun. This suggestion differs only from the complementizer theory to be presented below and from Chomsky's more recent "trace" theory in that the binding variable would then automatically assume all the syntactic and semantic features of the embedded sentence. Since this is roughly the line I shall take with respect to *that*, it might be worth pursuing in more detail. It is interesting to note that certain verbs of replying and commenting such as *reply*, *answer*, *add*, *comment*, *point out* etc. will not allow *that* to be deleted. This may have something to do with the different semantic structures within this class of verbs. For more detailed information on the notion of binding variable cf. McCawley (1970), Lewis (1972), R. Watts (1976).

and did not represent a lexical insertion substituting a feature bundle. As such I consider it to be a somewhat *ad hoc* method of explaining the presence and function of *that* and it was in fact soon given up in favour of Joan Bresnan's (1972) "complementizer analysis".

Bresnan noted that it would be difficult to explain such sentences as the following with a Chomsky-adjunction rule:

- (5) The fact *that* wages have increased is of little importance.
- (6) We reject the claim *that* the police were directly responsible for the violence.
- (7) I was plagued by the uncanny feeling *that* the roof would fall in any minute.
- (8) The belief *that* criminals have forfeited their right to a place in society is unfortunately widespread.
- (9) John worked on the assumption *that* Mary would bring the whisky.

Bresnan was of the opinion that such embedded sentences were not nominal. On the other hand, they are certainly not relative embeddings, since we cannot replace *that* by *which*. To explain such sentences she suggested that any and every embedded sentence must contain a complementizer node, which she symbolized as COMP and which she allowed to be generated in the base component.

Geoffrey Pullum and Deirdre Wilson (1977:773) have set up the following subcategorization of the COMP node:

- | | | |
|--------|--------------|-----------|
| 10) a. | →[COMP] | →[±THAT] |
| | b. →[+THAT] | →[±TENSE] |
| | c. →[+TENSE] | →[±PAST] |
| | d. →[-THAT] | →[±TO] |
| | e. →[+TO] | →[±FOR] |

It is clear from rule (10) that no attempt is made to explain what the nominal conjunction *that* is. It is simply generated as a feature [±THAT] which is replaced by the lexeme *that* at a later stage in the derivation. Rule (10), however, displays a more serious flaw. I know of no English sentences that contain tenseless embeddings commanded by the conjunction *that*, but precisely this type of fictive structure could be generated by rule (10). If we chose to retain the complementizer analysis (with the appropriate adjustments), we should have to assign to *that* a semantic feature relating to its obvious deictic function, as we should also need to do for pronominal occurrences of *that*. We should have to suggest a set of pragmatic inference rules to explain the fact that hearers understand nominal *that*-clauses to be statements of fact or statements governed by a personal attitude of the speaker's, since they are tensed propositions embedded into speaker-oriented attitudinal predicates or locu-

tionary predicates. Regarding nominal *that* as either a purely syntactic reflex or an empty particle substituting for the syntactic feature [+THAT] in the COMP node is thus tantamount to ignoring the semantic and pragmatic information implicit in its use.

Before we go on to consider relative and temporal embeddings with *that*, one further remark is in order concerning sentences (5)–(9). Note that apart from the noun *fact* all the other nouns preceding the conjunction *that* are nominalizations of speaker-oriented attitudinal or locutionary predicates. It stands to reason, then, that the embedded sentences with *that* in (5)–(9) are after all nominal embeddings.

3. Relative and temporal that

The complementizer analysis is rather appealing, since it allows an apparent simplification of the linguistic description of English. The conjunction *that* also occurs in relative and temporal clauses, as is evident from sentences (11)–(14). A more conventional type of grammar would classify it in this environment as a relative pronoun:

- (11) This is the house *that* Jack built.
- (12) Use the money *that's* lying on the table.
- (13) The girl *that* you were dancing with is my fiancée.
- (14) You'll never live to see the day *that* John comes on time.

It is perfectly possible to give plausible deep structures for (11)–(14) in which the embedded sentence introduced by *that* is generated as part of a noun phrase. From a traditional point of view the sentences here can be classified as "defining", i.e. they have the same status as definite articles and defining adjectives in that they create and specify a subset from a more general set. In most versions of transformational-generative grammar² a defining relative clause would be generated as part of the noun phrase and commanded by the determiner. But it is precisely this possibility which tempts Joseph Emonds (1976) to go much further than Bresnan and to generate every occurrence of *that*, whether it is a nominal conjunction or a relative pronoun in relative and temporal clauses, as a semantically empty complementizer. In (11)–(14), however, *that* is coreferent with a noun phrase in the matrix sentence — with *the house*, *the money*, *the girl* and *the day* — and as such it must assume the semantic and syntactic features of that noun phrase.

² Not all generativists are in agreement here; many argue that the restrictive or defining relative clause must be generated under a noun phrase. They are then forced to analyse the non-defining relative clause as a coordinate structure. Cf. e.g. Huddleston (1976).

Emonds's analysis depends on the very shaky acceptability of such sentences as the following:

(15) ?*That fellow that I just saw you with him is my brother.

(16) ?*Have you seen the film that I was telling you about it last week?

For me such sentences as (15) and (16) are simply not possible, and I have been told by American informants that in American English they sound somewhat esoteric. Yet it is on such sentences as these that Emonds bases his arguments. The reason is clear: according to the complementizer analysis nominal *that* is a semantically empty marker, so if relative and temporal *that* can be analysed in the same way, the grammar of English can be constructed much more elegantly and simply. If we are looking for simplicity, however, it might be more logical to look at things the other way round and to integrate the explanation of nominal *that* with that of relative and temporal *that*, as I shall do in the present paper.

What first needs to be explained is why, in contemporary English, the uninflected pronoun *that* is used at least as frequently as, if not more frequently than, *who(m)* and *which* in defining relative embeddings, and why, in non-defining relative embeddings, only *who(m)*, *which* and *whose* are acceptable.³ Note that temporal *that* is also restricted to defining temporal embeddings.

It might be possible to refute the complementizer analysis purely on the basis of data from contemporary English. The argument I shall use, however, will be diachronic. This is not because a diachronic analysis is necessarily more revealing than a synchronic one, but because the state of affairs in Old English with respect to relative and temporal *that* (or its equivalent at that stage of the language) is remarkably similar to the present situation and might thus be taken to reveal a structural characteristic peculiar to English throughout its development.

4. Relative and nominal connectives in Old English

As in modern English, the Old English nominal conjunction was *þæt*. The relative pronoun, on the other hand, was realized in a number of ways. It is generally thought that Proto-Germanic did not possess a relative pronoun. If this is so, it certainly goes some way towards explaining the variety of means in which relative structures were expressed in Old English. Torben Kisbye (1972:128) lists the following four principal types:

a. The demonstrative pronoun *se/sēo/þæt* is used.

b. The uninflected particle *þe* is introduced as a relative pronoun.

³ I should point out here that the odd occurrence of *that* as a clear non-defining relative pronoun can occasionally be found in contemporary English.

c. The demonstrative pronoun is inserted before the particle *þe*.

d. The particle *þe* precedes a pronoun which in effect belongs to the matrix sentence, but which appears in the relative clause immediately after *þe* to indicate the case and number of the relative particle.

The fourth type is a special case which need not concern us further here. The third type, with the combination of *se/sēo/þæt* and the particle *þe*, is generally used in sentences in which the demonstrative pronoun refers back to a noun phrase that has already appeared in the matrix sentence. The relative clause introduced by *þe* thus defines the demonstrative pronoun. The first two types are of more significance for the following discussion. In most, although not all, of the cases which I have studied *þe* is the equivalent of a *defining* relative pronoun, whereas *se/so/þæt* corresponds to a *non-defining* relative pronoun. Exactly as in modern English, however, *se/sēo/þæt* can frequently be found functioning as a defining relative pronoun, whereas cases of *þe* functioning as a non-defining relative pronoun are extremely rare indeed. In fact, we seem to have a parallel situation to that in modern English, except that a shift has taken place in the lexemes functioning as relative pronouns.

Before moving on to consider the historical development of *þæt* and *þe* from Old into early Middle English, we need to consider the logic that lies behind the distinction made between *þæt* as a nominal conjunction and *þe* as a relative pronoun, since in modern English *that* occurs in both positions.

In Old English *þæt* is not only the nominal conjunction, but also the grammatically neuter form of the demonstrative determiner and the demonstrative pronoun in the nominative and accusative singular cases. If the nominal conjunction *þæt* were only a semantically empty marker, it would have to be listed as a separate lexeme from the demonstrative determiner/pronoun in the lexicon. There would thus be two distinct entries, *þæt*₁ and *þæt*₂, for which different sets of syntactic and/or semantic features would have to be postulated. This would only complicate the lexicon unnecessarily, however. I believe that we need only consider one entry for *þæt*. It may be generated under the domination of the determiner node (i.e. as a demonstrative determiner) or on its own under the domination of NP (i.e. as a demonstrative pronoun). From the semantic point of view it simply denotes a notion of *unmarked spatio-temporal deixis*.⁴ It would then be interpreted as meaning "not in the proximity of the speaker" and could be pragmatically interpreted for either *anaphoric* or *cataphoric coreference*. Thus, whether *þæt* occurs as a neuter non-defining relative pronoun anaphorically coreferent with a neuter noun phrase in the matrix

⁴ I take spatio-temporal deixis to be equivalent to a concept of proximity, in spatial terms proximity to the speaker and in temporal terms proximity to the temporal orientation point in the text. It is for this reason that *se/sēo/þæt* also functions as the definite article. The marked term is *þēs/þeos/þis*.

sentence or as a neuter pronoun cataphorically coreferent with an embedded nominal sentence is immaterial. It is the self-same lexeme in both cases.

I shall illustrate this with the following two sentences:

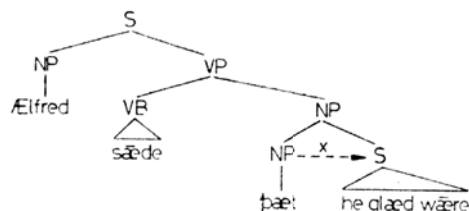
(17) *Ælfred sǣde þæt hē glæd wære.*

(Alfred said that he was happy).

(18) *Ælfred stæl pæs cyniges spere, þæt on wealle hangode.*

(Alfred stole the King's spear, which was hanging on the wall.)

A rough structural description of (17) can be given as follows:

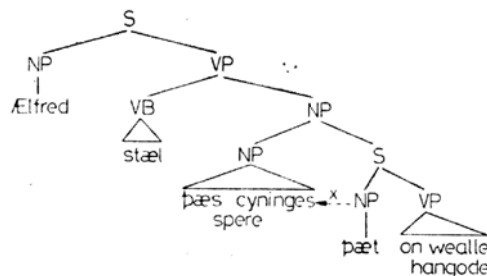


whereby x = cataphoric coreference

þæt realizes unmarked spatio-temporal deixis here and is cataphorically coreferent with the embedded sentence. Note also that the embedded sentence is dominated by NP and commanded by NP.

Consider now the rough structural description of (18):

(20)



whereby x = anaphoric coreference

Once again *þæt* realizes unmarked spatio-temporal deixis in (20) and is anaphorically coreferent with a preceding neuter noun phrase. Had the preceding noun phrase been feminine or masculine, *sēo* or *se* would have appeared in the embedded sentence. As in (19), the embedded S is dominated by NP and commanded by NP.

In point of fact, the analysis of the nominal embedding in (19) does not differ much from the complementizer analysis except in the following two essential points:

- There is no need to generate a COMP node and then to subcategorize it further as [+THAT].
- The equivalent of the COMP node, the NP dominated by NP and commanding S, is the same device as that used by Jacobs and Rosenbaum (1968) and Emonds (1976). It differs from the Emonds analysis in not being an empty dummy node and from the Jacobs and Rosenbaum analysis in not being, rather arbitrarily, filled by the pronoun *it*. Instead it contains the semantic denotation of unmarked spatio-temporal deixis.

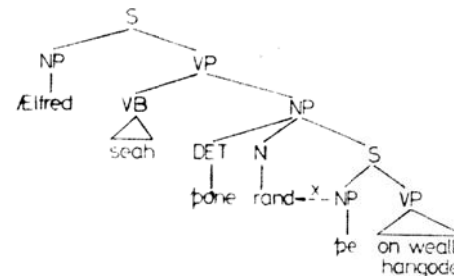
Under this analysis, then, the Old English nominal conjunction *þæt* is a deictic pronoun. Let us now consider the relative particle *þe* as it occurs in the following sentence:

(21) *Ælfred seah pone rand þe on wealle hangode.*

(Alfred saw the shield that was hanging on the wall.)

A rough structural description of (21), which takes into account the fact that the relative clause is defining, can be given as follows:

(22)



whereby x = anaphoric coreference

In (22) *þe* is anaphorically coreferent with the noun *rand* in the NP, and the function of the embedded S is to define a singular subset from all possible *randas*. Because it has an anaphoric function, it must be a realization of unmarked spatio-temporal deixis. Thus, even though it is uninflected, it bears all the semantic denotations of *rand*. The only difference between *þe* and *þæt* is structural. The embedded sentence in which *þe* appears is dominated by S, but commanded by DET^N, not by NP.

5. Defining and non-defining relative clauses in Old English

The deictic denotation is generally recognized by grammarians of Old English. The pronoun itself, however, is merely considered to be demonstrative. Campbell (1959) maintains that the alleged lack of a relative pronoun

in Old English "is largely made up for by the use of *se* either alone or followed by the particle *þe* (which shows it to be relational, not demonstrative in function), e.g. *se mon-dryhten se eow ðā māðmas geaf* 'the lord who gave you the treasures', *his feoh þæt tō lāfe biþ* 'his money which shall remain', *se ellen-gæst se þe in þystrum bād* 'the mighty spirit who lingered in darkness'" (1959: 291).

I interpret Campbell in the following way: If *se/sēo/þæt* is followed by *þe* (cf. Kisbye's relative pronoun type c. above), then it is a relative pronoun. I have already argued, however, that only the particle *þe* can have this function in such a structure. If, on the other hand, *se/sēo/þæt* occurs alone introducing an embedded clause (cf. Kisbye's relative pronoun type a. above), then it has the function of a demonstrative pronoun. In this case, the embedded clause is not embedded at all, but is an independent matrix sentence in apposition to the first. This type of interpretation, however, ignores the possibility that there were defining and non-defining clauses in Old English, just as there are in modern English. It also reflects a conviction that the syntax of Old English was paratactic rather than hypotactic. My short analysis of Old English relative and nominal clauses above provides arguments against both these assumptions. In any case, all the phrases quoted from Campbell are interpretable as containing defining relative clauses. This fact is signalled syntactically by the clause final position of the verb.

Andrew (1940: 35—47) argues that the role of word order in Old English has frequently been underestimated. With certain types of temporal embedding it was a crucial syntactic marker, as can be seen from the following sentence:

(23) *pā hē þæt ne gefafode, pā hēt hē hine lædan upp on sume dūne ond hine pær behēafdian.*

(When he did not agree, he had him led up onto a hill and there beheaded.) When the lexeme *pā* functions as a temporal conjunction, the verb moves to the verb final position, *pā* is repeated (in the sense of *then*) as an introduction to the matrix sentence and thus causes the inversion of the subject and the verb.

On the basis of an analysis of the syntactic role played by word order in Old English Andrew is able to provide very convincing arguments that Old English relative clauses could and should be split up into defining and non-defining. He concludes that *se/sēo/þæt* only fulfills the function of a demonstrative pronoun if it does *not* display anaphoric coreference with a noun phrase *immediately preceding* it in the matrix sentence, or if it displays cataphoric coreference with a relative clause immediately following (cf. Kisbye's relative pronoun type c. above). If there is an anaphoric coreference with an immediately preceding noun phrase, it must be understood as a non-defining relative pronoun. This is completely in line with my arguments so far. What I have done, however, is to go a step further and suggest that, whether as a

demonstrative pronoun, a non-defining relative pronoun or a nominal conjunction, *þæt* always realizes the same semantic denotation.

Both types of relative clause are contained in the following sentence:

(24) ...*se eahtatēoða dæg pæs mōnðes þe wē hātað Martius, pone gē hātað Hlyda.*

...the eighteenth day of the month that we call March, which you call Hlyda.

The translation into modern English shows the difference between the defining and the non-defining clauses very clearly.

6. Nominal, relative and temporal embeddings in the Peterborough Chronicle

With respect to the development from Old English to early Middle English there are a number of texts which could be investigated. In the present study, however, I shall concentrate on the *Peterborough Chronicle*, which recounts events in England from 1070 to 1154. It is important for a number of reasons. It shows the transition from Old to Middle English very clearly and was written in the East Midland dialect, which is generally taken to be the most important ancestor of modern English. It is assumed that the first part of the text from 1070 to 1122 was copied by a scribe, probably at Peterborough Abbey, from a text which had been compiled earlier, with the last few years filled in from memory.⁵ This is particularly clear in the so-called "First and Final Continuations" of the *Chronicle* covering the period from 1122 to 1154, which are assumed to have been written down at Peterborough. The "Final Continuation" in particular is in part a very vivid narrative in a form of English which differs from the first part of the *Chronicle* from 1070 to 1122. It is probable that the scribe was recording what had happened and was still happening during his lifetime.

6.1. 1070—1122

In the first part of the *Chronicle se/sēo/þæt* with the grammatical functions of determiner, demonstrative pronoun and relative pronoun are still inflected, although a tendency to mix up gender classes is in evidence. I take this to be less a question of changes in gender classes than an overt sign that grammatical gender was losing ground rapidly. Uncertainty concerning which noun belonged to which declension and gender class and the consequent "wrong" assignment of forms from the *se/sēo/þæt* paradigm are thought to

⁵ The text of the *Chronicle* to which I shall refer in this study is that edited by Cecily Clark (1958).

have been the result of a phonological weakening and obscuring of the relevant morpho-syntactic markings and to have led to a shift from grammatical to natural gender. In both "Continuations" this process has progressed so far that only the genitive *þes* is used more or less consistently. And even here the periphrastic form with *of* + NP appears to have taken over in many cases from the genitive.

Let us confine ourselves for the moment to the relative pronouns and the nominal conjunction. As might be expected, the relative pronoun system in the entries from 1070 to 1122 is more or less the same as in Old English, and the nominal conjunction is exclusively *þæt*, later *þet*.⁶ The entry for the year 1096 begins with the following sentence:

- (25) Eac on þison geare þa heafodmen þe þis land heoldan...⁷
In this year too the "headmen" who held this land...

þe functions in (25) as a defining relative pronoun. In the entry for the year 1097 we read the following somewhat more complex sentence:

- (26) Sum þære wæs Caduugaun gehaten þe heora weorðast wæs, se wæs Griffines broðer sunu cynges.

In (26) there is one occurrence of *þe* and one of *se*, and the word order in the sentence makes it rather difficult to tell whether *þe* may logically be called a defining relative pronoun. If we classify it as nondefining, we are left with no option but to take *se* as a demonstrative pronoun. As it stands, there seems to be no way out of the predicament, since *sum þære* 'one there' is defined by the predicative *Caduugaun gehaten* 'called Cadwgan'. If we translate the sentence exactly as it stands, we get:

- (27) One there was called Cadwgan, who was the worthiest of them, "who" (= he) was King Griffin's nephew.

On the other hand, the sentence opening with *se* appears to offer extra information about Cadwgan, whereas the sentence opening with *þe* appears to be closely related to *sum þære wæs* and to restrict the possible range of *sum þære*.

If we alter the order of the embedded sentences in (26) a little, we arrive at the following sentence:

- (28) Sum þære þe heora weorðast wæs wæs Caduugaun gehaten, se wæs Griffines broðer sunu cynges.

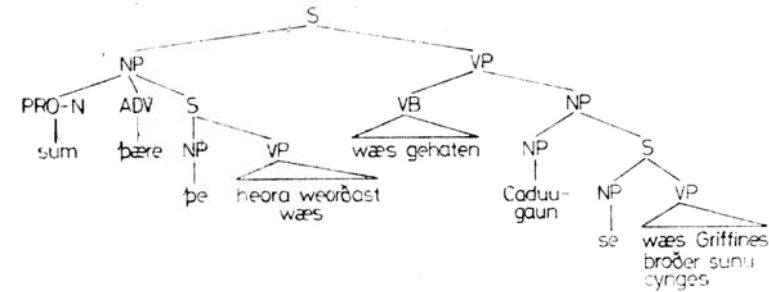
⁶ The *e* of *þet* is a sign that the original *æ* sound has undergone a process of weakening to the schwa vowel.

⁷ At this stage there was already some confusion as to how the nominal and verbal inflections should be spelt, since they must all have been weakened to the schwa vowel. Hence we have *þison* for *þissum* and *heoldan* for *heoldon*.

One there that was the worthiest of them was called Cadwgan, who was King Griffin's nephew.

Sentence (28) quite clearly contains a defining and a non-defining relative clause and can be given the following rough structural description:

(29)



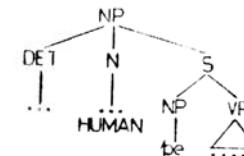
To reach (26) we need only extrapose the two nominal embeddings. Indeed the extraposition of relative and nominal clauses seems to have been a fairly common type of transformation and may go some way towards explaining the impression of parataxis which so many grammarians comment on. At all events, it should be clear that my analysis of (26) is easily able to preserve the distinction between defining *þe* and non-defining *se/sēo/þæt* in the first part of the *Peterborough Chronicle*,

6.2. The language of the "Continuations": the transitional stage

After the entry for 1122 the defining and non-defining relative pronouns are no longer clearly distinguished. There are still a few examples of *se/sēo/þæt* functioning as a non-defining (or even a defining) relative pronoun, but by this time the movement towards natural gender has progressed so far that a wholesale shift is underway throughout the determiner, relative and demonstrative pronoun system,⁸ for which I shall offer an explanation in section 8.

In the "Continuations" the Old English relative pronoun *þe* functions primarily in defining relative clauses and is coreferent with a preceding noun entailing the semantic denotation HUMAN. The abstract configuration of the NP can be given as follows:

(30)



⁸ This does not, however, appear to be the case with the temporal conjunctions.

The following sentence is a clear example of the structure in (30):

(31) Sume ieden on ælmes þe wæron sum wile rice men.

Some who were previously rich men were forced to accept alms.

On the other hand, the redundancy of the inflectional paradigm for *se/sēo/þæt* in a natural gender system and the consequent reduction to one uninflected definite determiner meant that these forms were no longer available as non-defining relative pronouns. Thus *þæt*, or *þet*, which was still in service as a nominal conjunction, was released, as it were, to do added service as a defining or non-defining relative pronoun coreferent with a preceding noun phrase not entailing the semantic feature HUMAN.

Why was it possible for *þet* to take on this function? The reason, I suggest, is directly deducible from my analysis of *þæt* as a pronoun realising unmarked spatio-temporal deixis. Coreference with a sentence structure in Old English (cf. section 4 above) was in effect equivalent to coreference with a neuter noun phrase. Thus the only form needed as a nominal conjunction was that marked as neuter, viz. *þæt*. Similarly, coreference with a preceding noun phrase that was marked as neuter meant that the pronoun itself also had to be marked in the same way.⁹ Both the nominal conjunction and the relative pronoun simply realised unmarked spatio-temporal deixis in the total sentence structure with a following embedded sentence or a preceding neuter noun phrase. In a natural gender system an embedded nominal sentence has the same status as a noun-phrase that denotes the semantic feature NON-HUMAN. Thus it was a completely logical move, once *se* and *sēo* had disappeared as relative pronouns, to maintain *þæt* in order to indicate coreference with a preceding noun phrase that entailed the feature NON-HUMAN. An example of *þet* in this function is given in the following sentence:

(32) ...and þe munecas herden þa horn blawen þet hi blewen on nihtes...
...and the monks heard the horns blow that they blew at night...

In the "Continuations" we also find a pointer towards the later predominance of *þet* as a relative pronoun in Middle English, since it is sometimes used with HUMAN antecedents in defining and non-defining clauses. A third relative pronoun *þa* is also in evidence in the "Continuations". It is used mostly in non-defining clauses where the antecedent NP entails the semantic feature HUMAN or in defining clauses where the antecedent is plural (particularly after the quantifier *ealle* 'all').

The transitional stage between Old and Middle English, which is manifested

⁹ Of course, if the noun phrase were masculine or feminine, then the relative pronoun would also have to be so marked.

rather clearly in the "Continuations" of the *Peterborough Chronicle*, provides ample evidence that the English relative pronoun *that*, whether defining or non-defining, is not merely a semantically empty syntactic complementizer. In the case of *þæt* and *þa* the semantic function of deixis is still in evidence. Both these and the relative pronoun *þe* are coreferent with a preceding noun phrase and must thus be understood as identical in terms of the semantic information which that noun phrase provides.

7. Opacity and transparency in language change

Before we trace the development of the relative pronouns from Old English into Middle English any further, two basic principles of generative studies in historical linguistics will be introduced:

- a. The *principle of simplicity* is of crucial importance to the generativist. The best grammar is also the simplest, whereby no value judgement should be attached to the word "simple" in this context. Linguists merely try to set up a grammar which will account for as large a range of linguistic data as possible with the smallest number of rules. Several generativists are of the opinion that all children learn their mother tongue according to the principle of simplicity, i.e. if they have a choice between one or more rules, they will accept the rule that appears to be the simplest. This last assumption, needless to say, has not gone unchallenged.
- b. Michael Peinovich (1979) has taken over and redefined Kiparsky's *principle of opacity and transparency* in relation to the generative description and explanation of changes in the linguistic system. An opaque rule is one whose effects are obscured by rules which follow it in the ordered sequence of rules needed for the penetration of a linguistic structure, where is a transparent rule is one which remains unaffected by subsequent rules. The differences between that which is opaque and that which is transparent are obviously not absolute, but are rather on a cline between the two extremes. Peinovich calls his principle "The Principle of Grammatical Transparency", which he defines as follows:

(33) Language learners tend to construct maximally transparent grammars.

(33) should not be taken to mean that every change necessarily leads to a simpler or more transparent form of the linguistic system which the child has to learn. Quite the contrary is the case. By making one part of the grammar more transparent opacity is often introduced in another area of the grammar, which will in its turn demand to be made transparent.

8. *Opacity and transparency in the "Continuations"*

The situation described above in section 6 in the system of relative pronouns, demonstrative pronouns and definite determiners in the transitional period, as this is revealed in the *Peterborough Chronicle*, is a classic case of the movement from transparency to opacity and thence back to transparency.

If I am correct in suggesting that *se/sēo/þæt* and *þe* simply denoted unmarked spatio-temporal deixis, then the system of Old English was reasonably transparent in this area of the grammar. The syntactic configurations in which the semantic denotation was realised determined whether *þe* or *se/sēo/þæt* was used. There were three syntactic criteria according to which the choice was made:

- whether or not the deictic coreference occurred in an embedded or a non-embedded structure (i.e. whether the nominal and relative elements were generated or the determiner and demonstrative pronoun),
- whether or not the coreference was appositional (i.e. whether coreference was from one NP to another NP — as in the case of the nominal conjunction, the non-defining relative pronouns, the definite determiner and the demonstrative pronoun — or from an NP to an N — as in the case of the defining relative pronoun),
- whether the coreference was anaphoric or cataphoric (i.e. whether the element concerned was a relative or demonstrative pronoun, or a nominal conjunction).

In appositional syntactic configurations the coreferent element had to show grammatical gender, surface structure case and number. In the case of the cataphoric nominal conjunction the situation was simple: an embedded sentence, with which the nominal conjunction was coreferent, was neuter, singular and either nominative or accusative, since such sentences would either be subject or object clauses (cf. sentences (1) — (3) in section 2 above). Hence only *þæt* could occur.¹⁰ In the case of the determiner, the demonstrative pronoun and the non-defining relative pronoun gender, case and number all had to be considered and the appropriate choice made from the full paradigm of *se/sēo/þæt*.

The significant steps on the road towards the situation revealed by the "Continuations" of the *Peterborough Chronicle* may be glossed roughly as follows:

- Phonological change in the vowel system of Old English led to the weakening and levelling out of the morpho-syntactic inflections in the noun classes. This created a situation of *opacity*, since nouns could no longer be placed into clear gender paradigms.

¹⁰ I shall later modify this statement slightly in the discussion of the temporal conjunctions.

- The "correct" forms of the determiners and demonstrative pronouns could no longer be placed automatically and were consequently confused. The *opacity* was increased.
- The move away from grammatical to natural gender represents a move towards *transparency*. The major distinction was now made between HUMAN and NON-HUMAN, and, among those nouns that denoted HUMAN, between MALE and FEMALE.¹¹
- As a result of the move towards natural gender the definite determiner gradually became the uninflected *þe*, thus representing a further move towards *transparency*.
- Concurrently, the non-defining relative pronouns were simplified along the lines of natural gender. Case and number also ceased to play a morpho-syntactic role, and the basic distinction was drawn between a non-defining relative pronoun denoting NON-HUMAN, viz. *þet*, and one denoting HUMAN, viz. *þa*. This I take to be a move towards *transparency*. The new relative pronoun *þet* denoting NON-HUMAN also assumed this function in defining clauses, thus taking over part of the area covered by the single Old English defining relative pronoun *þe*. For all defining coreference relationships with nouns denoting plurality, whether HUMAN or NON-HUMAN, *þa* was used, thus leaving *þe* to cover only coreference with singular nouns denoting HUMAN.

Taken as a whole, therefore, the relative pronoun system in the "Continuations" is decidedly *opaque*. The relevant pronouns with their associated denotations and syntactic functions can be summarized as follows:

	defining	non-defining
(34) NON-HUMAN sing.	<i>þet</i>	<i>þet</i>
NON-HUMAN plur.	<i>þa</i>	<i>þet</i>
HUMAN sing.	<i>þe</i>	<i>þa</i>
HUMAN plur.	<i>þa</i>	<i>þa</i>

Because of its complexity and consequent opacity a situation such as this is unstable. The first move towards restoring transparency was the disappearance of *þe* in (34). Not only was it used for one rather unique relative structure, but it was also rapidly gaining ground as a definite determiner. The total relative pronoun system was thus simplified by its disappearance. The next question was whether *þa* or *þet* should fill up the gap.¹² From a statistical point

¹¹ Things were obviously not quite as simple as I am making them out to be here. Obviously the distinction between MALE and FEMALE could be and was extended to animals. The Old English neuter noun *cild* also tended to retain its "neutrality", i.e. to be classified as NON-HUMAN.

¹² The reader is asked not to take this statement literally. In reality there was of course no gap to fill, since the movement from *þe* to *þet* (*þæt*) must in reality have been gradual.

of view it might seem that an extension of *þa* would be the most logical development. However, *þa* as a plural determiner was on the way out and an extension of it in the relative pronoun system, where it was primarily used with plural nominal antecedents, would, from a grammatical point of view, have represented the least logical move. What is more important about the system outlined above, however, is that the distinction between the defining and non-defining relative clauses has been all but obliterated. I take this fact to be a further and more important case of opacity.

9. Temporal embeddings as nominal and relative embeddings

Let us now turn our attention to the temporal conjunctions. An examination of the "Continuations" reveals that the total system of Old English conjunctions has not changed much in the transition to early Middle English. Sentences with *þa* in the sense of modern English *when* still occur exactly as in sentence (23). This can be seen in the following sentence:

- (35) *þa he to Engleland com þa was he underfangen mid micel wurtscipe.*
When he came to England, he was received with great honour.

Other temporal conjunctions in Old English were, for example, *siððan* 'after' or 'since', *þonne* 'when' or 'whenever', *oð* 'until' and *æf* 'before'. Most of these elements also appear as locative adverbs and prepositions. When functioning as conjunctions they are frequently followed either by a demonstrative pronoun and the relative pronoun, *þe*, or by the conjunction *þæt*. This observation is, of course, by no means new. It has also been made by Karl Heinz Wagner (1969: 52 ff.) in his transformational-generative study of Old English. Wagner, however, makes no attempt to give a structural description of such embeddings. He also confuses the issue by correctly noting that subordinate embeddings are in many cases relative embeddings in prepositional phrases, but then, having already postulated an abstract relative marker *Rel* analogous to *Imp* and *Q*, insisting that we can promptly do away with it, since whether a clause is relative or some other type of subordination will depend on "position in a particular structure". This latter point is, I believe, essentially correct, and my analysis of relative structures above and of temporal structures in this section show it to be so. One thus wonders why Wagner feels it necessary to postulate *Rel* in the first place at all.

The following two sentences are examples of this type of structure:

- (36) *Ðā ic ðā ð's eall gemunde, ðā gemunde ic ēac hū ic geseah — æf ðæm ðe hit eall forhergod wære ond forbærned — hū ðā ciricean giوند eall Angelcynn stōdon mādma ond bōca gefylda.*

When I remembered all this, I also remembered how I had seen — before it was all ravaged and burnt — how the churches throughout the whole of England were full of treasures and books.

(From King Alfred's letter to Bishop Wærfeð)

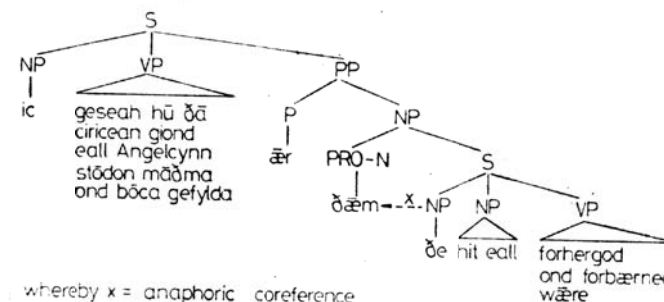
- (37) *[Hē] hergade ealne þone ende þe hē oferfērde, oð þæt hē cōm tō Beorhamstēde.*

[He (William the Conqueror)] ravaged the whole region that he traversed until he came to Berkhamstead.

(From the *Chronicles*, 1066)

A rough structural description for the relevant portion of (36) containing the temporal clause can be given as follows:

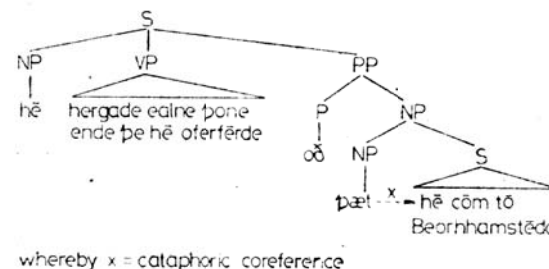
(38)



Admittedly, (38) does not look particularly elegant as a structural description of Alfred's sentence. The motivation for the initial NP generated as *ðe* in the embedded sentence is not immediately clear. But it is undeniably a relative structure, and there appears to be no other way of giving an adequate syntactic description of the phenomenon *ær ðæm ðe*.

Perhaps a clue is offered by the following structural description for (37):

(39)



The embedded sentence in (39) is a nominal sentence commanded by the preposition *oð*. Since *oð* always commands a noun in the surface structure accusative case, the nominal conjunction *þæt*, which I have already analysed as a deictic

pronoun, may remain, and we thus have the sentential analogue of a prepositional object. In (38), however, the preposition *æ̅r* commands a noun in the surface structure dative case. The dative of *þæt* is *þæ̅m* (or *þā̅m*), which cannot stand as a nominal conjunction. For this reason it is copied as a defining relative pronoun with which it is clearly coreferent. *þæ̅m* fulfills a cataphoric and *þe* an anaphoric function.

This analysis of temporal clauses in sentences such as (36) and (37) above provides further evidence for my initial assumption that nominal *that* and relative/temporal *that* are essentially the same spatio-temporal deictic lexeme. Nominal and relative embeddings are thus ultimately related through the semantic denotation of *that*, Old English *þæt*. It is also clear that temporal conjunctions have developed from prepositional phrases in which the noun phrase is commanded by a temporal preposition, and/or the noun phrase itself is a relative or nominal embedding.

In the "Continuations" of the *Peterborough Chronicle* there are several examples which provide evidence of this development. Beside *siððon þet* and *fram þet* 'since' and *to þet* 'till' or 'until' we find *æ̅r* 'before', *til* 'till' and *wile* 'while'. Note also that my analysis of the development from *se/sēo/þæt* to the definite determiner *þe* and the consequent introduction of *þet* as an uninflected defining or non-defining relative pronoun predicts exactly that, whereas in Old English we might have found *fram þæ̅m þe* and *to þæ̅m þe*, we will now find *fram þet* and *to þet*. Case and gender have been lost and with them the combination of the demonstrative pronoun *þæ̅m* and the relative pronoun *þe*. Instead we simply have the nominal conjunction *þet*.

However, in the "Continuations" we still find some examples of noun phrases fulfilling the grammatical function of temporal adverbs and followed by defining relative clauses, as in the following two examples:

- (40) *þa hwile þet se ærcebiscop wæs ut of lande, ... geaf se kyng þone biscoprice of Baðe an clerc.*

While the archbishop was out of the country, ... the King gave the See of Bath to a clerk.

- (41) *þat ilce dæi þat Martin Abbot of Burh sculde þider faren, þa sæclede he and ward ded.*

The very same day that Martin, Abbot of Peterborough, was to travel thither, he was taken ill and died.

Only this type of sentence has survived in modern English with the temporal defining *that*, such as we see in sentence (14). Before the temporal conjunctions became a fixed syntactic category without the accompanying nominal/relative *that*, however, the language had to go through a stage in which *that* became

all powerful in nominal, relative and temporal embeddings, thereby creating a further case of opacity. This development will be sketched out in the following section.

10. Overgeneralization of *þat* in Middle English

In this section I shall refer to two further early Middle English texts written in the East Midland dialect, a version of *Havelok the Dane* and the verse homily of the gospels that goes under the title of the *Ormulum*.

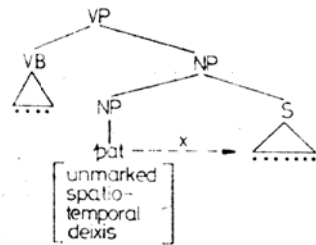
According to information given in the text the *Ormulum* was written by a monk named Ormin. The period in which it was written is generally thought to be sometime between 1200 and 1220, and Peterborough Abbey might very well have been the place at which it was written. Robert Palmatier (1969) has carried out a statistical analysis of the syntax of the *Ormulum* and has reached the conclusion that the linguistic system it reveals is remarkably similar to that of the "Continuations" of the *Peterborough Chronicle*. Nevertheless, it is significant that of the 261 relative clauses in the corpus analysed by Palmatier there are only 7 which are *not* introduced by the relative pronoun *þatt*. *Whatt* occurs 4 times, but always in non-defining relative clauses as the object of a preposition, and *whamm* occurs 3 times as the direct or indirect object of a non-defining relative clause.

The text of *Havelok the Dane* to which I refer (Skeat and Sisam 1939) was written at a later date than the *Ormulum*, sometime around 1280. But in the whole text, whether in defining or non-defining relative clauses, only the pronoun *þat* appears.

Here we have evidence that the disappearance of *þa* as a plural determiner must have contributed to its consequent disappearance within the relative pronoun system. Uninflected Old English *þe* as a defining relative pronoun has been substituted by uninflected Middle English *þat*. At the same time, however, — apart from significant isolated occurrences — *þat* has become generalized as an uninflected non-defining relative pronoun.

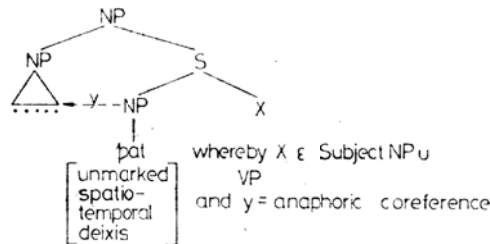
The opaque situation in the system of relative pronouns in the early Middle English period, as this was revealed in the *Peterborough Chronicle*, has been resolved. There is, however, a hitch. To coin a phrase, what one gains on the roundabout one loses on the swings. The radical simplification of this subsystem of the grammar of English has led to an overgeneralized use of *þat*. In each of its three uses as a conjunction it bears the semantic denotation of unmarked spatio-temporal deixis, but as a nominal conjunction it functions cataphorically and as a relative pronoun anaphorically. In addition, as a non-defining relative pronoun it is dominated by S and commanded by NP, whereas as a defining relative pronoun it is dominated by S but commanded by DET^N. The following three abstract structures represent this situation:

(42) a. nominal conjunction:

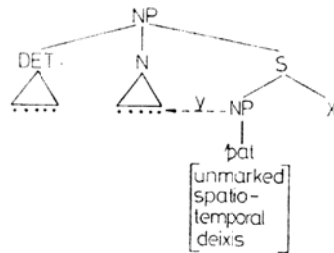


x = cataphoric coreference

b. non-defining relative pronoun



c. defining relative pronoun:



y = anaphoric coreference

This I take to be a renewed situation of opacity, since the generation of *þat* obscures the interpretation of the relevant syntactic structure and renders a correct pragmatic inference more difficult. Note, however, that (42) a. and (42) c. can hardly be confused, since in the case of (42) a. *þat* is dominated by NP and commanded by VB and may only be interpreted cataphorically, whereas in the case of (42) c. *þat* is dominated by S and commanded by DET N and may only be interpreted anaphorically. The obscurity is caused by the failure to distinguish between defining and non-defining relative pronouns.

In section 5 I argued the case for a distinction between the two types of relative clause in Old English. They could be clearly distinguished from one

another by the use of two types of relative pronoun, the one uninflected and defining (viz. *þe*) and the other inflected and non-defining (viz. *se/sēo/þæt*). In addition I argued that the non-defining relative pronoun might also be used in defining clauses, but not vice versa. The few cases of *whamm* and *whatt* in the *Ormulum* indicate that a further shift towards the re-establishment of this distinction was already underway, and indeed during the 14th. and 15th. centuries there emerged a new system of non-defining relative pronouns *what*, *who*, *whos*, *whom* and, above all, *the hwilc* (or *which*), all of which could equally well be used in defining clauses and all of which show semantic coreference with the antecedent NP.

The development of a new relative pronoun system is correlated highly with the decline in the use of *þat* in the system of temporal and causal conjunctions. In Middle English texts of the 13th. century we still find the sort of structure discussed in section 9 alongside what we now recognize as temporal and causal conjunctions. In *Havelok*, for example, we see *hwon þat*, *til þat*, *sipen þat* and *er þat* as well as *hwil*, *þanne*, *til*, *er* and *sipen*.

11. Conclusion

It has been my purpose in this paper to argue for a different type of analysis in sentences containing the conjunction *that* than that which is commonly offered in transformational-generative studies of the modified standard theory type. In particular, I have criticized those analyses which take all occurrences of grammatical lexemes introducing embedded structures to be semantically empty elements merely marking the embedding process as a lexical realization of the COMP node. Such elements as *that* are not generally felt to be meaningful by the native speaker and may be deleted in certain syntactic configurations. But this does not mean that they have no semantic content at all. I have postulated a denotation of *unmarked spatio-temporal deixis* to account for occurrences of *that* as a nominal or temporal conjunction and as a defining relative pronoun. Admittedly it is a rather weak denotation and may be felt to be a somewhat ad hoc solution. Yet its very weakness accounts neatly for its tendency to be deleted if the speaker feels that the coreference relationship may in any case be inferred from the utterance.

What is more significant is that the denotation I have postulated accounts for several apparently unrelated phenomena in those areas of the grammar of Old English related to demonstrative pronouns, determiners, relative pronouns, nominal and temporal conjunctions, etc. It is thus descriptively more adequate than other accounts of these phenomena. In addition it offers logical arguments with which to describe the historical development of the English conjunction system according to Peinovich's and Kiparsky's principles of simplicity and transparency.

Of course, it may still be argued that we do not need to consider historical data to present a descriptively adequate synchronic account of these phenomena in modern English. This may or may not be true. Yet I believe that the analyses presented by Bresnan, Emonds, Pullum and Wilson have little or nothing to say about the deictic function of the relative pronouns, or about the interpretation of factivity and/or attitudinal involvement in nominal sentences introduced by *that*, or about the occurrence of *that* in temporal and causal embeddings. A diachronic account, however, shows that the syntax of Old English in this area of the grammar was, at least in essence, very similar to that of modern English. It is not by accident that *that* occurs in so many apparently different structures in modern English, since, as I have shown, these structures are not unrelated. It is very profitable indeed to look at *that* from a semantic and from a syntactic point of view in the grammar of Old English. So it should be just as profitable to take the same or a similar approach in examining the present state of the language. The complementizer analysis was assumed to have introduced simplicity into the grammatical description of English. As we have seen with the analysis of *that*, however, simplicity is not everything and may easily lead to confusion elsewhere.

REFERENCES

- Andrew, S. 1966 (1940). *Syntax and style in Old English*. New York: Russell and Russell.
- Bach, E. and R. Harms. (eds). 1968. *Universals in linguistic theory*. New York: Holt, Rinehart and Winston.
- Bresnan, J. 1972. *Theory of complementation in English syntax*. MIT dissertation.
- Campbell, A. 1959. *Old English grammar*. Oxford: Clarendon Press.
- Chomsky, N. 1958. "A transformational approach to syntax". Fodor, J. and J. Katz. (eds). 1964. 211–45.
- Chomsky, N. 1977. "Conditions on rules of grammar". In Cole, R. (ed.). 1977. 3–50.
- Clark, C. (ed.). 1958. *The Peterborough Chronicle 1070–1154*. Oxford/London: Oxford University Press.
- Cole, R. (ed.). 1977. *Current issues in linguistic theory*. Bloomington/London: Indiana University Press.
- Davidson, D. and G. Harman. (eds). 1972. *Semantics of natural language*. Dordrecht: Reidel.
- Emonds, J. 1976. *A transformational approach to English syntax*. New York: Academic Press.
- Fodor, J. and J. Katz. (eds). 1964. *The structure of language: readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice Hall Inc..
- Hoenigswald, H. 1977. "Intentions, assumptions, and contradictions in historical linguistics". In Cole, R. (ed.). 1977. 168–94.
- Huddleston, R. 1976. *An introduction to English transformational syntax*. London: Longman.
- Jacobs, R. and P. Rosenbaum. 1968. *English transformational grammar*. Waltham/Toronto/London: Blaisdell.

- Jacobs, R. and P. Rosenbaum. (eds). 1970. *Readings in English transformational grammar*. Waltham/Toronto/London: Blaisdell.
- Kiparsky, P. 1968. "Linguistic universals and linguistic change". In Bach, E. and R. Harms. (eds). 1968. 170–202.
- Kisbye, T. 1972. *An historical outline of English syntax*. Copenhagen: Akademisk Boghandel.
- Lewis, D. 1972. "General semantics". In Davidson, D. and G. Harman, (eds). 1972. 169–218.
- MacCawley, J. 1970. "Where do noun phrases come from?". In Jacobs, R. and P. Rosenbaum. (eds). 1970. 166–183.
- Palmatier, R. 1969. *A descriptive outline of the "Ormulum"*. The Hague: Mouton.
- Peinovich, M. 1979. *Old English noun morphology*. Dordrecht: North Holland Press.
- Pullum, G. and D. Wilson. 1977. "Autonomous syntax and the analysis of auxiliaries". *Language* 53. 741–88.
- Skeat, W. and K. Sisam. (eds). 1939. *The Lay of Havelok the Dane*. Oxford: Clarendon Press.
- Wagner, K. 1969. *Generative grammatical studies in the Old English language*. Heidelberg: Julius Groos.
- Watts, R. 1976. *Lokative Präpositionen im Deutschen, Englischen und Zürichdeutschen*. Bern: Francke Verlag.