CRISTINA RESMINI[*]

Grammaticalisation paths of the Proto-Indo-European anaphoric pronoun *tód across the Germanic languages

Abstract

Secondo l'etimologia generalmente riconosciuta e nel contesto della grammaticalizzazione, si ritiene che, nel corso del tempo, i riflessi germanici del pronome anaforico protoindoeuropeo *tód abbiano assunto le funzioni di pronome dimostrativo, determinatore, pronome relativo, complementatore argomentale e congiunzione consecutiva. Sulla base dei dati empirici raccolti attraverso un corpus linguistico creato *ad hoc* relativo alle fasi più antiche delle lingue germaniche e alla luce di evidenze cross-linguistiche provenienti specialmente dal latino, questo articolo propone che i riflessi germanici di *tód, che inizialmente fungevano da pronomi dimostrativi, siano diventati determinatori laddove seguiti da un sostantivo e siano stati poi direttamente rianalizzati in complementatori argomentali attraverso la struttura del dittico inverso. L'argomentale, a sua volta, sarebbe stato successivamente rianalizzato per estensione in congiunzione consecutiva in contesti consecutivi e in pronome relativo improprio in presenza di antecedente.

Parole chiave: lingue germaniche, grammaticalizzazione, pronome anaforico protoindoeuropeo, pronome relativo, complementatore argomentale

Based on the accepted etymology and within a grammaticalisation framework, the Germanic reflexes of the Proto-Indo-European anaphoric pronoun *tód are believed to have diachronically acquired the functions of demonstrative pronoun, determiner, relative pronoun, argumental complementiser, and consecutive conjunction. According to the empirical data gathered from an extensive Early Germanic linguistic corpus created specifically for this study and considering cross-linguistic evidence especially from Latin, this paper suggests that the Germanic reflexes of *tód – which initially functioned as demonstrative pronouns – became determiners when followed by a noun and were directly reanalysed into argumental complementisers through a *dipthyque inverse* construction. In turn, the argumental complementiser was then reanalysed by extension into a consecutive conjunction in consecutive contexts, and into a relative *d-Pronomen* in the presence of an antecedent.

Key words: Germanic languages, grammaticalisation, proto-indo-european anaphoric pronoun, relative pronoun, argumental complementiser

[*] Cristina Resmini, Università di Napoli L'Orientale, c.resmini@unior.it.

1. Introduction

This article intends to investigate the grammaticalisation processes which have led the Present-Day Germanic reflexes of the Proto-Indo-European anaphoric pronoun *tód to acquire multiple functions (i.e. demonstrative pronoun, determiner, relative pronoun, argumental complementiser and consecutive conjunction).

To this purpose, this research traces the evolution of the reflexes of *tód throughout the earliest stages of the Germanic languages by means of a linguistic corpus which was created specifically for this study and which includes attestations from all three Germanic branches, dating from the second to the fourteenth century CE.

Since the corpus was meant to serve as the basis for a comparative analysis of the three Germanic branches, the guiding principle in building it was to ensure that each branch would be equally represented. For this reason, the corpus mainly focuses on one single language per branch, i.e. Gothic for the East Germanic branch, Old Norse for the North Germanic branch, and Old English for the West Germanic one. While in the first two cases the selected languages are the only possible option, the choice to examine Old English, instead of any of the other West Germanic languages, is deliberate and it is due to several reasons. Firstly, the Old English written production is much greater than that of any other West Germanic language and it thus represents the only West Germanic production that is comparable in size to that of Gothic and Old Norse. Secondly, most studies in the field seem to have often neglected Old English in favour of Old High German – where some relevant structures are undoubtedly more apparent. However, Old English presents us indeed with interesting dilemmas, as there exist contrasting accounts on a possible early grammaticalisation of relatives – which is exactly the crux of this study. Nevertheless, the linguistic corpus prepared for this research includes a number of relevant entries from other West Germanic languages (i.e. Old High German, Old Low Franconian, Old Frisian and Old Saxon), which appear to be consistent with what can be observed in Old English.

According to the accepted etymology and within the frame of a grammaticalisation process, the Early Germanic reflex of *tód – which was originally a demonstrative pronoun – become a determiner when followed by a noun and

was then supposedly reanalysed into a relative pronoun through the *dipthyque normal* structure (Viti 2013). In the presence of verbs such as *verba dicendi* and *sentiendi*, the relative thus assumedly grammaticalised by extension into an argumental complementiser (Axel-Tober 2017), through the *dipthyque inverse* structure, and then, by over-extension, into a consecutive conjunction, when found within a consecutive context.

While, as far as each individual language is concerned, this study's findings are indeed representative and consistent with the existing literature, the above-mentioned model in its entirety does not, however, apply to the empirical data gathered from the corpus. In fact, based on those data, the argumental and consecutive functions appear to have grammaticalised way before the relative one.

In order to explain such discrepancy, we must thus resort to cross-linguistic evidence (Narrog and Heine 2018). By comparing the Latin correlative structures (De Roberto 2010) with the Germanic ones, we could, in fact, argue that the *dipthyque normal* seems to have only given origin to the Germanic *w-Pronomen*. Following on this hypothesis, this paper then suggests that, on the one hand, the argumental complementiser was grammaticalised directly from the demonstrative, through a *dipthyque inverse* construction, and that, on the other hand, consecutive conjunctions and relative *d-Pronomen* derive from a later extension of the meaning of the argumental complementiser itself.

While the initial sections of this article present a detailed literature review [cf. 2. Theoretical Background] and a description of the linguistic corpus that has been compiled for this study [cf. 3. Methodology], the core of this paper is dedicated to the analysis of the data gathered from the corpus [cf. 4. Corpus Analysis] and to the proposition of a model which, based on that empirical evidence, aims to account for the distinct origins of argumental complementisers and relative pronouns [cf. 5 Discussion]. Finally, following a short summary of the main points this paper tries to make [cf. 6. Conclusion], a selection of the relevant corpus entries mentioned throughout the article is included in the Appendix section.

2. Theoretical Background

In Present-Day English, the word *that* covers multiple functions – including those of demonstrative and relative pronoun, determiner, and conjunction –

and the same can be said, for instance, of *das(s)*[1] and *dat* in, respectively, Present-Day German and Dutch. For this reason, in order to analyse the grammaticalisation paths which have led these morphemes to acquire all of the above-mentioned functions, it is necessary to provide an overview of the way in which the main Present-Day Germanic languages (i.e. English, German, Dutch, Danish, Norwegian, Swedish, and Icelandic) express those functions.

	EN	DE	NL	DA	NO	SV	IS
Demonstrative Pronouns/ Determiners (Proximal, Distal)	*this,* **that**	*der-die-**das**/dieser-diese-dieses/ derjenige-diejenige-dasjenige, der-die-**das** dort/ jener-jene-jenes*	*deze-dit, die-**dat***	*den-**det***	*den-**det**, denne-dette*	*den-**det** här, den-**det** där*	*þessi-þetta, sá-sú-**það***
Definite Articles	*the*	*der-die-**das***	*de-het*	*-en/ -et, den-**det***	*-en/-a/ -et*	*den-**det**, -en/-et*	*-inn/- n/ -ið, hinn*
Relative Pronouns (*w-Pronomen, d-Pronomen*)	*who- what,* **that**	*wie-was, der-die-**das***	*wie- wat, die-**dat***	*hvo- hvad, som, der*	*som*	*som*	*sem, er*
Conjunctions (ARG, CONS)	(**that**), *so* **that**	(**dass**), *so***dass**/ *damit*	(**dat**), *om***dat**/ *zo***dat**	(<u>*at*</u>), *så* <u>*at*</u>	(<u>*at*</u>), *Slik* <u>*at*</u>/ *Så* <u>*at*</u>	(<u>*att*</u>), *Så* <u>*att*</u>	(<u>*að*</u>), *svo* <u>*að*</u>

Tabella 1. Forms per function across the main Germanic languages.

2.1. From pronoun to article
Historical linguists and etymologists alike agree on the fact that the equivalents of *that* (i.e. *das(s)*, *dat*, *det*, and *það* – which are indicated **in bold** in Table 1) are indeed its cognates and that their common ancestor is the neuter form of the Proto-Indo-European anaphoric pronoun **só-, *sá-, *tód* (Sijs 2010).

PIE **tód* > PG **þat* > OE *þæt*, OHG *thaz/daz*, OLF *that*, ON *þat* >
EN *that*, DE *das*, NL *dat*, DA *det*, NO *det*, SV *det*, IS *það*

---

[1] The different distribution of *das* and *dass* is due to a mere spelling difference, probably introduced to avoid ambiguity (Axel-Tober 2017).

According to Quiles and López-Menchero (2011), the Proto-Indo-European *só-, *sá-, *tód* was an anaphoric demonstrative pronoun which tended to appear at the beginning of sentences, as it had probably originated from a relative pronoun. It expressed a *that*-deixis, thus referring to something that was generally *there*, without any specific spatial reference, and could be translated as either *this* or *that.*

As it is evident from Table 1, *that* and its cognates do indeed still act as demonstrative pronouns in all of the Present-Day Germanic languages taken into account and this indicates that the original deictic meaning of **tód* has been retained to this day, as well as its pronominal and anaphoric nature. Since it is therefore clear that the Present-Day Germanic demonstrative pronoun is the immediate reflex of the original **tód*, it seems acceptable to take this specific function as the starting point of this analysis.

Today, though, the Germanic demonstrative pronouns may also be used adjectivally as demonstrative determiners. At the time when Proto-Germanic split from Proto-Indo-European, in fact, the articles system had not yet been introduced in the mother language and, when the difference between indefinite and definite noun was considered to be crucial, the anaphoric pronoun **só-, *sá-, *tód* could also act as a definiteness marker. At that stage of Proto-Indo-European, however, the relationship between the demonstrative and the name to which it referred is assumed to have been of an appositional nature (Quiles and López-Menchero 2011), rather than of an adjectival one – which would otherwise imply that the category of determiners was already in existence. Indeed, it was only with the diachronic emergence of definite articles and the resulting appearance of the category of definiteness that a change in the syntactic structure of the languages occurred and a determiner phrase projection was created (Lyons 1999).

With the emergence of the category of determiners, the Germanic demonstrative pronoun acquired the function of demonstrative determiner, while its role as a definiteness marker was then transferred to the rising definite article, with fairly heterogeneous outcomes among the various languages. On the one hand, German retains the morpheme *das* also for the neuter article, while the English *the* and the Dutch *de* both just represent a different evolution of the original **tód* (Sijs 2010). On the other hand, Norwegian,

Icelandic, Faroese, Danish, and Swedish articles are generally expressed by means of suffixes – although the latter two languages also use the morpheme *det* as the neuter form of the full article, either as an alternative to (Danish) or together with (Swedish) the definite suffix.

The grammaticalisation path which led the anaphoric demonstrative pronoun to become a definite article is considered to be an "innovation," in that an entirely new grammatical category has emerged from this process (Lehmann 2002). At an initial stage, the demonstrative could be used freely (i.e. without context constraints) as an apposition of the head noun by which it was followed and had a [+deictic] feature. However, at some point, this construction was extended to contexts where the deictic reference was no longer relevant – i.e. when the demonstrative was used to indicate that its head noun had already been mentioned. In this specific context, where an anaphoric reference is made, the foregrounded [+anaphoric] feature of the demonstrative acted as the bridging context which brought about an extension of the meaning of the pronoun itself, through a context-induced reinterpretation. Since this new context was incompatible with the original deictic meaning of the demonstrative, though, the latter lost its [-deictic] function and gave rise to the innovated feature of [+definiteness], which is indeed based on anaphoric reference. It is thus during the final phase of the process that the new meaning was conventionalised and the demonstrative could therefore be used freely with that new acceptation, even in other contexts. As already mentioned, some Germanic languages have retained the same word both for the article and for the demonstrative, while, probably to avoid ambiguity, some others – like English and Dutch – have adopted a different and decategorised version of the demonstrative as their definite article. Yet, other languages (i.e. Norwegian, Icelandic, and Faroese) do not seem to have undergone this process at all [cf. Section 4.2].

2.2. From pronoun to subordinator

As far as the grammaticalisation of the anaphoric pronoun is concerned, however, the origin of definite articles is only the first of several differences between North and West Germanic languages. As shown in Table 1, in fact, while in English, German, and Dutch *that*, *das*, and *dat* may also be used –

among other forms – as relatives, in no Scandinavian language do the demonstratives *det* and það cover this specific function. This indicates the existence of two different isoglosses as far as the use of that and its cognates is concerned, but also raises questions about what might have caused the two sub-families to take these two diverging paths.

In the North Germanic languages, the relative function is generally expressed by relative particles [cf. Section 4.3.3], with the sole exception of Danish, which also features the relative pronoun *hvo, hvad*. This pronoun, is a cognate of the English *who, what*, the German *wer, was* and the Dutch *wie, wat*, and they all are reflexes of the original Proto-Germanic adjective interrogative-indefinite pronoun **qos, *qā̄, *qod* (Proto-Indo-European stem **kʷó-*) – which, in turn, had probably originated from an even more ancient Proto-Indo-European relative pronoun (Quiles and López-Menchero 2011).

<div align="center">

PIE **kʷó-* > PG **hʷa-* >
EN *who, what*; DE *wer, wie*; NL *wie, wat*; DA *hvo, hvad*

</div>

Fuß and Grewendorf (2014) refer to these Germanic reflexes as *w-Pronomen*, which usually head free relative clauses and are considered as the only "proper" relative pronouns. On the other hand, they also stress the existence, in the West Germanic languages, of the so-called *d-Pronomen*, which come from the surface-identical demonstrative and usually introduce headed relative clauses [cf. Section 5]. At a closer look, however, the English relative *that* can hardly be considered as a fully fledged pronoun [cf. Section 4.3.2]. Its invariable form, in fact, could lead to it being interpreted as a relative particle (and thus a relative complementiser), even though, from a diachronic perspective, it could also be seen as a "degenerate resumptive pronoun," i.e. a pronoun which has gradually lost its inflectional properties (Haspelmath 2001: 1494-1495).

From this brief account of the Present-Day Germanic relatives system, it is therefore clear that – at least as far as the West Germanic languages are concerned – there are two competing patterns of development in the formation of relatives, where the relative *w-Pronomen* derive from the interrogative-indefinite pronoun and the *d-Pronomen* originate from the anaphoric demonstrative one. For the purposes of this study, however, the most relevant forms are indeed those which result to be identical to the demonstrative.

The intent is, in fact, that of analysing if and how a demonstrative could be grammaticalised into a relative, especially because this pattern seems to be quite specific to this particular sub-group of languages.

Although, traditionally, Proto-Indo-European was assumed to lack subordinate structures and the hypotactic constructions in its daughter languages were believed to have descended from paratactic ones, most scholars nowadays agree on ascribing "at least relative clauses and some types of adverbial clauses" to Proto-Indo-European itself (Viti 2013: 89). The earliest reconstructed complex structure is, however, not yet a form of subordination, but a form of correlation, i.e. the correlative diptych identified by Minard (1936) in Vedic and later found by other linguists in other Indo-European languages, as well. Correlation is a relation between two clauses in which none of the two can exist independently from the other. While this makes it a relation that is unlike either coordination or subordination, Haudry (1979) maintains it might still be considered as the form of relation from which subordination stemmed. The correlative diptych is, in fact, neither an independent, nor an embedded structure (yet), but rather "a subtype of adjoined relative clause" (Hendery 2012: 18), where – in the case of a *dyptique normal* (Minard 1936) – "a preposed relative clause presents a noun phrase, which is anaphorically resumed in the subsequent main clause by a demonstrative pronoun" (Viti 2013: 92).

(1) a. DE *Was* *er* *macht,* *das* *mach-t er* *richtig.*
REL.ACC.N.SG he does DEM.ACC.N.SG does he right.ADV[2]
'What(ever) he does, he does that right'

b. DE *Was* *er* *mach-t,* *mach-t* *er* *richtig.*
REL.ACC.N.SG he does does he right.ADV
'What(ever) he does, he does right'

In particular, as shown in (1a), the *diptyque normal* consists in the adjoining of two simple – but not independent – clauses, the first of which is introduced by an element which belonged to the interrogative-indefinite pronoun-sys-

---

[2] All the glosses throughout the paper follow the Leipzig glossing rules and the information provided is generally restricted to what is specifically relevant for the scopes of this study.

tem (e.g. *was*), while the second one is introduced by an anaphoric demonstrative pronoun (e.g. *das*). However, over time, one of the two elements has been absorbed by the other, thus resulting in sentence (1b). The merger of the two elements – or the omission of one of the two – is what has historically led the correlative diptych to become a hypotactic structure, where the clause "*was er macht*" is embedded in the main clause (i.e. "*macht er richtig*"), as well as what has allowed for the originally interrogative-indefinite pronoun to become a proper relative pronoun.

Although this example manages to show the origin of relative subordinates, it does not yet explain if or how the demonstrative could ever become a relative – and the fact that the demonstrative pronoun is dropped would rather seem to imply it cannot. Let us thus consider the following examples:

(2) a. EN *I have read <u>the book</u> **that** you gave me.*
   b. DE *Ich habe <u>das</u>   <u>Buch</u> gelesen, **das** du   mir   gegeben   hast.*
      I   have ART.ACC.N.SG book read.PP DEM.ACC.N.SG you to.me given   have
      'I have read the book that you have given me'
   c. NL *Ik heb <u>het</u>   <u>boek</u> gelezen **dat** je   me   had   gegeven.*
      I   have ART.N.SG book read.PP DEM.N.SG you to.me have given.PP
      'I have read the book that you had given me'

These three sentences instantiate the use of *that*, *das*, and *dat* as relatives in Present-Day West Germanic languages. In order for those words to begin to be used as such, however, we have to assume a precedent correlative structure as in (3a), (3b) and (3c):

(3) a. EN **I have read <u>the book</u> **that** <u>which</u> you gave me.*
   b. DE **Ich habe <u>das</u>   <u>Buch</u> gelesen, **das**   <u>was</u>*
      I   have ART.ACC.N.SG book read.PP DEM.ACC.N.SG REL.ACC.N.SG
      du   mir   gegeben   hast.
      you   to.me given.PP   have
      **'I have read the book that which you have given me'
   c. NL **Ik heb <u>het</u>   <u>boek</u> gelezen **dat**   <u>wat</u>   je   me*
      I   have ART.N.SG book read.PP DEM.N.SG REL.N.SG you to.me
      had   gegeven.
      have   given.PP
      **'I have read the book that which you had given me'

In these cases, as the Present-Day outcomes show (2a), (2b), and (2c), the resumptive pronoun has prevailed over the proper relative pronoun – which has consequently been dropped – and the anaphoric demonstrative pronoun has become a relative pronoun in its own right.

Nevertheless, since English, German, and Dutch all retain both the *w-Pronomen* and the *d-Pronomen* forms of the relative we must consider this grammaticalisation process as a "renovation" (Lehmann 2002) that is only partly innovative, as indeed the new relative (the *d-Pronomen*) has never completely replaced the original proper relative. Furthermore, while in German and Dutch there is a complementary distribution among the two pronouns, in English, the relative complementiser *that* can always be replaced by a *w-Pronomen,* meaning that it is in an overlapping distribution with pronouns like *who* and *which* (Zimmermann 2012).

Beside the relative, however, there are two more grammaticalisation processes among the West Germanic languages in which the words *that*, *das*, and *dat* have been involved and these are the paths which have allowed for them to also acquire the functions of argumental complementiser and consecutive conjunction.

In historical linguistics studies, the origin and development of the argumental clause is far from being uncontroversial. From the Neogrammarians period and up to few years ago, the argumental complementiser in the West Germanic languages was traditionally believed to have derived directly from the Proto-Indo-European anaphoric pronoun **tód*. The anaphoric pronoun was, in fact, assumed to have been used cataphorically, when found within a paratactic structure and in the presence of *verba dicendi* and *verba sentiendi*.

(4)   a.   EN   *I know <u>that</u> you are wrong.*
      b.   EN   **I know <u>this</u>: **that** you are wrong.*

In example (4a), for instance, the presumed positioning of the clause boundary would be between *that* and *you* (i.e. "I know <u>that</u> | you are wrong"), and the demonstrative would thus indeed constitute a cataphoric reference to the following clause (i.e. "you are wrong"). Therefore, based on this traditional account, the process of reanalysis of the demonstrative into an argumental complementiser would consist in the cataphoric demonstrative moving from

the end of the first clause to the beginning of the second one, through a shift of clause boundary, and in the paratactic structure consequently developing into a hypotactic – albeit surface-identical – construction.

In recent years, however, this hypothesis has been abandoned in favour of a new one which sees the argumental complementiser as originating from the relative, instead. As most subordinates, in fact, the argumental clause is cross-linguistically based on the relative one (Haudry 1979) and it seems only fitting that the complementiser introducing it should have originated from some form of relative pronoun or complementiser. As further confirmation of that, it is believed that the argumental clause had derived from a correlative structure, i.e. the *diptyque inverse* (Minard 1936), where – contrarily to the *diptyque normal* – "the head noun is contained in the main clause, which precedes the relative clause" (Viti 2013: 92), or, in this case, the argumental one. This correlative structure thus consists of two adjoined clauses: the first one presents a demonstrative pronoun used cataphorically, while the second one is introduced by a relative (4b) – which can either be a relative complementiser as in English, or the neuter form of the relative pronoun, as in German and Dutch.

According to this second hypothesis, which has been especially put forward by Axel-Tober (2017), the argumental complementiser is indeed assumed to derive from a correlative structure where one element is the optionally silent demonstrative and the other element is the already grammaticalised relative *d-Pronomen*. In this case, the demonstrative could easily be dropped in order to simplify the structure and the *d-Pronomen* would thus prevail with no shift of clause boundary because it has always belonged to that second clause.

Therefore, we may maintain that, when in the presence of some specific categories of verbs [cf. Section 4.4] the relative *d-Pronomen* has been reanalysed as an argumental complementiser, through a context-induced "extension" of its meaning.

On the other hand, the grammaticalisation path which has seen *that* and its West Germanic cognates take on the function of consecutive conjunction is an "overextension," in that it represents the extension in meaning of an already extended element, such as the argumental complementiser. When in

the specific context of the consecutive construction (e.g. *so*+ADJ/ADV...*that*), in fact, the argumental complementiser is assumed to have undergone a "semantic bleaching" in which it has lost its semantic function and acquired a merely syntactic one.

As for the argumental complementiser and consecutive conjunction in the North Germanic languages, the forms reported in Table 1 are *at* (Danish, Norwegian), *att* (Swedish), and *að* (Icelandic). Although these morphemes are formally different from the Scandinavian demonstratives and even from the West Germanic *that*, *das*, and *dat*, a reconstruction which sees them as just different reflexes of the same Proto-Indo-European pronoun **tód* is not only plausible, but quite popular among scholars (Harbert 2007; Bisang and Malchukov 2020).

<div align="center">

PIE **tód* > PG **þat* > ON þat/at >
Demonstratives: DA *det*, NO *det*, SV *det*, IS *það* / Conjunctions:
DA *at*, NO *at*, SV *att*, IS *að*

</div>

Since, in fact, we would only need to postulate the dropping of the first consonant (Vigfusson 1874:28-29) for this development to be acceptable, we cannot exclude this possibility just on grounds of reconstruction. In this scenario, *at*, *att*, and *að* are cognates of *that*, *dass*, and *dat* and, although they seem to have mostly skipped the other grammaticalisation processes (i.e. demonstrative > definite article; demonstrative > relative[3]), they have been reanalysed into argumental complementiser and consecutive conjunction [cf. Section 4.4.3][4].

On the flip side, though, there seems to exist the possibility that *at*, *att*, and *að* all derive from the above-mentioned homographic preposition and conjunction, which, according to Vigfusson (1874:25-28), is a cognate of the

---

[3] Interestingly, though, while this is no longer true in Present-Day Icelandic, Vigfusson (1874: 29) mentions that the now archaic *at* (now replaced by *að*) could also be used as a less frequent alternative of the relative particle *er*, thus implying that the grammaticalisation of the demonstrative into a relative might indeed have happened, although no trace of it is left in Present-Day Icelandic.

[4] If this hypothesis were true, we would also have to assume a spelling difference between the demonstrative and the conjunction (not unlike that of the German *das/dass*), and to accept the existence in Present-Day North Germanic languages of a homographic preposition and conjunction.

English preposition *at* and has originated from the Proto-Indo-European preposition *ad*.

<div align="center">

PIE **ad* > PG **at* > ON *at* > DA *at*, NO *at*, SV *att*, IS *að*

</div>

The preposition *ad*, which meant *to*, *by* and *near*, was then reanalysed as a subordinating conjunction (translatable as the English conjunction *to*) and, if this second hypothesis could hold, further extended to the functions of argumental complementiser and consecutive conjunction.

In the first case, North and West Germanic languages would all belong to a single isogloss, as far as most of the grammaticalisation paths of the Proto-Indo-European **tód* are concerned. In the second case, on the other hand, we would be in the presence of two very distinct isoglosses among sister languages featuring cognate demonstratives. Although traditionally the first hypothesis seems to be the most accredited one, it is still interesting to analyse the validity of the second one, as it would appear to be plausible, not only from a linguistic point of view, but also from a historical one. Indeed, the discriminating factor in the matter is presumed to be the influence of Latin syntax on the Germanic languages, which has been of a major significance among the West Germanic languages, but virtually non-existent on the North Germanic ones.

2.3. Grammaticalisation and typology

This concept, however, raises the long-standing question of what could be the triggering factors which may lead to the grammaticalisation of some specific elements of a language. According to Heine and Kuteva (2010:380), grammaticalisation is the unidirectional process which allows for lexical forms to develop into grammatical forms, and for grammatical forms to evolve into even more grammatical ones. Assuming that every kind of linguistic evolution is the result of language change, when it comes to grammaticalisation, we have to consider that it is mainly motivated by the need to use "linguistic forms that are concrete, easily accessible, and/or clearly delineated to also express less concrete, less easily accessible, and less clearly delineated meaning contents" (Heine and Kuteva 2010: 381). In order for these processes to be set into motion, though, we need to also factor in a certain amount of creativity, "whereby

existing material is combined in novel ways to produce new artefacts" – i.e. new linguistic meaning and structures (Heine and Kuteva 2010: 381).

There are several accounts on how grammaticalisation processes may occur within a language, but most scholars have favoured either a transformational-generative approach or a typological perspective as the theoretical framework for their studies.

The transformational-generative approach is especially grounded in Chomsky's notions of "deep structure" and "surface structures" as the two coexisting levels of representation within a language. The first one contains all the information relevant to semantic interpretation, while the latter contains all the information relevant to phonetic interpretation alone (Chomsky 1988: 14-15). According to this approach, syntactic change is generally associated with child language acquisition, in the sense that it is seen as part of what happens in the transition of grammars from one generation to the next[5] (Campbell 1998:234-235). The child learners, in fact, are presumed to construct their own grammar on the basis of the input they receive from the adults around them. However, learners' acquisition of grammar will be limited to the surface structure of the input they have received and processed, and there is thus no guarantee that the grammar they have acquired does coincide with the deep structure of the adults' grammar. Furthermore, language speakers may at some point add rules to their grammar which would make it no longer optimal[6] and, it will thus be up to the children of the following generation to restructure the non-optimal adult grammar and make it more optimal, as they construct their own internal grammar. Indeed, Kroch (2001: 2) argues that language change is "by definition a failure in the transmission across time of linguistic features," while Roberts and Roussou (2007) claim that syntactic change can be interpreted as parametric variation. Finally, in more recent works, Biberauer & Roberts (2017: 134) attempt to integrate "the concerns of historical syntax with those of generative grammar" and

---

[5] According to Heine and Kuteva (2010: 382), "a *sine qua non* for grammaticalisation is the transmission of linguistic knowledge from one generation to the next."

[6] Optimality occurs when natural language grammars create structures which are designed to interface perfectly with other components of the mind, i.e. speech and though systems (Radford 2004: 9).

propose that the parameters of Universal Grammar are not pre-specified, as they, instead, supposedly emerge from the interaction subcounsciously operated by the acquierer between the three factors of language design, i.e. the Universal Grammar, the primary linguistic data and the non-domain-specific cognitive optimisation principles.

Although such descriptions of how syntactic change occurs are certainly valid ones, they cannot provide a satisfactory answer as to why certain elements of a specific language do indeed grammaticalise whereas others do not, nor as to how unrelated languages may take similar paths, while related ones behave differently.

It thus appears evident that there must be some other factor at play, besides the ones we have already mentioned, and that is language typology. The diachronically typological approach is in fact the one that has been adopted for this study.

Typology is "the study of structural differences and similarities between languages" (Velupillai 2012: 15) and, while synchronic typology aims at listing and categorising all possible patterns, diachronic typology attempts to explain why such patterns exist and why they change (Pat-El 2021). According to Bybee (2006), some universal mechanisms lead, in fact, to pathways of change which then reflect on synchronic linguistic structures and, since these patterns are often cross-linguistically similar, their synchronic outcomes may indeed bear some resemblance. In particular, the typological features of a language are expected to correlate with some aspects of grammaticalisation, in that "certain criteria of grammaticalisation may apply differently to different types of language" and this correlation does not appear to be unidirectional, as it is argued both that "typological features influence aspects of grammaticalisation" and that "grammaticalisation motivates structural features that can be typologised" (Narrog and Heine 2018: 1-2). Therefore, whereas generative approaches tend to reduce grammaticalisation "to an essence that is universal and not amenable to typological influences" (Narrog and Heine 2018: 4), from a typological perspective, it is argued that "the likelihood for a certain grammaticalisation process to appear is at least to some extent dependent on structural properties of the language" (Dahl 2018: 95). Furthermore, Narrog and Heine (2018: 14-15) main-

tain that there exists a general tendency in languages to "follow already trodden grammaticalisation paths and to reproduce or flesh out established grammatical categories [...] rather than to create entirely new structures and categories," while they also claim that when languages diverge from this "tendency towards conservatism, or inertia," it is most likely due to intense language contact.

When it comes to similarities in development between languages, however, typology, tends to reject the hypothesis of language contact and prefers to explain them as parallel independent developments due to "similarities in preconditions – either internal, i.e. shared structural properties, or external – shared ecologies, or universal cognitive properties" (Dahl 2018: 79).

As descendents of a common ancestor, in fact, the Germanic languages are expected to bear varying degrees of resemblance to all of the other Indo-European languages, even though they belong to different branches and have not had direct or significant contact with them after they split from the mother language. Moreover, towards the end of the twentieth century, some scholars identified an European linguistic area, i.e. the so-called "Standard Average European", or SAE, which mainly consists of the Romance, Germanic, Balto-Slavic, and Balkan languages, as well as the non-Indo-European westernmost Finno-Ugrian languages. As a linguistic area, the Standard Average European comprises a number of geographically contiguous languages that "share structural features which cannot be due to retention from a common proto-language and which give these languages a profile that makes them stand out among the surrounding languages" (Haspelmath 2001: 1492).

Following on these concepts, the shared properties within either a language family or a language area seem to have a major role also as far as grammaticalisation is concerned. For this reason, it is necessary to broaden the scopes of this study and further analyse the already discussed grammaticalisation paths, from an Indo-European perspective.

As mentioned earlier, articles are typically considered as a late Proto-Indo-European feature which only appeared after the daughter languages had already started to separate from the mother language. This explains why Latin and the early Homeric Greek lacked definite articles, while later forms of

Ancient Greek, the Romance languages, and even the Present-Day Germanic languages do not. However, these languages belong to different Indo-European branches and they have introduced the articles system at very different times. Interestingly, however, the grammaticalisation path which has brought them to acquire the definite article is exactly the same as the one we have already described and started with the language-specific reflex of one of the Proto-Indo-European demonstratives.

PIE *tód* > GR τό (ART.NOM.N.SG)
LAT *ille* (DEM.NOM.M.SG) > IT *il* (ART.M.SG)

Besides *só-, *sá-, *tód*, in fact, Proto-Indo-European also featured a second anaphoric demonstrative pronoun, i.e. *is*, *ieh$_2$-*, *id* – from which, for instance, the Latin pronoun *is*, *ea*, *id* originated and which, in turn, had derived from an even older demonstrative of stem *i-* (Quiles and López-Menchero 2011). However, according to Quiles and López-Menchero (2011), in Proto-indo-European, there was another pronoun which had stemmed from the anaphoric stem *i-*, i.e. the original relative pronoun *ios, *iā, *iod*, from which many Indo-European languages, including Indo-Iranian, Greek, Phrygian and Slavic, have derived their relative pronoun. While *ios, *iā, *iod* usually introduced appositive-explicative relative clauses, however, there also existed a second relative pronoun, which introduced attributive-restrictive clauses and came from the Proto-Indo-European interrogative-indefinite pronouns *qis*, *qid* (substantive form) and *qos*, *qā, *qod* (adjective form) – from which, as already mentioned, the Present-Day Germanic *w-Pronomen* derive. From *qis*, *qid*, on the other hand, the Latin *quis*, *quid* originated, which then produced reflexes in the Romance languages, such as the Italian *chi*, *che (cosa)* and the French *qui*, *que*.

As we can thus see from this account, Proto-Indo-European itself is assumed to have featured a complementary distribution between two relative pronouns which had originated from competing patterns, i.e. one from a demonstrative stem and the other from an interrogative-indefinite one. Although in Present-Day SAE languages, relatives based on a demonstrative are not common (Haspelmath 2001:1494), from a typological perspective, we may justify the fact that (West) Germanic demonstratives are assumed

to have been reanalysed into relatives, on grounds of the apparently similar origin of one of the Proto-Indo-European relatives.

As for the formation of relative, argumental, and consecutive clauses, the plausibility of the grammaticalisation paths which we have described for the (West) Germanic languages is corroborated by cross-linguistic evidence. In fact, complex sentences in Latin are believed to have originated from correlative structures, where the merger of two elements into one gave rise to different types of subordinate clauses and the merged element became a subordinator.

More specifically, according to De Roberto (2010), relative clauses in Latin derived from the *diptyque normal* structure, where the pronoun *is* had an anaphoric function and was, over time, absorbed by the interrogative-indefinite pronoun *qui*, which was thus grammaticalised into a relative pronoun (5).

(5) LAT ***Qui***   *bene* *amat*, (*is*)   *bene*   *castigat.*
          REL.NOM.M.SG well   loves   PERS.NOM.3SG   well   chastises
          'Who well loves, he well chastises'

                                                                        (De Roberto 2010)

Following on this explanation, when in the presence of *verba dicendi* and *verba sentiendi*, the already grammaticalised relative pronoun was also reanalysed into an argumental complementiser through the *diptyque inverse* structure (De Roberto 2010), where the demonstrative element was used cataphorically and, once again, dropped and absorbed by the relative (6). Moreover, just as in the (West) Germanic languages, the grammaticalised argumental complementiser was reanalysed into a consecutive conjunction, by "overextension."

(6) LAT ***Qui***   *bene* *amat*, (*is*)   *bene*   *castigat.*
          REL.NOM.M.SG well   loves   PERS.NOM.3SG   well   chastises
          'Who well loves, he well chastises'

                                                                        (De Roberto 2010)

(6) LAT *Legati*   *renuntiaverunt* *hoc,*   ***quod***   *Pompeium in*
          envoys   reported   DEM.ACC.N.SG REL.ACC.N.SG Pompeius in
          potestate   *haberent.*
          power   had.3PL
          'The envoys reported this: that they had Pompeius in their power'

                                                                        (De Roberto 2010)

Indeed, it thus appears that the only difference between the Germanic and the Italic grammaticalisation of relative pronouns, argumental complementisers and consecutive conjunctions, is the element they have selected to undergo those paths – in the first case, the demonstrative prevailed, whereas, in the second one, the interrogative-indefinite was favoured. Although it remains uncertain (while still acceptable, as discussed above) why that would be the case for the Germanic languages, in typological accordance with the grammaticalisation paths undergone by the Italic relative pronoun, it is plausible to assume that the element being reanalysed into an argumental complementiser in the Germanic languages was indeed the relative, and not the demonstrative[7].

While, at this point, the general outline of the above-described grammaticalisation paths seems to be – at least on principle – fairly acceptable, the aim of this study is that of trying to shed some light on the multiple controversies pertaining to this topic.

3. Methodology

The linguistic corpus which has been compiled specifically for this study consists of about eight hundred sentences from the earliest attested stages of the main Germanic language. These sentences were mostly gathered from three different sources: the *Early Indo-European Online* repository by the University of Texas at Austin, the *RuneS* website, and the handbook of Germanic Philology *Old English and its Closest Relatives* by Orrin W. Robinson.

The prime criterion employed to select the relevant sentences was the presence within each one of them of at least one element which, according to the Present-Day English translation provided, could be rendered as *that*. At a later stage, the collected entries were further selected, by only keeping those which covered one of the functions that have been instantiated so far (i.e. demonstrative pronoun, demonstrative determiner, definite article, relative pronoun, argumental complementiser, and consecutive conjunction). Any token which, despite its function, did not have a relevant origin (e.g. Old

---

[7] This hypothesis does, however, still need testing.

English *swaswa* is a consecutive conjunction meaning *so that*, but it clearly does not derive from any form of the Proto-Indo-European *tód*) was disregarded.

For what concerns the selection of the languages to analyse, as already discussed [cf. Section 1], the cardinal principle was to provide a balanced account of each Germanic branch, in order to make sure any statistic result would be significant. Beside Gothic (East Germanic branch) and Old Norse (North Germanic branch), Old English was thus chosen among the West Germanic languages as it is the most represented one. Furthermore, Old English happens to pose major issues especially as far as the grammaticalisation of relatives is concerned, but has not been investigated as much as other languages, like Old High German, have. Nonetheless, some runic inscriptions and a handful of illustrative sentences from other West Germanic languages (i.e. Old High German, Old Low Franconian, Old Frisian, Old Saxon, and Middle English) are also included in the corpus.

From a chronological point of view, the oldest dated entries in the corpus are taken from the few surviving runic inscriptions in Older Fuþark, which date from the second half of the second century CE to the end of the seventh century. The corpus also features some interesting Younger Fuþark (700-1325) and Fuþork (700-1100) inscriptions. While it is debated whether Older Fuþark should be defined as Proto-Norse or as a form of Common-Germanic (Robinson 1992), it is clear that Younger Fuþark inscriptions are purely Scandinavian, whereas Fuþork is considered as a form of Old English. Notably, the oldest runic inscriptions typically consist of one or two descriptive words and serve the purposes of "naming objects, naming the owners of objects, naming the carver of the runes, naming the person memorialised, or some combination of the above" (Robinson 1992: 95-96).

As for non-runic texts, the corpus covers quite an extensive period of time (roughly from the fourth century to the late fourteenth century) and a wide range of text genres. In particular, as far as Gothic is concerned, the passages which have been inserted in the corpus are taken from the Gospels in Wulfila's Bible (fourth century) and from the *Skeirins* – a fifth century commentary on the Gospel of Mark. Religious content may however also be

found in some West Germanic entries, and especially in those taken from the following texts:

- An Old English poetic paraphrase of the *Genesis* written by an English monk around 1000;
- The Old Saxon alliterative epic poem *Heliand* (first half of the ninth century), where an account of the life of Jesus is provided, through the combination of all four Gospels;
- The ninth century Old High German eschatological poem *Muspilli*;
- The word-for-word interlinear translations from Latin of the Gospels and the Book of Psalms in, respectively, Old High German and Old Low Franconian[8], for which no certain date can be provided;
- The Middle English historical and religious poem *Cursor Mundi* (1300).

With respect to religious texts in Old Norse, on the other hand, we are mainly referring to the myths of the Germanic heathendom which may be found in some excerpts of a 1270s manuscript of Eddic poetry which is included in the corpus, as well as in some passages from the Prologue of Snorri Sturluson's (1179-1241) prosaic *Edda*.

Besides religious content, however, a significant number of the entries in the corpus are taken from historical accounts, such as the Old English *Anglo-Saxon Chronicle* (years 755, 897, and 991) and their most recent realisation, i.e. the *Peterborough Chronicle* (year 1087) – of which some passages written in the Middle English period (year 1137) are also included. Similar accounts also existed in Old Norse literature and some examples may be found in sentences taken from an account of the Battle of Stamford Bridge in 1066 and in the Íslendingabók, or the *Book of Icelanders*, which was written by the priest Ari Þorgilsson (1067-1148) and dealt with the early history of Iceland.

Among Old Norse literary texts, we may also find the so-called "family sagas," which usually narrate the story of the families that first settled in

---

[8] While some scholars claim that this language may be called "Old Dutch," Robinson (1992: 205) argues that it rather is "the oldest stage of the Limburgic dialect of Dutch, which differs substantially from the standard language." For this reason, it is thus preferable to define this early phase as "Old Low Franconian."

Iceland. This genre is represented in the corpus through passages from the *Brennu-Njáls Saga* (composed around 1250-127 by an unknown author) and from a fourteenth century manuscript of the *Egils Saga*, whose author might again be Snorri Sturluson. In addition to family sagas, however, the corpus also contains some excerpts from "legendary sagas," like the *Grettis Saga* and the *Hrólfs saga* (both taken from fourteenth century manuscripts), which tell stories of heroes and warriors and whose genre greatly resembles that of epic poems such as the Old English *Beowulf*, of which some passages taken from a 1000 lone manuscript are provided.

The corpus also includes entries from two Old English elegiac poems from the tenth century *Exeter Book* (i.e. *The Seafarer* and *The Wanderer*), a lecture on the poet Caemdon by the Venerable Bede (671-735), and King Alfred's (c. 848-899) account of *The Voyages of Ohthere and Wulfstan*, as well as some sentences taken from more recent sources, like the Old Frisian 1300 *Asegabook* (i.e. the legal code of the Frisians inhabiting the Rustringian region of Frisia) and Geoffrey Chaucher's *Tales of Canterbury*, from the Middle English period (1387).

Once all the relevant sentences had been collected, the tokens that were translatable as *that* were analysed by part of speech and by the syntactic function they covered within the context of the sentence. Based on the syntactic function of the entries, the corpus was thus divided into four separate tables[9]: demonstratives, articles, relatives and conjunctions. All of the entries featured in these tables were ordered chronologically, based on the assumed dating of the manuscript from which they were taken, and a number of variables were included so as to provide a more detailed analysis of the tokens in question.

As it is evident from the selection included in the Appendix section, each entry is associated with a progressive number, in order to make any reference throughout the paper more accessible. The first few columns indicate the language and text from which the sentence was taken, as well as its dating. The last two columns, on the other hand, feature the translation of the sentence and the source from which the entry itself was collected. Although,

---

[9] A selection of each of the four resulting tables is presented in the Appendix section of this paper.

more often than not, the translation was already provided in the consulted repositories, adjustments were made when a more literal translation was needed to render the structure of the original sentence more apparent. Finally, the primary sources that had been consulted to compile the corpus [cf. Primary sources in Bibliography] were associated with a letter and the relevant letter for each of the entries was thus inserted in the "Source" column.

However, there are some differences among the tables which are worth analysing separately. The "Demonstrative" table indicates the form from which the token came (e.g. *sa* comes from *sa, so, þata*), its semantic agreement by case, gender, and number and, finally, its syntactic function – i.e. demonstrative pronoun, demonstrative determiner, or person pronoun. In this table, the "Part of Speech" column is not included, as it was the same (i.e. demonstrative pronoun) for all of the entries.

Similarly, since the syntactic function was the same for all the entries in both the "Article" and "Relative" tables, only the origin of the token and the semantic agreement are indicated.

In conclusion, the "Conjunction" table includes both the syntactic function of argumental complementiser and that of consecutive conjunctions, and the entries are therefore analysed by origin of the token and syntactic function. While they were all uninflected forms and the semantic agreement was thus irrelevant, the verb by which the argumental clause is introduced was highlighted, in the case of argumental complementisers, as well as the possible presence of a consecutive structure (e.g. *so*+ADV/ADJ *that*), in the case of a consecutive conjunction.

4. Corpus analysis

4.1. Demonstratives

As already discussed, the Proto-Indo-European anaphoric pronoun **só-, *sá-, *tód* is the common ancestor of at least some of the demonstrative pronouns and determiners which we find in the Present-Day Germanic languages. As shown in Table 2, this is true for the earliest stages of the Germanic languages, as well. Although, in fact, the Gothic form **jáins** and the Old Norse *hann* and *hinn* do not derive from **só-, *sá-, *tód*, most of the demonstratives represented in the corpus are indeed recognised as its reflexes (e.g. **in bold**,

the Gothic *sa*, *so*, *þata*, the Old English *sē*, *sēo*, *ðæt* and the Old Norse *sá* and *sjá*).

Forms	Gothic			Old English		Old Norse				
	sa, so, þata	*jáins*	Tot.	*sē, sēo, ðæt*	Tot.	*hann*	*hinn*	*sá*	*sjá*	Tot.
No. Instances	35	3	38	28	28	55	2	135	1	198
Demonstrative Pronoun	13	0	13	12	12	8	2	82	1	93
Demonstrative Determiner	8	2	10	12	12	1	0	11	0	17
Third-Person Pronoun	14	1	15	4	4	46	0	42	0	88

Table 2. Distribution of deictics within the corpus.[10]

As it is evident from the table, in all of the Germanic branches, the demonstratives tended to function not only as deictic pronouns, which could then be used both cataphorically (7) and anaphorically (8), but also as demonstrative determiners (9) and third-person pronouns (10). While in most Present-Day Germanic languages this latter use has been dropped, the determiner function has been preserved and even consolidated.

(7) GOT **Sa** *ist* *sunus* *meins* *sa* *liuba* *[...].*
            this is   son   my   the   beloved[11]
        'This is my son, the beloved.'

                                              [cf. Appendix, no. 3]

(8) GOT *Jah* *þatei* *gasaƕ* *jag* *gahausida* **þata** *weitwodeiþ.*
            and   what   saw.3SG and   heard   that   testify.3SG
        'And what he saw and heard, that he testifies.'

                                              [cf. Appendix, no. 4]

[10] In line with the choices that have been made regarding the corpus, Table 2, as well as most of the tables included in this paper, only features instances from one language per Germanic branch (i.e. Gothic, Old English and Old Norse). Although, as already mentioned in Section 3, the intent behind this decision was to provide an even representation of each branch, this has not always been the case, since the linguistic corpora of the different languages are obviously not equivalent.

[11] Relevant elements are often not glossed for morpheme, as the relative information can be found either in the textual explanation of the examples or in the provided tables.

(9) ON *Fyrir* **þá**   <u>sǫk</u> *[...].*
        for   that reason
        'For that reason.'

[cf. Appendix, no. 8]

(10) EN *Mid*   **þǣm** *þæt* *manige [...]* *forðferdon.*
         among   them (the) many   died
         'Among them, many [...] died'

[cf. Appendix, no. 6]

The aim of this section is therefore to provide an overview of how the demonstratives are represented within the corpus, while analysing each of the three Germanic branches separately and making some considerations on the evolution of the demonstratives systems in the different languages.

4.1.1. East Germanic

Table 3 accounts for the use of demonstratives in Gothic, in that each lemma instantiated in the demonstratives part of the corpus is analysed by case, number and gender and the number of occurrences per syntactic function is registered. The table also features information on the texts where such function was featured and the dating of those texts.

	Inflection			Demonstrative Pronoun		Demonstrative Determiner		Third-Person Pronoun		
				No.	Text, dating	No.	Text, dating	No.	Text, dating	**Tot.**
ijos	NOM	–	PL	0	–	0	–	1	*Gospels*, 4th cent.	1
jainaim[12]	DAT	M	PL	0	–	~~1~~	*Gospels*, 4th cent.	~~1~~	*Gospels*, 4th cent.	~~2~~
jainans[13]	ACC	M	PL	0	–	~~1~~	*Gospels*, 4th cent.	0	–	~~1~~

[12] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 3.

[13] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations is therefore not relevant for the purposes of this study and it is not included in the sum total reported in the last row of Table 3.

	CASE									
sa	NOM	M	SG	2	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	0	-	3
soh	NOM	F	SG	0	-	1	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	2
þai	NOM	M	PL	1	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	3
þaim	DAT	M	PL	1	*Gospels*, 4th cent.	0	-	5	*Gospels*, 4th cent.	7
		N		1	*Gospels*, 4th cent.	0	-	0	-	
þamma	DAT	M	SG	0	-	1	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	2
þata	NOM	N	SG	1	*Gospels*, 4th cent.	0	-	1	*Gospels*, 4th cent.	11
	ACC	N	SG	3	*Gospels*, 4th cent.; *Skeireins*, 5th cent	2	*Gospels*, 4th cent.	3	*Gospels*, 4th cent.	
þat(uh)				1	*Gospels*, 4th cent.	0	-	0	-	
þis	GEN	N	SG	1	*Skeireins*, 5th cent.	0	-	0	-	1
þize	GEN	N	PL	1	*Gospels*, 4th cent.	0	-	0	-	1
þo	NOM	N	SG	1	*Skeireins*, 5th cent.	0	-	0	-	4
		N	PL	0	-	0	-	1	*Gospels*, 4th cent.	
	ACC	F	SG	0	-	1	*Skeireins*, 5th cent.	0	-	
		N	PL	0	-	1	*Gospels*, 4th cent.	0	-	
Tot.				**13**		**8**		**14**		**35**

Table 3. Distribution of Gothic deictics in the corpus.

The table provides a first confirmation of what has been discussed so far, in that – as expected – most of the instantiated lemmas are but inflected forms of the Gothic demonstrative *sa*, *sō*, **þata**, which originated from the Proto-In-

do-European **só-, *sá-, *tód* and, just like it, expressed a *that*-deixis. Interestingly, the only exceptions are the forms *jainans* and *jainaim* which represent the distal-deixis inflected forms of *jáins*, *jaina*, *jainata* and both derive from the Proto-Germanic **jainaz*, **jenaz* – just as the Present-Day German distal demonstrative *jener*, *jene*, *jenes* does (Keretchashvili 2021). The very fact that this form exists in German and is in a complementary distribution with the proximal demonstrative *der*, *die*, *das* (i.e. a reflex of **só-, *sá-, *tód*) seems to indicate that the demonstratives deriving from *jáins*, *jaina*, *jainata* and those deriving from *sa*, *sō*, *þata* might have been in a complementary distribution in Gothic, as well.

If this were the case, we could assume that the Gothic demonstrative forms deriving from **só-, *sá-, *tód* were already fully grammaticalised as demonstrative determiners as early as in the fourth century. In fact, as results from Table 3, the inflected forms of *sa*, *sō*, *þata* function as demonstrative pronouns – used both cataphorically (7) and anaphorically (8) – as well as as demonstrative determiners, when in adjectival position. Moreover, the Gothic *sa*, *sō*, *þata* also appears to have largely been used as a third-person pronoun.

There does not appear to be a pattern by which certain forms alone might take on specific functions, while others could not. While, within the pool of sentences analysed in the corpus, some forms do not feature all of the functions taken into consideration, it is reasonable to assume that this is just imputable to the fact that the entries gathered for the corpus were too few and/or – having been randomly chosen – not representative enough.

Despite its being a closed-corpus language, Gothic still represents the optimal point of departure for this research as it is the earliest attested Germanic language. On the one hand, we can, in fact, assume a closer resemblance with the non-attested Proto-Germanic language and expect Gothic to thus provide us with invaluable insight on those earlier stages of the Germanic languages. On the other hand, however, we need to bear in mind that all the Gothic texts we have today are translations from Koine Greek and that, on these grounds, the written form of the Gothic language has been largely influenced and even shaped by it.

Nevertheless, seeing how Koine Greek demonstratives worked further confirms the patterns shown in the corpus. In Greek, in fact, the grammat-

icalisation of demonstratives into determiners had occurred already in the Classical period and was preserved throughout the Koine Greek time so that the demonstratives could be used indifferently as pronouns or adjectives.

While a third-person pronoun existed in Greek, however, it was rarely used and often replaced by the oblique cases of the demonstrative αὐτός, αὐτή, αὐτό (Rydberg-Cox 2021). This is not unlike what seems to have happened in Gothic, where – despite the existence of the "proper" third-person pronoun *is, eis, ita* (Keretchashvili 2021) – the demonstrative *sa, sō,* **þata** is often preferred.

4.1.2. West Germanic

With regards to the Old Enlgish demonstratives, as seen earlier on in this chapter, all of the instances of demonstratives in the corpus are inflected forms of *sē, sēo,* ðæt – i.e. an Old English reflex of the Proto-Indo-European **só-, *sá-, *tód.* In Old English sē, *sēo,* ðæt used to express a that-deixis, while today, after the dropping of the distal pronoun *yon* (from the Proto-Germanic **jainaz, jenaz*), its descendent *that, those* has taken over the distal-deixis.

	Inflection			Demonstrative Pronoun		Demonstrative Determiner		Third-Person Pronoun		Tot.
				No	Texts, Dating	No	Texts, Dating	No	Texts, Dating	
ðæt	ACC	N	SG	0	-	1	*The Voyages of Ohthere and Wulfstan* (848-899)	0	-	1
þā	ACC	-	PL	0	-	2	The Venerable Bede (671-735)	0	-	2
þa	ACC	F	SG	1	*Genesis*, 1000	1	*Genesis*, 1000	0	-	2
þām	DAT	M	SG	0	-	0	-	1	*Exeter Book*, 10th cent.	1
þam	DAT	-	PL	0	-	2	Peterborough Chronicle, 1087	0	-	2
þāra	GEN	-	PL	1	*Anglo-Saxon Chronicle* (y. 897)	0	-	1	The Venerable Bede (671-735)	2

þǣm	DAT	-	PL	0	-	1	*Anglo-Saxon Chronicle* (y. 897)	0	-	4
	DAT	N	SG	0	-	1	*The Voyages of Ohthere and Wulfstan* (848-899)	0	-	
	NOM	N	SG	1	*Anglo-Saxon Chronicle* (y. 897)	1	*Anglo-Saxon Chronicle* (y. 897)	0	-	
þæs	GEN	N	SG	4	*Anglo-Saxon Chronicle* (y. 755, 897); *Beowulf's Prologue*, 1000	0	-	0	-	4
þæt[14]				3	*The Voyages of Ohthere and Wulfstan* (848-899); *Anglo-Saxon Chronicle* (y. 991); *Beowulf's Prologue*, 1000	0	-	1	*The Voyages of Ohthere and Wulfstan* (848-899)	10
				3	The Venerable Bede (671-735); *Anglo-Saxon Chronicle* (y. 755); *Exeter Book*, 10th cent.	3	*The Voyages of Ohthere and Wulfstan* (848-899); *Genesis*, 1000	0		
Tot.								3		28

Table 4. Distribution of Old English deictics in the corpus.

As shown in Table 4, the few instances of demonstratives featured in the corpus were almost equally used both pronominally and adjectivally – thus showing that also in Old English the grammaticalisation of demonstrative pronouns into determiners had already taken place.

Conversely, there are only three sentences in which the demonstrative is used as a third-person pronoun and that leaves the question open about

---

[14] Spelling variation of *þæt*.

whether the pool of entries was again too small and not representative enough, or whether those were merely sporadic uses which had never really become established. Either way, based on some extracts from the *Canterbury Tales*, this specific use seems to have already been dropped by the Middle English period.

As for the other West Germanic languages, as reported in the corpus through few but significant examples, the Old Saxon *that*, the Old Frisian *thi*, *thiu*, *thet* and the Old High German *daz*, *thaz* – as well as their inflected forms – all functioned as both demonstrative pronouns and determiners with deictic function when used in adjectival position. In addition, Old High German featured a second demonstrative pronoun, *thiz*, which had an emphatic function.

Old Low Franconian, on the other hand, stands out for being the only West Germanic language in which the grammaticalisation of determiners appears to have not yet occurred. The demonstrative *thie*, therefore, was probably only used as an anaphoric pronoun with deictic function (11).

(11) OLF   <u>*Ginathi*</u> <u>*in*</u>   <u>*uuarheide sina*</u>   *uue*   *sal*   **thia**   *suocan?*
            grace   and   truth   his   who   shall   those   question
            'His grace and truth, who shall question those?'

                                                         [cf. Appendix, no. 9]

Nevertheless, this seems merely due to a delay in the rise of the category of definiteness in the earlier stages of the Dutch language, as in Present-Day Dutch it is perfectly acceptable to use demonstratives in adjectival position (12).

(12) NL   **Die** <u>*zwarte*</u> <u>*fiets*</u> *is* *de* *mijne.*
            that   black   bike   is   the   mine
            'That black bike is mine.'

Among the Germanic languages analysed so far, however, Present-Day Dutch is also the only one in which a demonstrative still acts as the singular neuter third-person pronoun. This is, however, merely due to the assimilation of the two forms which occurred in Modern times.

4.1.3. North Germanic

As shown in Table 2, the Old Norse demonstratives system is composed of four main pronouns, with distinct origins and/or meanings. The most represented pro-

noun in the corpus is *sá*, *sú*, *þat*, which is a reflex of the Proto-Indo-European **só-*, **sá-*, **tód* and expresses a *that*-deixis. A similar form, of the same derivation, is constituted by the proximal demonstrative pronoun *sjá/þessi*, *sjá/þessi*, *þetta*, which, however, only appears in one instance throughout the corpus. Another common demonstrative in the corpus is *hinn*, *hin*, *hið*, which – despite its derivation from the Proto-Germanic distal demonstrative pronoun **jainaz*, like the Gothic *jáins*, *jaina*, *jainata* (Keretchashvili 2021) – is assumed to have lacked any specific spatial deixis and was often used as the second referent in constructions like "the one ... the other," where it was contrasted with another referent marked by either *sá* or *sjá* (Krause and Slocum 2005). Finally, another way to express the demonstrative function in Old Norse was through the third-person pronoun *hann*, *hon*, *þat.* This pronoun was really a demonstrative as its declension constituted a suppletive system, where its masculine and feminine singular forms were built from the demonstrative *hann* (again originating from the Proto-Germanic **jainaz),* whereas the neuter singular and the plural of all genders stemmed from the *þ*-forms found in the declension of the demonstrative *sá*, *sú*, *þat* (Krause and Slocum 2005).

Given this premise, it is therefore clear why so many entries featuring *hann* would have the function of third-person pronoun. Nonetheless, because of its direct origin from **só-*, **sá-*, **tód*, the main focus in this chapter will remain the demonstrative *sá*, *sú*, *þat* (it is, in fact, impossible to draw any conclusion based on the single instance of the other reflex of **só-*, **sá-*, **tód*, i.e. *sjá/þessi*, *sjá/þessi*, *þetta).*

				Demonstrative Pronoun		Demonstrative Determiner		Third-Person Pronoun		
	Inflection			No	Texts, Dating	No	Texts, Dating	No	Texts, Dating	**Tot.**
hinn[15]	NOM	M	SG	~~2~~	*Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.	0	-	0	-	~~2~~

[15] This form does not derive from the Proto-Indo-European *só-*, **sá-*, **tód.* Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 5.

	Case	Gen	Num							Total
sá	NOM	M	SG	2	*Poetic Edda*, 1270s	2	*Battle of Stamford Bridge* (y. 1066); *Hrólfs saga*, late 14th cent.	3	*Poetic Edda*, 1270s	7
sú	NOM	F	SG	2	*Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275)	0	-	0	-	2
þá	ACC	F	SG	0	-	1	*Edda* (1179-1241)	0	-	4
		M	PL	2	*Battle of Stamford Bridge* (y. 1066); *Poetic Edda*, 1270s	0	-	1	*Brennu-Njáls Saga* (1250-1275)	
þann	ACC	M	SG	2	*Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.	4	*Íslendingabók* (1067-1148); *Edda* (1179-1241); *Poetic Edda*, 1270s;	2	*Poetic Edda*, 1270s; *Egils Saga*, late 14th cent.	8
þat	NOM	N	SG	6	*Íslendingabók* (1067-1148); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Grettis Saga*, 14th cent.	0	-	6	*Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.	50
	ACC	N	SG	32	*Íslendingabók* (1067-1148); *Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.; *Egils Saga*, late 14th cent.	2	*Poetic Edda*, 1270s;	4	*Edda* (1179-1241); *Hrólfs saga*, late 14th cent.	

Form	Case	Gen	Num							Total
þau[16]	NOM	N	PL	2̶	Íslendingabók (1067-1148); *Brennu-Njáls Saga* (1250-1275)	0	-	1̶	*Poetic Edda*, 1270s	4̶
	ACC	N	PL	1̶	*Poetic Edda*, 1270s	0	-	0	-	
þæim	DAT	M	PL	0	-	0	-	1	*Battle of Stamford Bridge* (y. 1066)	1
þær	NOM	F	PL	0	-	0	-	4	*Poetic Edda*, 1270s	5
	ACC	F	PL	1	Edda (1179-1241)	0	-	0	-	
þeim	DAT	M	SG	1	*Poetic Edda*, 1270s	0	-	2	*Poetic Edda*, 1270s	21
		M	PL	1	*Poetic Edda*, 1270s	0	-	7	Íslendingabók (1067-1148); *Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Hrólfs saga*, late 14th cent.	
		F	PL	1	*Poetic Edda*, 1270s	0	-	0	-	
		N	PL	1	*Edda* (1179-1241)	0	-	8	*Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.	

[16] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 5.

Form	Case	Gender	Number							Total
þeir[17]	NOM	M	PL	6	Íslendingabók (1067-1148); *Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Hrólfs saga*, late 14th cent.	~~1~~	Íslendingabók (1067-1148);	~~40~~	*Battle of Stamford Bridge* (y. 1066); Íslendingabók (1067-1148); *Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Grettis Saga*, 14th cent.; *Hrólfs saga*, late 14th cent.	~~47~~
þeira[18]	GEN	M	PL	0	-	0	-	~~2~~	*Edda* (1179-1241); *Hrólfs saga*, late 14th cent.	6
		N		0	-	0	-	~~4~~	*Edda* (1179-1241); *Brennu-Njáls Saga* (1250-1275)	
þeiri	DAT	F	SG	0	-	0	-	1	*Brennu-Njáls Saga* (1250-1275)	1
þeirri	DAT	F	SG	1	Íslendingabók (1067-1148)	0	-	0	-	1

[17] This form does not derive from the Proto-Indo-European *só-*, **sá-*, **tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 5.

[18] This form does not derive from the Proto-Indo-European *só-*, **sá-*, **tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 5.

þess	GEN	M		0	-	1	*Poetic Edda,* 1270s	0	-	4
		N	SG	3	Íslendingabók (1067-1148); *Egils Saga,* late 14th cent.; *Hrólfs saga,* late 14th cent.	0	-	0	-	
þet	NOM	N	SG	3	*Battle of Stamford Bridge* (y. 1066)	0	-	1	*Battle of Stamford Bridge* (y. 1066)	6
	ACC	N	SG	2	*Battle of Stamford Bridge* (y. 1066)	0	-	0	-	
því	DAT	N	SG	23	*Battle of Stamford Bridge* (y. 1066); Íslendingabók (1067-1241); *Edda* (1179-1241); *Poetic Edda,* 1270s; *Grettis Saga,* 14th cent.; *Egils Saga,* late 14th cent.; *Hrólfs saga,* late 14th cent.	1	*Poetic Edda,* 1270s	1	*Hrólfs saga,* late 14th cent.	25
Tot.				**83**		**11**		**41**		**135**

Table 5. Distribution of Old Norse deictics in the corpus.

As reported in Table 5, the most represented function in the corpus, as far as the inflected forms of *sá, sú,* **þat** are concerned, is that of demonstrative pronoun – and unsurprisingly so, since that was the original function even in Proto-Indo-European. What, on the other hand, seems noteworthy is that, in comparison, its adjectival use as a demonstrative determiner is significantly under-represented. Although, as already mentioned, this could just be due to the random selection of entries, the difference in the number of occurrences is large enough to at least hypothesise that this function had not yet been fully established at the time in which

the texts included in the corpus were written. This latter hypothesis seems to be supported also by the fact that, in the Present-Day North Germanic languages, the demonstrative is equally used both as a pronoun and as an adjective.

Moreover, as already seen in Old English and especially in Gothic, even the "proper" demonstrative *sá*, *sú*, **þat** was often used as a third-person pronoun – a use which has been dropped in Present-Day Scandinavian languages, where, conversely, the descendents of the Old Norse pronoun *hann* have stabilised as the only third-person pronouns.

4.2. Articles

As discussed before, in Proto-Indo-European, the relationship between the demonstrative and the name to which it referred is believed to have been an appositional one (Quiles and López-Menchero 2011). While we cannot reconstruct the grammatical category of the article in Proto-Indo-European, as well as in the earliest attested stages of ancient Indo-European languages such as Homeric Greek, we do however find definite articles, for instance, in later stages of Greek itself. We may therefore still assume that – probably once the structure (anaphoric demonstrative pronoun + noun) and its meaning had stabilised – the nature of the demonstrative changed from appositional to adjectival, thus giving rise to the category of determiners.

The [+definiteness] feature, which generally distinguishes definite articles, could be expressed by the appositional demonstrative already in Proto-Indo-European (at least in contexts where it was salient) and has been the triggering element which gave rise to the category of determiners. It was thus only when, at a later stage, the determiner lost its [-deictic] feature – while preserving its [+definiteness] one – that it became a definite article. For this reason, articles may be considered as less marked then demonstrative determiners, as they lack the [-deictic] feature. Nonetheless, it is still unknown at what point in the history of the individual Germanic languages this renovation came into being.

Skrzypek, Piotrowska, and Jaworski (n.d.: 26-27) maintain that, whereas "there does seem to exist a difference between the demonstrative, in which the definite article originates, and the definite article itself," such difference is "not clear-cut, and therefore the stages of grammaticalisation are not discrete." They argue, in fact, that the development of the demonstrative

"forms a continuum from a purely deictic element to a distinct definite article" (Skrzypek, Piotrowska, and Jaworski n.d.: 27). From this perspective, we may thus expect some features belonging to the demonstrative as well as some of those proper to the article to have, at some point, coexisted within the grammaticalising demonstrative.

The main purpose of this chapter is therefore that of comparing the corpus data which have been annotated as either demonstrative determiners or definite articles [Table 6] and trying to decide, for each one of the languages taken into account, whether or not this latter function was already present and what its degree of grammaticalisation was at the time instantiated. Moreover, the origin of the North Germanic definite articles will be discussed, as well as the definite articles systems in the Present-Day Germanic languages.

Forms	Gothic	Old English	Old Norse						Tot.
	sa, so, þata	*sē, sēo, ðæt*	*hann*	*hinn*	*sá*	*sjá*	Tot.		
Demonstrative Determiners [cf. Section 2]	8	12	1	0	11	0	17		**37**
Definite Articles	52	12	1	6	12	1	22		**86**

Table 6. Distribution of determiners in the corpus.

4.2.1. East Germanic

Table 7 features the instances of Gothic demonstrative determiners and definite articles found in the corpus and analyses each lemma based on its inflection and occurrence.

				Definite Article		Demonstrative Determiner		
	Inflection			No	Texts, Dating	No	Texts, Dating	Tot
jainaim[19]	DAT	M	PL	0	-	Ⰶ	*Gospels*, 4th cent.	Ⰶ

[19] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations is therefore not relevant for the purposes of this study and it is not included in the sum total reported in the last row of Table 7.

Form	Case	Gen	Num	N	Source	N	Source	Total
jainans[20]	ACC	M	PL	0	-	‡	*Gospels*, 4th cent.	‡
sa	NOM	M	SG	6	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	7
soh	NOM	F	SG 6 *Gospels*, 4th cent.	0	-	1	*Gospels*, 4th cent.	7
so				0	-			
þai	NOM	M	PL	6	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	7
þaim	DAT	M	PL	3	*Gospels*, 4th cent.	0	-	3
þamma	DAT	M	SG	10	*Gospels*, 4th cent.	1	*Gospels*, 4th cent.	11
þana	ACC	M	SG	4	*Gospels*, 4th cent.	0	-	4
þans	ACC	M	PL	3	*Gospels*, 4th cent.	0	-	3
þata	NOM	M	SG	1	*Gospels*, 4th cent.	0	-	4
	ACC	M	SG	1	*Gospels*, 4th cent.	2	*Gospels*, 4th cent.	
þe	INSTR	N	SG	1	*Skeireins*, 5th cent	0	-	1
þis	GEN	M	SG	2	*Gospels*, 4th cent.	0	-	3
		N		1	*Gospels*, 4th cent.	0	-	
þizai	DAT	F	SG	1	*Gospels*, 4th cent.	0	-	1
þize	GEN	M	PL	0	-	0	-	2
		N	PL	2	*Gospels*, 4th cent.	0	-	
þizos	GEN	F	SG	1	*Gospels*, 4th cent.	0	-	1

[20] This form does not derive from the Proto-Indo-European *só-*, *sá-*, *tód*. Its attestations is therefore not relevant for the purposes of this study and it is not included in the sum total reported in the last row of Table 7.

þo	ACC	F	SG	2	*Gospels*, 4th cent.	1	*Skeireins*, 5th cent.	5
		N	PL	1	*Skeireins*, 5th cent.	1	*Gospels*, 4th cent.	
þos	ACC	F	PL	1	*Gospels*, 4th cent.	0	–	1
Tot.				52		8		60

Table 7. Distribution of Gothic determiners in the corpus.

As a first indication, all of the instances of definite articles are inflected forms of the Gothic demonstrative *sa*, *sō*, **þata**, while, in two occurrences, the demonstrative determiner is expressed through the already discussed forms *jainans* and *jainaim* [cf. Section 2]. Granted that the amount of data is not sufficient to draw any relevant conclusion on this point, it is safe to say that the reflex of **so-*, **sa-*, **tód* is at least the prevalent option in the formation of articles. This strong connection between the demonstrative and the article is, moreover, confirmed by the case of Greek, where the demonstrative **ὅδε**, **ἥδε**, **τόδε** is composed of the definite article **ὅ**, **ἥ**, **τό** (< PIE **so-*, **sa-*, **tód*) and the enclitic suffix **-δε**. Greek, in fact, was also the first Indo-European language to grammaticalise definiteness and to thus introduce definite articles. Indeed, we can find examples of overlapping distribution between the original demonstrative **ὅ**, **ἥ**, **τό** and the homographic definite article already in the *Odyssey*, as well as in the latest books of the *Iliad* (Chantraine 1958:165).

Despite the large difference in the number of occurrences of Gothic demonstrative determiners and definite articles in the corpus, in the absence of any reason to suppose otherwise, we may assume that any inflected form of the Gothic *sa*, *sō*, **þata** could be used to express both functions. Furthermore, written Gothic was a translational language based on a source language (i.e. testamentary Greek) which did feature definite articles and it is thus reasonable to presume that the target language would use them, as well.

Since, however, the identical surface of demonstrative determiners and definite articles is a confusing factor, in order to identify any difference of use between the two, we need to have a closer look at the meanings expressed within the specific contexts in which they occur.

Example (13) presents two occurrences of demonstratives used in adjectival position, both showing a clear [+deictic] feature: in the first case, in fact, "*jainans*" refers to some specific days, while, in the second case, the deictic function is even more explicit, in that "*soh gilstrameleins*" ("this taxing") constitutes an anaphoric reference to the "*gameljan*" ("being taxed") introduced in the previous sentence.

(13) GOT *Wa   þan   in dagans   jainans, urrann gagrefts fram   kaisara*
            happened and in days   those   rose   decree   from Caesar
            Agustau,   gameljan allana midjungard. Soh þan gilstrameleins frumista
            Augustus, be.taxed all   world   this and taxing   first
            warþ   at   [...].
            happened when
            'And it happened in those days (that) a decree from Caesar Augustus
            (stating that) all the world should be taxed rose. And this taxing first
            happened when […]'

                                                                          [cf. Appendix, no. 1-2]

(14) GOT *was-uh   þan   neƕa   pasxa,   so   dulþs   Iudaie.*
            was-and   then   near   Passover the   feast   Jew.GEN.PL
            'And then the Passover, the feast of the Jews, was near.'

                                                                          [cf. Appendix, no. 11]

As for example (14), however, it is beyond doubt that the determiner *so* is to be interpreted as an article: we could, in fact, translate it as either "*a* feast of the Jews" or "*the* feast of the Jews," depending on the importance of the feast from a cultural point of view, but we could never render it as "*that* feast of the Jews" unless we make the reference more specific, e.g. "*that* feast of the Jews during which they commemorate their liberation from slavery in Egypt" (in this latter case, the [+deictic] feature could only be motivated by emphasis, as mentioned in the previous example). However, in the absence of such addition and of any emphatic purpose in the passage taken from the corpus, it is clear that the function of the determiner *so* in this sentence can only be that of an article.

Even these few examples, in fact, seem to provide sufficient evidence to conclude that the signature features of definite articles were indeed present in the Gothic determiner *sa*, *sō*, **þata** and that, when it was used as such, it lacked, on the other hand, the [-deictic] feature which is proper to the

demonstrative. It is thus reasonable to suppose that, not unlike the "polyfunctional *that*" in Present-Day English, the Gothic surface-identical *sa*, *sō*, **þata** could function as a demonstrative pronoun, demonstrative determiner, third-person pronoun, and as a definite article. Although some degree of ambiguity is still present at times, we could argue that, for the most part, each function had specific features and contexts of use.

4.2.2. West Germanic

According to corpus data relative to Old English, both the entries annotated as demonstratives and those annotated as definite articles are all inflected forms of *sē*, *sēo*, *ðæt* and Table 8 provides a compared analysis of their occurrences in the corpus, with specific attention to the lemmas involved.

Unfortunately, however, the instances gathered in the corpus, as far as definite articles and demonstrative determiners are concerned, is too limited to make any relevant consideration and the fact that Table 8 shows that the definite article is only expressed through three forms (i.e., in **bold**, *ðā*, *þāra*, and *þæt*) is just due to their under-representation within the corpus. All the inflected forms of the Old English *sē*, *sēo*, *ðæt* served, in fact, both as the demonstrative pronoun/determiner *that*, *those*, and as the definite article *the*. Although *sē*, *sēo*, *ðæt* did have a peculiar declension, which may be considered as suppletive, since the various forms are quite difficult to predict, the only true difference between the demonstrative forms and the definite article ones, is that the article was unstressed and the masculine nominative form was thus spelled as *se* – rather than *sē*.

	Inflection			Definite Article		Demonstrative Determiner		
				No.	Texts, Dating	No.	Texts, Dating	**Tot.**
ðæt	ACC	N	SG	0	-	1	*The Voyages of Ohthere and Wulfstan* (848-899)	1
ðā	NOM	-	PL	3	*The Voyages of Ohthere and Wulfstan* (848-899); *Beowulf*, 1000	0	-	3

þā	ACC	-	PL	0	-	2	The Venerable Bede (671-735)	2
þa[21]	ACC	F	SG	0	-	1	*Genesis*, 1000	1
þam	DAT	-	PL	0	-	2	Peterborough Chronicle, 1087	2
þāra	GEN	-	PL	1	*Anglo-Saxon Chronicle* (y. 897)	0	-	1
þǣm	NOM	N	SG	0	-	1	*Anglo-Saxon Chronicle* (y. 897)	3
	DAT	N	SG	0	-	1	*The Voyages of Ohthere and Wulfstan* (848-899)	
		-	PL	0	-	1	*Anglo-Saxon Chronicle* (y. 897)	
þæt[22]	NOM	N	SG	3	*The Voyages of Ohthere and Wulfstan* (848-899); *Anglo-Saxon Chronicle* (y. 897)	0	-	10
	ACC	N	SG	4	The Venerable Bede (671-735); *The Voyages of Ohthere and Wulfstan* (848-899); *Exeter Book*, 10th cent.	3	*The Voyages of Ohthere and Wulfstan* (848-899); *Genesis*, 1000	
Tot.				**11**		**12**		**23**

Table 8. Distribution of Old English determiners in the corpus.

The examples which can be found in the corpus seem to show that there indeed existed some difference – albeit subtle – in the contexts in which the

[21] Spelling variation of *þā*.
[22] Spelling variation of *ðæt*.

Old English sē, *sēo*, ðæt could be used, based on whether it expressed deixis
(15) or definiteness (16).

(15) OE *[...] hīe   wǣron micle swīþor gebrocede   on **þ��m** *þrim* *gēarum*
             they were   much more   afflicted   in those three years
             mid cēapes   cwilde   ond monna.
             by   cattle.GEN.SG pestilence and man.GEN.PL
             'They were much more afflicted in those three years by pestilence of
             cattle and men.'

[cf. Appendix, no. 5]

(16) OE *[...] hē   forlēt **þæt** *hūs*   *þæs*   *gebēorscipes* [...]*
             he   left   the   house the.GEN.SG feast
             'He left the house of the feast'

[cf. Appendix, no. 12]

There are, however, also some more ambiguous instances, where a "partial
distributional overlap" may be identified, as in the example reported by Cris-
ma (2020: 39-40), taken from the *Catholic Homilies* (c. 990) by the English abbot
Ælfric, where the phrase "ðæt beclysede geat" could be equally rendered as
either "*the* locked gate" or "*that* locked gate."

(17) OE *Eft   Ezechiel se witega   geseah on his witegunge an belocen geat*
             then Ezechiel the prophet saw   in his prophecy a locked gate
             on Godes   huse. &   him cwæð to sum engel. **Þis** *geat*
             in God.GEN house and him said to some angel this gate
             ne   bið nanum men   geopenod ac   se hlaford ana færð
             NEG is no.DAT men.DAT   opened   but the Lord   one goes
             *in   þurh **þæt** *geat*. &   eft   ut   færð & hit bið*
             in   through that gate and back   out goes and it is
             *belocen   on ecnysse. **Ðæt**   *beclysede* *geat**
             locked   in eternity **The/That** locked   gate
             on Godes   huse getacnode þone halgan mægðhad þære
             in God.GEN house betokened the holy   virginity the.GEN
             eadigan   Marian
             blessed.GEN   Mary.GEN
             'Also Ezechiel the prophet saw in his prophecy a locked gate in God's
             house, and an angel said to him, "This gate shall be opened to no man, for
             the Lord alone will go in through that gate and then back out, and it shall

be locked for ever". The/That locked gate in the house of God signified the holy virginity of the blessed Mary.'

(Crisma 2020:39-40)

This ambiguity between the surface-identical demonstrative determiners and definite articles has often led scholars to wonder if there even existed a proper definite article (and, by extension, the category of definiteness altogether) in Old English. As Allen (2019) claims, in fact, we must not attempt to impose Present-Day categories on any earlier stage of the languages.

Before we can question the status of the Old English definite article, however, we need to understand why the case of English would be more prominent than that of the other languages where such ambiguity can be found – like, for instance, Gothic and Present-Day German.

As already discussed, the Gothic *sa*, *sō*, *þata* could be considered, at least to some extent, as an ambiguous determiner, in that it might not always be possible to distinguish when it served as a deictic marker from when it was used as a definiteness one. Unlike Old English, however, as we have already mentioned, Gothic is known to have had direct and massive contact with a language which already featured grammaticalised definiteness (i.e. Greek) and it is thus more straightforward to assume that it, too, presented definite articles.

Moreover, since the Gothic language died out quite soon, one can only speculate about whether at some point the two functions might have split into distinct forms. On the other hand, a compelling reason why the status of the Old English article is object of debate is indeed the presence, in Middle English, of a new form of the article (i.e. the invariant *þe*[23]), which was "phonologically and semantically distinct from the distal demonstrative[24]," which by that time had, in turn, developed into the invariant ðat and specialised as the distal demonstrative (Crisma 2020: 38). It is, in fact, the very existence of two distinct forms in later stages of the English lan-

---

[23] As we will see in [Section 4.3.2], in Old English, *þe* was also a relative particle.
[24] Definite articles which are distinct from the demonstrative are defined as "discrete definite articles" (Crisma 2020: 44).

guage that have brought the invariant *þe* to be regarded as the "proper" definite article.

(18) OE **Þe** *wrecce*   *men* *of* **þe** *land*
        the wretched men of the land
        'The wretched men of the land'

[cf. Appendix, no. 14-15]

Although its role as a definiteness marker is beyond doubt (18), however, it does not constitute sufficient reason to state that the previous form (i.e. *sē*, *sēo*, *ðæt*) did not cover that function, as well. The Present-Day German *der*, *die*, *das* is, in fact, both a demonstrative pronoun and a definite article[25] and neither of its functions is called into question. Therefore, the issue at hand is not really the overlapping distribution of the two functions, but rather whether grammaticalised definiteness had already been accomplished in the Old English period (Crisma 2020).

In her research, Crisma (2020) confirms that evidence of grammaticalised definiteness could be found in Old English prose (eighth century), but not in the Old English poetry and concludes that this grammaticalisation process must have occurred during the Old English period, centuries before the split of the two forms. On these grounds, although we cannot pinpoint the exact moment in which definiteness was grammaticalised, we may generally say that, roughly after the eighth century, when *sē*, *sēo*, *ðæt* was used to indicate definiteness, from a syntactical point of view, it was indeed a definite article.

While the morphological split could be "a case of parallel decay in form and meaning" (Crisma 2000: 76), in that both forms, *þæt* and *þe*, lost their inflectional properties and, respectively, their [-definiteness] and [-deictic] features, from a typological point of view, the grammaticalisation of definiteness is a process which has occurred in a significant number of languages, including non-Indo-European ones.

As for the other West Germanic languages, the Old Saxon *that*, the Old Frisian *thi*, *thiu*, *thet*, and the Old High German *daz*, *thaz* were all examples of "near-articles," in that they were demonstrative pronouns which, just like the Gothic *sa*, *sō*, *þata* and the Old English *sē*, *sēo*, *ðæt*, could be used as definite

---

[25] "A demonstrative pronoun with article function is called a "near-article" (Crisma 2020: 44).

articles. Besides *daz, thaz*, however, Old High German also featured a second determiner which could also be used as a relative pronoun, i.e. *der, ther*.

Finally, as already mentioned, Old Low Franconian did not yet feature grammaticalised definiteness and therefore did not have definite articles. On the other hand, Present-Day Dutch does present a definite article (i.e. *de, het*[26]), also because, as we have seen, it is not a direct ancestor of Dutch.

4.2.3. North Germanic

Table 9 presents a comparison of the Old Norse entries from the corpus which have been annotated as either demonstrative determiners or as definite articles. While, in the corpus, the demonstrative determiner is almost always expressed by the Old Norse demonstrative *sá, sú, þat* (except for one single instance where the third-person pronoun *hann, hon, þat* is used), the representation of the definite article is more heterogeneous. We can find, in fact, a restricted number of cases in which it is expressed by the demonstratives *hinn, hin, hið* and *sjá/þessi, sjá/þessi, þetta*.

Nevertheless, as shown in Table 9, the inflected forms of *sá, sú, þat* do prevail also in the realisation of the definite article, while the other options analysed occur far too sporadically for us to make any significant consideration in those regards.

| | Inflection | | | Definite Article | | Demonstrative Determiner | | Tot. |
				No.	Texts, Dating	No.	Texts, Dating	
Hin[27]	NOM	F	SG	1	*Edda* (1179-1241)	0	-	2
	ACC	M	SG	1	*Battle of Stamford Bridge* (y. 1066)	0	-	

Form	Case	Gen	Num					Total
~~hinn~~[28]	NOM	M	SG	~~3~~	*Battle of Stamford Bridge* (y. 1066); *Hrólfs saga*, late 14th cent.	0	-	~~3~~
~~hinum~~[29]	DAT	M	PL	~~1~~	*Hrólfs saga*, late 14th cent.	0	-	~~1~~
sá	NOM	M	SG	1	*Edda* (1179-1241)	2	*Battle of Stamford Bridge* (y. 1066); *Hrólfs saga*, late 14th cent.	3
þá	ACC	F	SG	0	-	1	*Edda* (1179-1241)	2
		M	PL	1	*Edda* (1179-1241)	0	-	
þann	ACC	M	SG	0	-	4	*Íslendingabók* (1067-1148); *Edda* (1179-1241); *Poetic Edda*, 1270s;	4
þat	NOM	N	SG	1	*Egils Saga*, late 14th cent.	0	-	3
	ACC	N	SG	0	-	2	*Poetic Edda*, 1270s	
þæim	DAT	M	SG	1	*Battle of Stamford Bridge* (y. 1066)	0	-	1
þeim	DAT	M	SG	1	*Poetic Edda*, 1270s;	0	-	1
~~þeir~~[30]	NOM	M	PL	0	-	~~1~~	*Íslendingabók* (1067-1148);	~~1~~

[28] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations are therefore not relevant for the purposes of this study AND THEY are not included in the sum total reported in the last row of Table 9.

[29] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations is therefore not relevant for the purposes of this study AND IT IS not included in the sum total reported in the last row of Table 9.

[30] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód*. Its attestations is therefore not relevant for the purposes of this study and it is not included in the sum total reported in the last row of Table 9.

þeira[31]	GEN	N	PL	ǂ	Íslendingabók (1067-1148)	0	-	ǂ
þeiri	DAT	F	SG	1	*Edda* (1179-1241)	0	-	1
þenn	ACC	M	SG	2	*Battle of Stamford Bridge* (y. 1066)	0	-	3
þenna	ACC	M	SG	1	*Battle of Stamford Bridge* (y. 1066)	0	-	
þess	GEN	M	SG	1	Íslendingabók (1067-1148)	1	*Poetic Edda*, 1270s	4
		N		2	*Hrólfs saga*, late 14th cent.	0	-	
því	DAT	N	SG	1	Íslendingabók (1067-1148)	1	*Poetic Edda*, 1270s	2
Tot.				**13**		**11**		**24**

Table 9. Distribution of Old Norse determiners in the corpus.

From a closer look at the sentences exemplified in the corpus, we can also see that, in addition to the significant instability in the formation of the definite article, there are virtually no instances in which the [+definiteness] feature indisputably prevails over the [+deictic] feature.

The slight difference in meaning between the forms annotated as demonstrative determiners (19) and those annotated as definite articles (20) is, in fact, not relevant enough to suppose that the demonstrative *sá, sú, **þat*** could indeed serve as a definiteness marker with no [-deictic] feature.

(19) ON *[...]* *Óláfr* *inn* *Digri* *gørði* *skýrt* *at* *hverr* *maðr* *skyldi* *gjalda* *[...]*
Olaf the thick made definite that each man should pay
nema *konur*
except women

---

[31] These forms do not derive from the Proto-Indo-European *só-, *sá-, *tód* and are therefore not relevant for the purposes of this study. These instances are not included in the sum total reported in the last row of Table 9.

eða **þeir** <u>*menn*</u> *er hann næmi frá.*
or   those men whom he   exempts from
'Olaf the Thick made definite that each man should pay [...] except wom-
en or those men whom he exampts.'

<div align="right">[cf. Appendix, no. 7]</div>

(20) ON *Kenndo þér þenn hinn   myckla meðr* **þæim** <u>*blá*</u> <u>*kyrtli*</u> *oc hin*
know you that this.one noble with the   blue kirtle and the
faghra   hialm [...]?
impressive helmet
'Do you know that noble man with the blue kirtle and impressive helmet [...]?'

<div align="right">[cf. Appendix, no. 13]</div>

On the other hand, we cannot rule out the possibility that the category of definiteness was grammaticalised, by virtue of the fact that Old Norse did feature a postpositive definite article (21), i.e. *inn, in, it* – whose declension followed that of the demonstrative *hinn, hin, hi*ð (Krause and Slocum 2005).

(21) ON *[...] Svá kom   at þeir týndu guðs   nafni ok víðast um*
so occurred that they forgot God.GEN name and widely around
*verǫld-**ina** fansk   eigi sá maðr er deili   kunni á*
world-ART could.be.found not DET man who details knew of
skapara sínum.
creator his
'So it happened that they forgot God's name and in most places around the world a man could not be found who knew the details of his creator.'

<div align="right">(*Snorra Edda*, Krause and Slocum 2005)</div>

While *hinn, hin, hi*ð expressed distal deixis and derived from an Indo-European pronoun combining variants of the stems **ko-* and **eno-* (cf. the Greek distal demonstrative ἐκεῖνος), the article *inn, in, it* was a Proto-Scandinavian innovation (Stroh-Wollin 2020). This article could be found in some *Older Fuþark* runic inscriptions [cf. *hino* in Appendix, no. 10], according to Stroh-Wollin (2020), derived from an anciently inherited demonstrative based on the stem *en-* (< Proto-Indo-European **eno-*), which had originally been strengthened through the addition of a Proto-Germanic root expressing proximal deixis (**hī-* or **hi-* < Proto-Indo-European **ko-*). As we can thus notice, this account represents typological evidence of the grammaticalisation of demonstratives into definite articles, regardless of the stem from which the demonstrative came.

As for the Present-Day Scandinavian languages, as already seen [cf. Section 2], they all present the suffixal form derivating from *-inn, -in, -it* as their sole or – at least – primary definiteness marker. While, in fact, Danish and Swedish do also feature a free-standing article whose form is compatible with a possible origination from *sá*, *sú*, *þat*, Nowegian and Icelandic only use the suffix. However, since the suffix *-inn, -in, -it* is assumed to have originated from the above-mentioned article *inn, in, it* (Stroh-Wollin 2020), we could argue that all the North Germanic languages have undergone the grammaticalisation path by which a demonstrative became a definite article (Skrzypek, Piotrowska, and Jaworski n.d.).

4.3. Relatives

Based on the relatives system in the Present-Day Germanic languages, we may assume that one of the West Germanic relatives (i.e. the *d-Pronomen*) derived from the Proto-Indo-European **so-, *sa-, *tód*, while there is no trace of such relative form in the Scandinavian languages. What we do not know, however, is how and why this development came into being, as well as when it occurred and what languages it involved.

Through an in-depth analysis of some salient entries from the corpus, this section thus tries to verify whether the demonstrative had already been grammaticalised into a relative in the early stages of the Germanic languages and, if so, in which languages. The main criterion that will be used to test the relevant sentences is the relationship between the main clause and the supposed relative one. In order to consider a relative as such, in fact, the two clauses must form at least a correlative structure (i.e. either a *diptyque normal* or a *diptyque inverse*), in that, in any other case, the two clauses would just be coordinates and the element heading the second clause would have no reason to be interpreted as anything different from a mere anaphoric demonstrative.

FORMS	GOTHIC		OLD ENGLISH		OLD NORSE		
	sa, so, þata	*sa + ei*	*sē, sēo, ðæt*	*þe*	*sa + er*	*er*	*sem*
Inflected	6	17	5	0	8	0	0
Non-inflected	0	0	0	31	0	46	7
Tot.	**24**		**37**		**61**		

Table 10. Distribution of relative elements in the corpus.

At a first look, all the Gothic, Old English, and Old Norse tokens featured in the corpus which have been annotated as relatives belong to three categories: demonstratives, relative particles, or a combination of both [Table 10]. All the demonstratives instantiated are reflexes of **so-, *sa-, *tód*, while the so-called "combinations" are compound forms composed of a demonstrative element (again from **so-, *sa-, *tód*) and a relative particle. As for their specific forms, as could be expected, all the demonstratives and the compounds are inflected, while the relative particles are clearly indeclinable.

4.3.1. East Germanic

As shown in Table 10, based on the corpus data, Gothic seems to have expressed relatives either through the simple demonstrative *sa, so, þata* or through the compound form made of *sa, so, þata* itself and the relative particle *ei.* In addition, Table 11 features an analysis of the specific lemmas involved and provides information on their occurrence within the corpus.

				Demonstrative		Demonstrative + Relative Particle			
	Inflection			No	Texts, Dating	No	Texts, Dating	Form	Tot.
sa	NOM	M	SG	1	*Gospels*, 4th cent.	5	*Gospels*, 4th cent.	*saei*	6
so	NOM	F	SG	0	-	3	*Gospels*, 4th cent.	*sei*	3
þai	NOM	M	PL	1	*Gospels*, 4th cent.	1	*Skeireins*, 5th cent.	*þaiei*	2
þaim	DAT	M	PL	1	*Gospels*, 4th cent.	0	-	-	1
þamma	DAT	M	SG	2	*Gospels*, 4th cent.	0	-	-	2
þana	ACC	M	SG	1	*Gospels*, 4th cent.	0	-	-	1
þata	NOM	N	SG	0	-	4	*Gospels*, 4th cent.	*þatei*	4
þizai	DAT	F	SG	0	-	1	*Gospels*, 4th cent.	*þizaiei*	1

þo	ACC	F	SG	0	-	1	*Gospels*, 4th cent.	*þoei*	2
		N	PL	0	-	1	*Skeireins*, 5th cent.	*þoei*	
þos	ACC	F	SG	0	-	1	*Gospels*, 4th cent.	*þozei*	1
Tot.				6		17			23

Table 11. Distribution of Gothic relative elements in the corpus.

According to Krause and Slocum (2006), Gothic had no independent pronoun which could function specifically as a relative and the relative marker function was thus covered by the particle *ei*[32], when "conjoined to the demonstrative." This compound is described as the relative pronoun for the third person, while, to create first and second-person pronouns the demonstrative was replaced by the appropriate form of the personal pronoun (e.g. *ik+ei, þu+ei*). In addition, Krause and Slocum (2006) argue that, just like in Present-Day English, the relative pronoun always "derives its case from its function in the relative clause."

As previously stated, however, the goal of this corpus-based analysis of relatives is to verify whether the demonstrative pronoun derived from **so-, *sa-, *tód* had already been grammaticalised into a relative in the early stages of the Germanic languages. To this aim, we must disregard the above-mentioned compounds, in that in such forms the relative function is expressed by the particle and there is no indication that the demonstrative could be anything different from an originally free-standing antecedent of the relative, which has later been conjoined to the relative particle.

Conversely, we will need to focus on the few cases in which the simple demonstrative is used and, for once, the low number of instances might indeed be an extremely significant indicator. Let us thus consider each of the six cases individually.

The first observation we can make on these passages is that in (22) and (23) the supposed relative clause features non-finite verb forms (i.e., more

---

[32] Although its etymology is unclear, *ei* is presumed to have derived from the Proto-Indo-European anaphoric stem **i-* (Keretchashvili 2018).

specifically, two participles which could thus be considered as adjectives), whereas in (24), (25), (26), and (27) it does not present a verb at all. In either case, in the absence of a proper verb, the <u>supposed relative</u> cannot be defined as a clause. Nonetheless, the following sentences will be analysed in an attempt to further prove that, although the demonstrative in **bold** might be interpreted and translated in Present-Day English as a relative pronoun, it functioned as a determiner in Gothic.

(22) a. GOT *usgeisnodedun*   *þan*   <u>*allai*</u>   **þai**
         were.astonished.3PL   and   all.NOM.M.PL   DEM.NOM.M.PL
         <u>*hausjandans*</u>   *is*   *ana frodein jah*   *andawaurdjam is.*
         hear-PRSP.NOM.M.PL   him   at   wisdom and   answers   his
       'And all those hearing him were astonished at his wisdom and answers.'
                                              [cf. Appendix, no. 16]

      b. GR   *ἐξίσταντο*   *δὲ*   *πάντες*   *οἱ*
         were.astonished.3PL   then   all.NOM.M.PL   DEM.NOM.M.PL
         ἀκούοντες   *αὐτοῦ […]*
         hear-PRSP.NOM.M.PL   him
       'Then all those hearing him were astonished.'
                                                  (Luke 2:47)

(23) a. GOT   <u>*Iesu*</u>   *sokeiþ*   <u>*Nazoraiu*</u>   *þana*   <u>*ushramidan.*</u>
         Jesus   seek.2PL   of.Nazareth   DEM.ACC.M.SG   crucify.PP.ACC.M.SG
       'You seek Jesus of Nazareth, the crucified.'
                                               [cf. Appendix, no. 17]

      b. GR   *Ιησους*   *ζητειτε*   *τον*   *Ναζαρηνος*   *τον*
         Jesus   seek.2PL   DEM.ACC.M.SG   of.Nazareth   DEM.ACC.M.SG
         εσταυρωμενον
         exposed.PP.ACC.M.SG
       'You seek Jesus the Nazarene, the exposed'
                                                 (Mark 16:6)

As just mentioned, since "*hausjandans*" (22) and "*ushramidan*" (23) are verbal adjectives, the demonstrative pronoun which precedes and agrees with them could be described as a determiner[33] heading the appositional determiner

---

[33] As seen in Section 4.1.1 and Section 4.2.1, Gothic demonstratives were "near-articles" as they could indeed function both as demonstrative determiners and as definite articles.

phrases "*þai hausjandans*" and "*þana ushramidan.*" Interestingly, but perhaps also unsurprisingly, the Greek model (22b, 23b) from which these passages have been translated present the exact same constructions.

On the other hand, while with sentences (24), (25), (26), and (27) we could still be in the presence of appositional determiner phrases (there is, in fact, no evident correlative structure), we might assume that the verb *to be* is simply implied and that a relative clause is thus indeed present.

(24) GOT *[...]* <u>*atta*</u>   <u>*izwar*</u> **sa**   <u>*ufar*</u>   <u>*himinam.*</u>
                       father.NOM.M.SG your DEM.NOM.M.SG above heaven.DAT.M.SG
                     '[...] your father, the one that is above heavens.'
                                                     [cf. Appendix, no. 18]

However, while this reasoning could hold for sentence (24), where the supposed relative pronoun – which, nonetheless agrees with its antecedent – could also be seen as deriving its case from the relative itself, it cannot be accepted with regards to sentences (25), (26) and (27). As we can see, in fact, the demonstrative always agrees with the noun it follows and does not derive its case from its function in the <u>supposed relative clause</u>.

(25) GOT [...] *iþ*   <u>*jainaim*</u>   **þaim**   <u>*uta*</u>   *in gajukom*
                but   DEM.DAT.M.PL DEM.DAT.M.PL without in parables
                allata wairþiþ
                all   become.3SG
              '[...] But to them – those without (it) – all is done in parables.'
                                            [cf. Appendix, no. 19]

(26) GOT *[...]* <u>*attin*</u>   <u>*izwaramma*</u> **þamma**   <u>*in*</u>   <u>*himinam.*</u>
                 father.DAT.M.SG your   DEM.DAT.M.SG in   heaven.DAT.M.SG
                 '[...] Your father, the one in heaven.'
                                            [cf. Appendix, no. 20]

(27) GOT *[...]* *bidei*   *du* <u>*attin*</u>   <u>*þeinamma*</u> **þamma**
                 pray.imp.2sg to father.DAT.M.SG thy   DEM.DAT.M.SG
                <u>*in*</u>   <u>*fulhsnja,*</u>
                in   secret.DAT.N.SG
              '[...] Pray to your father, the one that is in secret.'
                                          [cf. Appendix, no. 21]

One final and conclusive proof of this can be provided by example (28) – the natural continuation of sentence (27) – in which a proper relative clause is present (there is, in fact, the verb "*saiƕiþ*") and, consequently, the proper relative pronoun *saei* is used.

(28) GOT *[...] jah atta þeins saei   saiƕiþ   in   fulhsnja,*
      and father thy REL.NOM.M.SG see.3SG in   secret.DAT.N.SG
      usgibiþ   þus   in   bairhtein
      reward.3SG   thee   in   brightness
      '[...] And your father who sees in secret, rewards you in brightness.'
      [cf. Appendix, no. 22]

Based on this analysis, it could thus be argued with a fair degree of certainty that Gothic demonstratives were not grammaticalised into relatives and that the only relative marker was indeed the compound form featuring the relative particle *ei*.

4.3.2. West Germanic

As far as Old English is concerned, Table 12 shows that its relative marker appears to have been expressed either by the relative particle *þe* or by an inflected demonstrative pronoun. However, differently from Gothic and Old Norse, the Old English relative particle *þe* is believed to have derived, itself, from the instrumental case of the demonstrative *sē, sēo, ðæt* (Keretchashvili 2018). However, based on the inflection of second-person pronoun (Keretchashvili 2018), its origin from the dative/instrumental singular form seems to be acceptable, as well.

	Inflection			POS	Texts, Dating	Tot.
ðe[34]	UNINFLECTED			Relative Particle	*Anglo-Saxon Chronicle* (y. 991); *Genesis*, 1000	2
sē	NOM	M	SG	Demonstrative Pronoun	*Anglo-Saxon Chronicle* (y. 755); *Beowulf*, 1000	2

[34] Spelling variation of *þe*.

þæt	NOM	N	SG	Demonstrative Pronoun	_The Voyages of Ohthere and Wulfstan_ (848-899); _Anglo-Saxon Chronicle_ (y. 991)	2
	ACC	N	SG	Demonstrative Pronoun	_Genesis_, 1000	1
_þe_	UNINFLECTED			Relative Particle	The Venerable Bede (671-735); _Anglo-Saxon Chronicle_ (y. 755, 897); _The Voyages of Ohthere and Wulfstan_ (848-899); _Exeter Book_, 10th cent.; _Peterborough Chronicle_, 1087	29
þone	ACC	M	SG	Demonstrative Pronoun	_The Voyages of Ohthere and Wulfstan_ (848-899)	1
Tot.						**39**

Table 12. Distribution of Old English relative elements in the corpus.

As for the demonstrative, it only headed a limited number of supposed relative sentences in Old English and the ones included in the corspus will be individually analysed.

In examples (29), (30), and (31), although a finite verb is present in each supposed relative clause, the relation of the two clauses is not of a correlative type. The clause headed by the demonstrative can, in fact, be interpreted as either an apposition (29) or an attribute (30, 31) of the noun phrase to which it refers and, as a result, unlike correlated and subordinate clauses, they could stand on their own.

(29) OE _[...] alle būtan ānum_, **_sē_**   _wæs þæs aldormonnes godsunu._
         all but   one,   DEM.NOM.M.SG was his   noble   godson
         '[...] all but one – that one was his noble godson.'

                                                                        [cf. Appendix, no. 23]

(30) OE _Þonne is ān port on sūðeweardum þǣm_   _lande,_
         there is a port in the.south.of DEM.DAT.N.SG land
         þone   _man_   _hǣt_   _Scīringes_ _hēal._
         DEM.ACC.M.SG one   calls   Skiringssal
         'There is a port in the south of that land – that (land) one calls Skiringssal.'

                                                                        [cf. Appendix, no. 24]

(31) OE *Geworhton* ðā | *Wedra*   *lēode*
      built   the   of.Weders people
      <u>*hlǣw*</u> *on* *hliðe,* | **sē**   <u>*wæs*</u> <u>*hēah*</u> <u>*ond*</u> <u>*brād,*</u> *[...]*
      mound upon hill   DEM.NOM.M.SG was high and wide
      'The people of the Weders built a mound upon the hill – that (mound)
      was high and wide.'
                                   [cf. Appendix, no. 26]

As for example (32), it is composed of one single clause and "*eall þæt*" is thus probably to be analysed as a syntagma used to expand the indefinite pronoun *all.*

(32) OE *[...]* *and* <u>*eall*</u> **þæt**   <u>*to*</u> <u>*fæsle*</u> <u>*frea*</u> *ælmihtig*
          and all DEM.ACC.N.SG for progeny lord almighty
      <u>*habban*</u> <u>*wolde;*</u>
      hav   would
      'and everyone (=all those) the Lord almighthy would have for progeny.'
                                   [cf. Appendix, no. 27]

The same could be said of sentence (33), where we can also find a second demonstrative functioning as a resumptive element, which might indicate that, even though the model of a correlative construction seems to be present, the structure had not yet been grammaticalised.

(33) OE <u>*Eal*</u> **þæt**   <u>*his*</u> <u>*man*</u> <u>*āþer*</u> <u>*oððe*</u> <u>*ettan*</u> <u>*oððe*</u> <u>*erian*</u>   <u>*mæg,*</u>
      all DEM.NOM.M.SG a man either or   graze or   plough can
      þæt   *lið*   *wið*   *ðā* *sǣ.*
      DEM.NOM.M.SG   extends alongside the sea
      'Everything (all that) a man can either graze or plough – that extends
      alongside the sea.'
                                   [cf. Appendix, no. 25]

While it is thus clear that the inflected demonstrative was not a relative in Old English, the status of the relative *þe* is, to this day, a central issue in Germanic studies.

(34) OE *he* *wǣs* *milde* *þam*   <u>*goum*</u> <u>*mannum*</u>   **þe** *god* *lufedon.*
      he was gentle DEM.DAT.PL good man.DAT.PL REL God love.3PL
      'He was gentle with those good men who love God.'
                                   [cf. Appendix, no. 28]

As example (34) – among many others which can be found in the corpus – seems to indicate, the particle *þe* probably was a relative marker, just like the Gothic (and, as we will see, the Old Norse) relative particles. However, what is not clear is whether it was indeed a demonstrative which had been grammaticalised into a relative or something else entirely. During the Middle English period, in fact, it took over the function of the definite article, and was replaced as a relative marker by the form *þat* (Lee 2006) – which, again, derived from the demonstrative *sē, sēo, ðæt*. Furthermore, the Middle English *þat*, as well as the Old English *þæt*, could also be used as argumental complementisers and this is one of the main reasons why we could argue that the West Germanic *d-Pronomen* did not directly originate from the demonstrative, but rather from the complementiser itself[35].

Following on this hypothesis and on Bacskai-Atkari's study (2019), it would appear reasonable to maintain that the English *d-Pronomen that* was neither a decategorised pronoun nor a relative particle, but rather a relative complementiser.

As for the other West Germanic languages, they all expressed what in the corpus has been annotated as a relative through forms deriving from the demonstrative. These forms are either uninflected ones as the Old Saxon *that* (which also functions as a complementiser) and the Old Frisian relative particle *ther*, or inflected, as the Old Low Franconian *thie*, and the Old High German *thie, thar* and *der, ther*. However, none of these forms functioned as a true relative marker.

The sole exception was, in fact, the Old Frisian *thi, thiu, thet* (35), but we must, however, consider that this passage is taken from a fourteenth century manuscript, which is thus more recent than most of the texts included in the corpus.

(35) OF *ther   hi on eskriuin hede tha tian bodo,   tha*
        there he on written had the ten commandments DEM.ACC.PL
        skolde   hi lera   tha   Israheliska   folke.
        should   he teach the   Israelite   folk
        'On there he had written the ten commandments – the ones he should teach the Israelite folk.'

[cf. Appendix, no. 29]

---

[35] This hypothesis will be discussed in detail in Section 5 of this paper.

4.3.3. North Germanic

Finally, in Old Norse, as we can see from Table 13, there are two main types of relative markers: the relative particles *er*[36] and *sem*[37] and the compound forms composed of the demonstrative *sá*, *sú*, **þat** and the relative particle *er* itself.

	Inflection			POS	Texts, Dating	Tot.
er[38]	UNINFLECTED			Relative Particle	*Battle of Stamford Bridge* (y. 1066); *Íslendingabók* (1067-1148); *Edda* (1179-1241); *Bren-nu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Egils Saga*, late 14th cent.; *Grettis Saga*, 14th cent.; *Hrólfs saga*, late 14th cent.	~~46~~
sás	ACC	M	SG	Demonstrative *sá* + Relative Particle *er*	*Poetic Edda*, 1270s	1
sem[39]	UNINFLECTED			Relative Particle	*Hrólfs saga*, late 14th cent.	~~6~~
				(Demonstrative Antecedent +) Relative Particle	*Battle of Stamford Bridge* (y. 1066)	~~1~~
þœrs	NOM	F	PL	Demonstrative *sá* + Relative Particle *er*	*Poetic Edda*, 1270s	1

[36] The Old Norse relative particle *er* derives from the Proto-Germanic masculine third-person pronoun **iz* (Keretchashvili 2018).

[37] The Old Norse relative particle *sem* has uncertain origins, possibly from the Proto-Germanic adjective **saman*, meaning *same* (Keretchashvili 2018). The main difference between *er* and *sem* is that the latter was used especially in the later texts.

[38] This form does not derive from the Proto-Indo-European *só-*, **sá-*, **tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 9.

[39] This form does not derive from the Proto-Indo-European *só-*, **sá-*, **tód*. Its attestations are therefore not relevant for the purposes of this study and they are not included in the sum total reported in the last row of Table 9.

þanns	ACC	M	SG	Demonstrative *sá* + Relative Particle *er*	*Poetic Edda*, 1270s	3
þats	ACC	N	SG	Demonstrative *sá* + Relative Particle *er*	*Poetic Edda*, 1270s	2
þeims	DAT	M	SG	Demonstrative *sá* + Relative Particle *er*	*Poetic Edda*, 1270s	1
Tot.						**8**

Table 13. Distribution of Old Norse relative elements in the corpus.

According to Krause and Slocum (2005), the relative particles were often preceded by a form of *sá*, *sú*, *þat* which could alternatively belong to the relative clause (when compounded with the particle *er*), or to the preceding clause – thus bearing no relation to the case represented by the particle. In either case, the demonstrative is always an antecedent of the relative and not a relative marker itself. Since there is, moreover, no instance of use of the simple demonstrative as a relative, we might conclude that, based on corpus data, Old Norse did not have a grammaticalised relative pronoun which derived from the demonstrative *sá*, *sú*, *þat*.

Although this account would seem to be confirmed also by the distribution of the relatives in the Present-Day Scandinavian languages – where, in fact, no relative *d-Pronomen* is present – we cannot completely rule out the possibility that the grammaticalisation of the demonstrative into a relative might indeed have happened at a later stage than Old Norse and that the resulting *d-Pronoun* was then dropped some time before the Present-Day stage of the languages.

A possible confirmation in this direction would be the already mentioned account by Vigfusson (1874:29), in which he claims that the Icelandic archaic argumental complementiser *at* (which probably derived from the demonstrative *sá*, *sú*, *þat*) could also be used as a less frequent alternative of the relative particle *er*.

4.4 Conjunctions

This section intends to analyse the conjunctions which are believed to have originated, at least in some Present-Day Germanic languages, from the Proto-Indo-European demonstrative *só-, *sá-, *tód* – and, more specifically, the argumental complementiser and the consecutive conjunction.

Given the identity of form among demonstratives, relatives and conjunctions [cf. Section 2], the derivation of the latter from the first seems to be rather uncontroversial as far as the West Germanic languages are concerned. On the other hand, however, the origin of the Old Norse conjunction *at* and its reflexes across the Present-Day Scandinavian languages is, as previously anticipated, still object of debate.

Furthermore, since the previous chapter showed that relatives did not seem to be grammaticalised in the early stages of any of the Germanic languages, it is particularly interesting to verify whether argumental complementisers and consecutive conjunctions were.

This chapter will thus analyse the forms annotated in the corpus as either argumental complementisers or consecutive conjunctions and try to establish their origin (especially as far as the North Germanic languages are concerned) and their degree of grammaticalisation.

As shown in Table 14, we can see at a first look that all of the forms instantiated (except for the Gothic compound relative *sa+ei*) are invariant ones. While we recognise the relative particles *ei* and *sem* in, respectively, Gothic and Old Norse, and the now uninflected Old English demonstrative *ðæt/þæt*, we can also find some other forms which are worth discussing individually in the dedicated sections.

FORMS	GOTHIC				OLD ENGLISH	OLD NORSE		Tot.
	ei	*þatei*	*þei*	*sa + ei*	*ðæt/ þæt*	*at*	*sem*	
ARG	3	9	0	2	17	48	1	**80**
CONS	6	0	3	0	6	28	0	**43**
Tot.	23				23	77		121

Table 14. Distribution of conjunctions in the corpus.

Despite the individual differences among the different languages, however, all argumental complementisers do share, as a common feature, the context in which they can be found. Such context is given by the verbs introducing them, which all belong to specific semantic categories. Although the general label of *verba dicendi* and *sentiendi* – which has been used so far – is still applicable, as it is evident from Table 15, there are in fact many more categories involved.

Semantic Category	Verbs	GOT			OE	ON		Tot.
		ei	*þatei*	*sa + ei*	*ðæt/ þæt*	*at*	*sem*	
Predicate Adjectives	be	0	0	0	0	1	0	**1**
	be best	0	0	0	0	1	0	**1**
	be far	0	0	0	0	1	0	**1**
	be fitting	0	0	0	1	0	0	**1**
	be likely	0	0	0	0	2	0	**2**
Verba Dicendi	claim	0	0	0	0	1	0	**1**
	say	0	5	0	8	5	0	**18**
	tell	0	1	0	0	2	0	**3**
Verba Desiderandi	desire	0	0	0	0	1	0	**1**
	want	0	0	0	0	1	0	**1**
Verba Existimandi	decide	0	0	0	0	1	0	**1**
	determine	0	0	0	0	1	0	**1**
	know	0	1	0	2	3	0	**6**
	mean	0	0	0	0	1	0	**1**
	reason	0	0	0	0	1	0	**1**
	settle (on)	0	0	0	0	1	0	**1**
	suspect	0	0	0	0	1	0	**1**
	understand	0	0	0	0	2	0	**2**
Verba Iubendi	command	1	0	0	0	0	0	**1**

Verba Sentiendi	hear	0	1	0	2	0	0	**3**
	see	0	0	2	0	8	0	**10**
	seem	0	0	0	1	0	1	**2**
	watch	0	0	0	0	1	0	**1**
Others	come about	0	0	0	0	1	0	**1**
	discover	0	0	0	0	2	0	**2**
	happen	0	0	0	0	2	0	**2**
	make clear	0	0	0	0	1	0	**1**
	offer	0	0	0	2	0	0	**2**
	pray	1	0	0	0	0	0	**1**
	strike	0	0	0	1	0	0	**1**
	reveal	0	0	0	1	0	0	**1**
	write	1	1	0	0	0	0	**2**
Non-ARG		0	0	0	0	6	0	**6**
Tot.		**3**	**9**	**2**	**17**	**48**	**1**	**80**

Table 15. Semantic categories of verbs introducing argumental complementisers in the corpus.

In conclusion, while there are still many issues to be addressed with reference to argumental complementisers, consecutive conjunctions seem to be relatively unproblematic, in that their grammaticalisation is considered to be an extension of the meaning of the complementiser, when in the contexts of a consecutive construction, such as *so*+ADJ/ADV...*that*.

4.4.1. East Germanic

As reported in Table 16, Gothic argumental and consecutive clauses were headed by the relative particle *ei* – which was occasionally compounded with other forms, mainly taken from the demonstrative *sa*, *so*, *þata*.

However, while in a restricted number of instances (in the corpus there are only two) this compound was indeed inflected and corresponded to the relative pronoun *sa*+*ei* [cf. Section 4.3.1], the compound *þatei* is an invariant form (composed of the nominative neuter singular form of the demonstra-

tive *þata* and the particle *ei*) which seems to have been a subordinating conjunction in its own right (Krause and Slocum 2006). Similarly, the form *þei* was itself an invariant compound functioning as a subordinating conjunction, but no part of it seems to have derived from the demonstrative: in fact, its *þ*- stem (to which the particle *ei* was added) derived from the Proto-Germanic pronominal stem **þa*- (Keretchashvili 2018).

	Inflection			POS	Argumental Complementiser		Consecutive Conjunction		
					No.	Texts, Dating	No.	Texts, Dating	Tot.
ei	UNINFLECTED			Particle	3	*Gospels*, 4th cent.	6	*Gospels*, 4th cent.; *Skeireins* 5th cent.	9
þammei	DAT	N	SG	Relative Pronoun (> *saei*)	2	*Gospels*, 4th cent.	0	-	2
þatei	UNINFLECTED			Conjunction	9	*Gospels*, 4th cent.	0	-	9
þei	UNINFLECTED			Conjunction	0	-	3	*Gospels*, 4th cent.	3
Tot.					14		9		23

Table 16. Distribution of Gothic conjunctions in the corpus.

Based on this account, we can thus deduce that in Gothic the simple demonstrative was never used as a conjunction, which clearly indicates that it never did grammaticalise into either of the forms we are considering in this chapter (i.e. argumental complementiser and consecutive conjunction).

Nonetheless, we may notice a certain degree of similarity between relative and argumental clauses, as they are both headed by conjunctions or pronouns which contain the relative marker *ei*.

4.4.2. West Germanic
As for the Old English argumental and consecutive clauses, they seem to have all been headed by an invariant form of the demonstrative *sē, sēo, ðæt* – and,

more specifically, the nominative neuter singular form ðæt, whose alternative orthography *þæt* is also attested, and often preferred [Table 17].

	Inflection	POS	No.	Argumental Complementiser Texts, Dating	No.	Consecutive Conjunction Texts, Dating	Tot.
ðæt			1	*The Voyages of Ohthere and Wulfstan* (848-899)	0	-	
þæt	UNINFLECTED	CONJ	16	*Anglo-Saxon Chronicle* (y. 755, 897); *The Voyages of Ohthere and Wulfstan* (848-899); *Exeter Book*, 10th cent.; *Beowulf*, 1000	6	The Venerable Bede (671-735); *The Voyages of Ohthere and Wulfstan* (848-899); *Exeter Book*, 10th cent.; *Beowulf*, 1000	23
	Tot.		17		6		

Table 17. Distribution of Old English conjunctions in the corpus.

Since all the forms instantiated in the corpus are thus derived from the Proto-Indo-European **só-, *sá-, *tód,* we will analyse some examples in order to verify the degree of grammaticalisation into a conjunction of the demonstrative itself.

(36) OE *For þon*   *cnyssað nū*
therefore strikes now
*heortan   geþōhtas, | **þæ** ic hēan   strēamas,*
heart.GEN.F.SG thought.ACC.M.PL   ARG I humble.ADJ stream.ACC.M.PL
sealtȳþa   gelāc | sylf   cunnige;
seawaves.GEN.F.PL tumult   myself test.1SG
'Therefore it strikes now the thoughts of the heart that I test myself the humble streams, the tumult of sea waves. '

[cf. Appendix, no. 31]

As it is evident from example (36), the Old English *þæt* does seem to be fully grammaticalised as an argumental complementiser, and – as mentioned at the beginning of this chapter – even when the matrix verb is not strictly a *verbum dicendi* or *sentiendi*, just like in the Present-Day variety of English.

However, while the grammaticalisation path of the argumental complementiser into a consecutive conjunction might be fairly straightforward

on principle, it does not seem to have occurred at this stage of Old English. Throughout the corpus, in fact, the conjunction is almost never found in the typical consecutive context (i.e. *so*+ADJ/ADV...*that*) and seems to cover more of an explicative function (37) – thus resulting as something which is still in between an argumental complementiser and a proper consecutive conjunction (as a matter of fact, in the use of the demonstrative we can still identify a trace of the *diptyque inverse* correlative structure mentioned in Section 2). On the other hand, in the one instance in which it is accompanied by the construction *so*+ADJ, it does indeed function as a fully fledged consecutive conjunction (38).

(37) OE *Þā hē þæt   þā   sumre tīde dyde,*
     and he DEM.ACC.N.SG then one   time did.3SG
     þæt hē forlēt þæt   hūs   þæs   gebēorscipes [...]
     CONJ he left ART.ACC.N.SG house ART.GEN.M.SG feast.GEN.M.SG
     'And then one time he did this – that he left the house of the feast.'
                                                    [cf. Appendix, no. 30]

(38) OE *For þon nis   þæs mōdwlonc mon ofer eorþan, [...]*
     because is.not so   proud.ADJ man on   earth
     þæt hē ā   his sǣfōre | sorge   næbbe,
     CONJ he never his sea.travel anxiety has
     'because not a man on earth is so proud that he never has his sea travel anxiety.'
                                                    [cf. Appendix, no. 31]

As we have already seen in the previous chapter [cf. Section 4.3.2], the Old English *þæt* went through a vocalic sound change (*æ* > *a*) and, over the Middle English period transformed first into *þat* and then into the Present-Day form *that.* On the other hand, the fact that the Middle English relative complementiser took the form of *þat*, as well, while the particle *þe* was dropped (it lost the relative marker function, but transformed into the definite article) seem to imply that the relative did grammaticalise afterwards and on the basis of the grammaticalisation of the argumental complementiser.

Furthermore, the fact that the relative *d*-Pronoun did not derive directly from the demonstrative seems to be confirmed by the forms used in Old High

German, as well. The Old High German demonstratives and complementisers were in fact surface-identical (i.e. *daz*, *thaz*), while the relative had a slightly different form (i.e. *der*, *ther*). In addition, as we can see, Old High German did preserve both a common gender (for masculine and feminine) and neuter form in the complementiser, which could explain why the supposedly derived relative had an inflection, as well. On the other hand, as we know, the Middle English relative *þat* did not, probably because of its derivation from an invariable complementiser such as *þæt* itself.

As for the other West Germanic languages, just like in Old English, the Old Frisian complementiser *thet* and the homographic relative correspond to the nominative neuter singular form of the demonstrative *thi, thiu, thet*. The same can also be said for the Old Low Franconian demonstrative *thie*, *that* and the conjunction *that*, as well as for the Old Saxon *that*, which was used as the only form from the beginning.

4.4.3. North Germanic

As we can see from Table 17, Old Norse argumental complementisers and consecutive conjunctions were generally expressed by the conjunction *at*, which could, however, also be compounded with an adverb to specify its meaning (e.g. the adverb *svát* gives the subordinate a consecutive meaning).

Among the instances included in the corpus, the sole exception is the use of the relative particle *sem*, but Krause and Slocum (2005) do mention that the relative particles (especially *er*) and the conjunction at did at times overlap in their function. However, the interchangeability between the two functions and the relative forms could be just the other side of the same phenomenon which we have found in Gothic, as well as in some of the early West Germanic languages, i.e. the affinity between relative and argumental complementisers. While in Old Norse, in fact, the two functions could be expressed alternatively by *at*, *er* or *sem*, in the other languages instantiated the same element could be used to render both functions.

	Inflec-tion	POS	**Argumental Complementiser**		**Consecutive Conjunction**		Tot.
			No.	Texts, Dating	No.	Texts, Dating	
at	UNIN-FLECT-ED	Conjunc-tion	48	*Battle of Stamford Bridge* (y. 1066); *Íslendingabók* (1067-1148); *Edda* (1179-1241); *Bren-nu-Njáls Saga* (1250-1275); *Poetic Edda*, 1270s; *Egils Saga*, late 14th cent.; *Grettis Saga*, 14th cent.; *Hrólfs saga*, late 14th cent.	26	*Battle of Stam-ford Bridge* (y. 1066); *Edda* (1179-1241); *Po-etic Edda*, 1270s; *Egils Saga*, late 14th cent.; *Grettis Saga*, 14th cent.; *Hrólfs saga*, late 14th cent.	74
svát		Adverb (*svá*) + Conjunc-tion (*at*)	0	-	2	*Poetic Edda*, 1270s	2
sem[40]	UNIN-FLECT-ED	Relative Particle	₮	*Hrólfs saga*, late 14th cent.	0	-	₮
	Tot.		48	28			76

Table 18. Distribution of Old Norse conjunctions in the corpus.

As example (39) and (40) show, the conjunction *at* seems to be fully grammat-icalised as both an argumental complementiser and, unlike Old English, also as a consecutive conjunction. In sentence (40), as in many other examples throughout the corpus, the conjunction *at* is, in fact, preceded by the adverb *svá*+ADJ/ADV and, at least based on the instances found in the corpus, this structure seems to have been far more common in Old Norse than in Old

---

[40] This form does not derive from the Proto-Indo-European *só-, *sá-, *tód and is therefore not relevant for the purposes of this study. This instance is not included in the sum total reported in the last row of Table 18.

English – thus probably implying a greater degree of grammaticalisation in the first.

(39) OE Óláfr *inn Digri* <u>*gørði skýrt*</u> ***at*** *hverr maðr skyldi*
       Olaf the thick made definite ARG each man should
       gjalda konungi hálfa mǫrk.
       pay king.DAT.M.SG half mark
       'Olaf the Thick made definite that each man should pay half mark to the king.'

                                                       [cf. Appendix, no. 32]

(40) OE *[...] ok* <u>*svá mikit*</u> *gerðisk af því* ***at*** *þeir vildu*
         and so far went with dem.DAT.M.PL CONJ they desired
       eigi nefna guð.
       no.longer mention god
       '[...] And went so far with this, that they no longer desired to mention God.'

                                                       [cf. Appendix, no. 33]

As already anticipated [cf. Section 2], however, it is argued that *at* might not really derive from the Proto-Indo-European **só-, *sá-, *tód* – the reasons behind this uncertainty being the following:
- While in the West Germanic languages one single form (with very minor orthographic differences, e.g. the German *das*/*dass*) is used to express all the functions we are taking into account, in all the Present-Day Scandinavian languages, there is some formal difference between the language-specific demonstrative and complementiser;
- An homographic preposition and infinitive conjunction (meaning *to*, *by* and *near*, and deriving from the Proto-Indo-European preposition **ad*) exists in all the Present-Day Scandinavian languages – and it is argued that this could be the true origin of *at*;
- Assuming, as we previously have [cf. Section 2], that the relatives were grammaticalised directly from the demonstrative and thus before the argumental complementisers, it would be hardly acceptable to instantiate that the North Germanic languages had skipped (also according to the corpus data) the intermediate passage, i.e. the grammaticalisation of the relative *d-Pronouns*, but then featured that of the complementisers.

This notwithstanding, from a phonetic point of view, the derivation of *at* from the Proto-Indo-European *só-, *sá-, *tód*, and then especially from the Proto-Germanic *þat*, would be perfectly acceptable if we hypothesise – as already discussed [cf. Section 2] – the loss of the initial consonant.

Furthermore, from a typological perspective, while we have proof from the West Germanic languages of complementisers deriving from demonstratives, we have no such evidence as far as the derivation of complementisers from the Proto-Indo-European preposition *ad* is concerned. In Old English, for instance, the reflex of *ad* has preserved its original meaning until today in the form of the preposition *at*.

In conclusion, should the argumental complementiser have indeed grammaticalised before the relative, the third reason for doubting of the origin of the Old Norse *at* would be easily discredited: it would, in fact, be more plausible to believe that the North Germanic languages might have missed the final step of the grammaticalisation process, rather than thinking they might have just skipped a previous one.

On that note, it is also worth stressing that, as already discussed, according to Vigfusson (1874:29), the archaic Icelandic argumental complementiser *ad* (i.e. the reflex of the Old Norse at from which the Present-Day Icelandic *að* derives) could also be used as a relative. This fact could either be a further proof of the affinity between the relative and the complementiser (as seen, for instance, between the Old Norse *er* and *at*) or an indicator that, at some point in the history of the Scandinavian languages, complementisers did grammaticalise into relatives – thus generating a competing form with the already established relative particle *er*. Either way, however, based on the evidence we have from Old Norse and the Present-Day North Germanic languages, we can maintain that, despite the initial instability of the two forms, the particle *er* and the conjunction *at* have prevailed in the end, respectively as relative marker and argumental complementiser.

5. Discussion

In an attempt to make sense of the results described in the previous section of this paper, this chapter will set the evidence gathered from the linguistic corpus against the existing literature presented in Section 2. In doing so, this

study will put forward some hypotheses that aim to explain the discrepancies between the empirical data and other scholars' accounts.

Taking into consideration both the earlier phases and the Present-Day stage of the Germanic languages and based on the analysis carried out in the previous chapter [cf. Section 4.1], we can conclude that, although the determiner function of the demonstrative developed at different times in each individual language (especially interesting is the already mentioned late grammaticalisation of determiners in Old Low Franconian and their incomplete establishment in Old Norse), nowadays they all feature demonstrative forms which can be used equally as pronouns or as adjectives.

Furthermore, we know that the use of the demonstrative as a third-person pronoun has been largely dropped over time and we can assume that Old English was one of the first Germanic languages to stop employing it as such – provided that it had ever been an established use at all.

While the distribution of demonstratives was still significantly unstable in the earliest stages of most Germanic languages (as already seen, for instance, the Old Norse demonstrative could be expressed through several competing forms, with slight differences in meaning), the origin of at least one of their demonstratives from the anaphoric Proto-Indo-European *só-, *sá-, *tód* is rather uncontroversial and, from the point of view of historical linguistics, the explanation of its evolution into a demonstrative determiner when accompanied by a noun [cf. Section 2] seems relatively unproblematic.

As for articles [cf. Section 4.2], based on the analysis of corpus data and on the contributions made by other scholars, like Crisma (2020) and Stroh-Wollin (2020), we may maintain that all three branches of the Germanic languages did undergo a grammaticalisation process which led their language-specific reflex of a Proto-Indo-European demonstrative (be it *só-, *sá-, *tód*, or a combination of other demonstrative stems, e.g. the Old Norse *inn, in, it* from the Proto-Indo-European *ko-+*eno-*) to become a definite article. On these grounds we can also claim that, albeit at different moments in time, Gothic, Old English and Old Norse all became languages with grammaticalised definiteness.

Although the definite articles systems of the Present-Day Germanic languages are fairly heterogeneous, even within the same branch (e.g. the Ger-

man "near-article" vs. the English and Dutch "discrete articles"), we can now maintain that both the West Germanic free-standing articles and the Scandinavian definite suffixes all originate from a deictic element.

With regards to the relative function, on the other hand, it appears that, according to the corpus data, no early Germanic language (except for Old Frisian in the fourteenth century) featured relative pronouns which had been directly grammaticalised from any of the reflexes of the demonstrative *só-, *sá-, *tód*[41] [cf. Section 4.3].

On these grounds, we might deduce that, this grammaticalisation path must have occurred at a later stage and only in some languages – signally in those in which the demonstrative is today also a *d-Pronomen*. Nonetheless, as already seen, Vigfusson's account (1874) leaves the question open as to whether the same patterns were undertaken by the Scandinavian languages, as well – thus implying that their *d-Pronomen* has gone lost over time.

Therefore, though *that* and its cognates do undoubtedly feature this function today (at least in the West Germanic languages) the origin and development, as well as the nature of the relative function, raise quite a few questions: was the relative originally a demonstrative? how and why did it become a relative? can the *that*-relative be described as a true relative pronoun, from a syntactic point of view? etc.

Although some of these questions have already been answered by other scholars [cf. Section 2], this study attempts to give its own corpus-based interpretation, while aiming to account for the Present-Day distribution and status of relatives across the Germanic languages.

As far as the origin of Proto-Indo-European relatives is concerned, several scenarios have been hypothesised, including the possibility that Proto-Indo-European did not have relative clauses at all and that any relative marker featured in the daughter languages was thus an innovation introduced after their split from the mother language. This notwithstanding, there is general agreement on the stems from which Indo-European relatives might have originated – either already in Proto-Indo-European, or at a later stage – and

---

[41] Old English features a relative particle whose form might have derived from the demonstrative [cf. Section 4.3.2]

these are the anaphoric **i-* and the interrogative-indefinite **kʷó-/*kʷí-*. While some scholars believe that one of the two derived forms was the original Proto-Indo-European relative pronoun and the other was an innovation independently made by some languages, Clackson (2007) maintains that both co-existed in Proto-Indo-European and were in complementary distribution with one another (Cotticelli-Kurras 2020).

As Quiles and López-Menchero (2011) argue, the Proto-Indo-European relative pronoun **ios, *iā, *iod* (from the relative stem **ió-*) gave origin to – among others – the Ancient Greek relative pronoun *ὅς, ἥ, ὅ* and the Proto-Germanic concessive/conditional conjunction **iaba, *iabai* (although it clearly did not retain the relative meaning), and was represented in every early Indo-European language except for Proto-Italic and possibly Anatolian[42], where it had probably gone lost (Quiles and López-Menchero 2011)[43]. It should also be noticed, however, that Hellenist scholars such as Monteil (1963) claim that, beside the relative pronoun deriving from **ió-*, Homeric Greek could also express the relative function by means of a pronoun deriving from the demonstrative **só-, *sá-, *tód*[44], possibly just like the Germanic reflexes of **tód* can be used as relative pronouns, as well.

As for the second relative pronoun, it derived, as already mentioned, from the interrogative-indefinite pronouns **qis*, **qid* (substantive form) and **qos*, **qā*, **qod* (adjective form) – and, ultimately, from the Proto-Indo-European **kwo/i-* stem. These forms, in turn, originated interrogative and indefinite pronouns such as the Greek *τίς, τί* (and the clitic *τις, τι* for the indefinite) and the Latin *quis, quid*, but also the Proto-Germanic root **hʷa-*, from which all the Present-Day Germanic interrogative *w-Pronomen* stem. Interestingly,

---

[42] The origin of the Hittite *ya-* from the Proto-Indo-European **jo-* is debated.

[43] Since it was present in almost every early Indo-European language, Quiles and López-Menchero (2011) believe that **ios, *iā, *iod* was indeed one of the two original Proto-Indo-European relative pronouns. It is more plausible to assume that it went lost in individual branches (i.e. Proto-Italian and possibly Anatolian) than to believe that all the other languages made the same innovation independently from one another.

44 Over time, however, the relative stemming from **ió-* prevailed and became standard in Classic Greek.

however, whereas the Greek *τίς, τί* maintained the indefinite semantic and did not acquire a relative function, the Latin *quis, quid* did at some point grammaticalise into a relative pronoun, through the *diptyque normal* construction.

Assuming, thus, that, as De Roberto (2010) claimed, Latin only introduced relatives through the grammaticalisation of its interrogative-indefinite pronoun[45], we should be able to instantiate an earlier stage, before the grammaticalisation occurred [cf. example (5) in Section 2]. In this example (41a) we can still see, in fact, the personal pronoun *is* anaphorically referring to the interrogative-indefinite pronoun *quis.*

(41) a. LAT ***Quis***   *bene* *amat,* <u>*is*</u>   *bene*   *castigat.*
                IND.M.SG well loves, PERS.M.3SG well chastises
                'Someone well loves. He well chastises.'
      b. LAT ***Qui***   *bene* *amat,* *bene* *castigat.*
                REL.M.SG well loves, well chastises
                'Who well loves, well chastises'
      c. EN ***That***   *loves* *well,* *well* *chastises.*
      d. EN ***Who***(*ever*) *loves* *well,* *well* *chastises.*

We can thus assume the following step to have been the one described by De Roberto (2010), in which the anaphoric element is absorbed by the interrogative-indefinite pronoun, which thus acquires a new form (from *quis, quid* to the relative *qui, quae, quod*) and becomes a relative conjunction with an optional [+indefiniteness] feature (41b).

Following on this example, and in the light of what has been discovered so far, however, it could prove useful to address the grammaticalisation of Germanic relatives and their distribution by putting forward a different hypothesis, which is partly contrasting with the account given in the literature review [cf. Section 2].

As we know, besides the *w-Pronomen* deriving from the interrogative-indefinite pronoun *k^wó-* (> *hwo-* > *wh-*), the Present-Day West Germanic languages also feature *d-Pronomen*, i.e. relatives which derive from a Pro-

---

[45] Assuming that Proto-Italic had indeed lost the relative deriving from *ios, *$iā$, *iod*, it is likely that Latin initially lacked relatives.

to-Germanic demonstrative and, ultimately, from the Proto-Indo-European anaphoric pronoun **só-, *sá-, *tód*. Remarkably, the same competition we can find today among the West Germanic relatives (i.e. between *w-Pronomen* and *d-Pronomen*) is supposed to have occurred in Proto-Indo-European, as well. As we have seen, in fact, Proto-Indo-European is assumed to have featured both an attributive-restrictive relative originating from the interrogative-indefinite pronoun and an appositive-explicative relative derivating from a demonstrative (Quiles and López-Menchero 2011).

While in Proto-Indo-European the two competing relatives seem to have been in a complementary distribution, however, as we have seen in Section 2, the status of the relatives in the Present-Day Germanic languages appears to be somewhat more complicated. For the moment, though, let us only consider that, in English, *w-Pronomen* may introduce either free or headed relative clauses, while *d-Pronomen* may only introduce headed ones.

On these grounds, if we were now to translate the previous sentence in English, we would have but one choice: the only grammatically acceptable way to preserve the *diptyque normal* structure is, in fact, interpreting the first clause as a free relative (36d) and thus render the Latin *qui* with the English *w-Pronoun who* (which has indeed an [+indefiniteness] feature).

Therefore, since the use of *d-Pronomen* in free relative clauses would be ungrammatical (41c), we could argue that, contrarily to previous assumptions deriving from the existing literature, the *diptyque normal* construction was not related to the grammaticalisation of demonstrative pronouns into relative *d-Pronomen*, but rather to the grammaticalisation of the Proto-Germanic (among other languages, e.g. Latin and Hittite) interrogative-indefinite pronoun (deriving from **hʷa-*) into a relative *w-Pronomen*.

Furthermore, this interpretation would be typologically more acceptable, since this grammaticalisation path would be identical to the one which occurred in Latin and the interrogative-indefinite pronoun would be the one prevailing over the anaphoric element in both the Italic and the Germanic branch.

On the other hand, based on the structure of the sentence, we could interpret the grammaticalisation of demonstrative pronouns into relative *d-Pronomen* as originating from the *diptyque inverse* construction and as mediated by the grammaticalisation of the argumental complementiser.

More specifically, let us take the grammaticalisation of the argumental complementiser as our starting point. In Latin (42a) [cf. example (6) in Section 2], we find a demonstrative at the end of the first clause and the already grammaticalised relative (through the *diptyque normal* construction) at the beginning of the second one, which the absorbed the demonstrative and was grammaticalised into an argumental complementiser.

(42) a. LAT *Legati renuntiaverunt <u>hoc,</u>   **quod**   Pompeium in*
              envoys reported   DEM.ACC.N.SG REL.ACC.N.SG Pompeius in
              potestate   haberent.
              power   had.3 PL
              'The envoys reported this: that they had Pompeius in (their) power.'
     b. LAT *Legati   renuntiaverunt **quod** Pompeium in   potestate haberent.*
              envoys reported   ARG   Pompeius in power   had.3PL
              'The envoys reported that they had Pompeius in (their) power.'
     c. EN   **The envoys reported <u>this,</u> | **that** they had Pompeius in their power.*
     d. EN   *The envoys reported | **that** they had Pompeius in their power.*

Conversely, according to the hypothesis that this study has just proposed, in the Germanic languages we would have two demonstratives (42c) – the second of which was supposedly the anaphoric element which, in the presence of specific categories of verbs like that of the *verba dicendi*, prevailed over the other demonstrative and was grammaticalised into an argumental complementiser (42d).

However, if we were to compare an argumental subordinate (42c) with a relative one featuring a *d-Pronomen* (43a), we would obtain two surface-identical structures – the only difference being that the relative pronoun is always preceded by its antecedent, while the argumental clause constitutes the argument of the predicate.

(43) a. EN *I have read <u>this</u> <u>(book)</u> | **that** you gave me.*
     b. DE *Ich habe <u>das</u>   <u>Buch</u> gelesen, **das**   du   mir*
             I   have ART.ACC.N.SG book read.PP DEM.ACC.N.SG you t o . m e
             gegeben   hast.
             given   have
             'I have read the book that you have given me.'
     c. NL *Ik heb <u>het</u>   <u>boek</u> gelezen **dat**   je   me   had   gegeven.*
             I   have ART.N.SG book read.PP DEM.N.SG you to.me have given.PP
             'I have read the book that you have given me.'

The striking similarity between the two constructions hints at their possible common origin and it is thus believable that the relative *d-Pronomen* could simply be an argumental complementiser which has been reanalysed when in the absence of specific categories of verbs (like *verba dicendi*, *sentiendi, iubendi,* etc.) and whose meaning has thus been extended.

Furthermore, although the structure of sentence (43a) seems to be modelled on the *diptyque inverse* construction (and contrarily to what has been previously outlined in Section 2), from this perspective, it would no longer be necessary to assume the absorption of an antecedent by the *d-Pronomen* in order to explain its reanalysis into a subordinator. According to this hypothesis, in fact, the *d-Pronomen* would merely be an extension of the argumental complementiser and would therefore have been a subordinator from the beginning.

Therefore, while rejecting Axel-Tober's account, this interpretation revisits the Neogrammarians' theory – according to which the argumental complementiser had grammaticalised directly from the demonstrative [cf. Section 2]. Contrarily to the traditional account however, this study proposes that the demonstrative from which the complementiser came had maintained its anaphoric nature and had thus always belonged to the second clause (42c).

As for the other main West Germanic languages, i.e. German (43b) and Dutch (43c), they seem to have taken the same grammaticalisation path.

From a cross-linguistic point of view, the *diptyque normal* thus appears to have originated relatives heading free relative clauses (i.e. indefinite relative pronouns or *w-Pronomen*), while the *diptyque inverse* construction has arguably been the model structure for the grammaticalisation of argumental complementisers and relatives introducing headed relative clauses (i.e. *d-Pronomen*).

While wondering why this only applies to West Germanic languages and not to the Scandinavian ones might lead us into analysing a false problem (the North Germanic languages – except for Danish – only use relative particles and do not have any *w-Pronomen* either), the grammaticalisation of the "Latin *d-Pronomen*" (i.e. the relative pronoun introducing a headed relative clause) is surely worth discussing.

Based on example (44), in fact, they seem to have followed the same grammaticalisation pattern we have described for the West Germanic *d-Pronomen*, in that the structure of the argumental subordinate (42b) seems to be identical to the relative one reported below.

(44) LAT *Magistra* <u>*hoc*</u>   *laudat* **quod**   *discipulus fecit.*
          Teacher   DEM.ACC.N.SG praises   REL.ACC.N.SG student   did
          'The teacher praises the thing (=this) which the student did.'

Unlike the West Germanic languages, however, the "Latin *d-Pronomen*" maintained the same surface-identical form as the indefinite relative pronoun probably because the argumental complementiser from which it came had, in turn, originated from the indefinite-relative itself.

Although we cannot be sure as to what the reason for the indefinite relative to grammaticalise into an argumental complementiser was, and we do not know for definite why the Germanic complementiser came, instead, from the demonstrative, it seems acceptable to conclude that the Germanic *d-Pronomen* have taken the form of the demonstrative because they came from the argumental complementiser, which had originated from the demonstrative, in the first place.

After all, as we have already seen [cf. Section 4.4.1] shared features between relative and argumental clauses could be found even in Gothic, as both clauses are headed by conjunctions or pronouns containing the relative marker *ei*. While, in fact, just like in Latin, neither form derives from the grammaticalisation of the demonstrative, we may interpret the similarity between relative and argumental heads as a typological feature, which is independent of the ultimate derivation of either elements.

As for the argumental complementiser itself, based on the corpus data analysed in the previous chapter, it seems to have grammaticalised directly from the demonstrative derived from the Proto-Indo-European **só-, *sá-, *tód* in all of the languages taken in consideration, except for Gothic – which, however, separated from Proto-Germanic quite early and did not have any significant contact with the other Germanic languages afterwards.

On the other hand, the fact that this grammaticalisation path did occur in the other Germanic languages at a stage when the *d-Pronouns* were not yet

grammaticalised seems to further confirm the hypothesis that the grammaticalisation process of relatives was a later extension of that of the argumental complementisers.

Furthermore, also thanks to typological evidence, we may be convinced that the Old Norse *at* did in fact derive from the Proto-Indo-European **só-, *sá-, *tód* [cf. Section 4.4.3].

Finally, as far as consecutive conjunctions are concerned, their grammaticalisation process, too, is considered to be another extension of that of the argumental complementisers. As such, however, the grammaticalisation degree of the conjunction used to head consecutive clauses seems to differ from language to language. In almost all of the Old English entries, in fact, the supposed consecutive clauses do lack their signature structure, i.e. *so*+ADJ/ADV... *that* – which, on the other hand seems to have already been quite common in Old Norse.

6. Conclusion

This study has attempted to examine the existing literature [cf. Section 2] in the light of the data that have been collected from the early stages of the main Germanic languages and gathered into a linguistic corpus [cf. Section 3, Appendix].

As it is evident from what has been said so far in this paper, an in-depth analysis of the corpus has often led towards new interpretations which are at least partly contrasting with the initial account described in the literature review. At the same time, the entire process of putting forward some hypotheses and trying to verify them, has given rise to new questions, which could represent possible future developments of this research.

Lang.	Demonstrative Pronouns		Demonstrative Determiners		Articles		Relatives		Conjunctions	ARG	CONS	Tot.
GOT	*sa, so, þata*	13	*sa, so, þata*	8	*sa, so, þata*	52	-	0	-	0	0	**73**
OE	*sē, sēo, ðæt*	13	*sē, sēo, ðæt*	12	*sē, sēo, ðæt*	11	*þe*(?)	~~31~~	*þæt*	17	1	**54**
OHG	*daz, thaz*	1	*daz, thaz*	0	*daz, thaz*	0	-	0	*daz, thaz*	1	0	**2**

OS	*that*	0	*that*	1	*that*	0	-	0	*that*	0	0	**1**
OF	*thi, thiu, thet*	2	*thi, thiu, thet*	0	*thi, thiu, thet*	2	*thi, thiu, thet*	1	*thet*	1	0	**6**
OLF	*thie*	1	-	0	-	0	-	0	*that*	0	0	**1**
ON	*sá, sú, þat*	83	*sá, sú, þat*	11	*sá, sú, þat*	13	-	0	*at*	48	28	**183**
Tot.		113		32		78		1		67	29	**320**

Table 19. Distribution of relevant forms per function in the corpus.[46]

Therefore, based on a careful study of the linguistic corpus [Table 19] and of the subject matter, we could tentatively describe the grammaticalisation paths of the Proto-Indo-European *só-, *sá-, *tód* across the Germanic languages as follows.

In agreement with the existing literature, we can maintain that the anaphoric demonstrative pronoun *só-, *sá-, *tód* became a demonstrative determiner in the earliest stages of most Germanic languages through a grammaticalisation process that allowed for it to be followed by a noun, despite its being a pronoun. Although each language probably developed this function at different times, based on the entries found in the corpus, we cannot find any such instance in Old Low Franconian and the representation of the determiner in Old Norse seems to have been unstable [cf. Section 4.1].

In each individual Germanic language, then, the demonstrative determiner deriving from *só-, *sá-, *tód* lost its [-deictic] feature and acquired the [+definiteness] one, thus beginning to function as a definiteness marker (which generally maintained the same form as the determiner). In Old Norse, however, this same process also involved a Proto-Scandinavian demonstrative pronoun which later produced Present-Day Scandinavian definite suffixes [cf. Section 4.2].

As for the following steps, this study distances itself from what has been

[46] This table provides an overview of the occurrences per grammaticalised function of the language-specific forms deriving from *só-, *sá-, *tód.* This table only includes the instances whose indicated function was confirmed by the corpus analysis described in Section 4. Once again, Gothic, Old English, and Old Norse are the most represented languages in this table, as, due to a specific methodological decision [cf. Section 1 and 3], they are the most represented languages in the corpus.

suggested by other scholars by proposing that the grammaticalisation of argumental complementisers might have taken place before that of the relatives. Arguably, in fact, in all of the Germanic languages taken into consideration – except for Gothic – the demonstrative **só-, *sá-, *tód* grammaticalised into an argumental complementiser through a *dipthyc inverse* construction [cf. Section 4.4, Section 5].

Furthermore, based on the evidence provided in Section 4.4.3, this paper supports the hypothesis that the Old Norse *at* and its Present-Day descendents [Table 20] are in fact all reflexes of the anaphoric **tód.*

According to this study, two more functions have then originated from the complementiser, through a *dipthyc inverse* structure, i.e. the relative marker function and that of the consecutive conjunction. These two grammaticalisation paths – interpretable as independent patterns of extension of the meaning of the complementiser – could be described as later processes which had generally not been established (especially as far as the relatives are concerned) in the time span covered by the linguistic corpus compiled for this study [cf. Section 4.3, Section 4.4].

Finally, with regard to the origin of relatives, this research claims that, contrarily to other existing accounts, the *dipthyc normal* structure only gave rise to the relative *w-Pronomen* – i.e. the type of realative pronoun which normally heads free relatives – while, as already mentioned, the relative *d-Pronomen* – which is normally used for headed relatives – derived from the extension of the meaning of the complementiser in the presence of an antecedent of the relative itself.

FUNCTION	EN	DE	NL	DA	NO	SV	IS
Demonstrative Pronoun/ Determiner	*that* *<*tód*	*das* *<*tód*	*dat* *<*tód*	*det* *<*tód*	*det* *<*tód*	*det* *<*tód*	*það* *<*tód*
Definite article	<u>the</u> *<*tód*	*das* *<*tód*	<u>het</u>[47] *<*tód*	*det<*tód* +suffi<DEM	suffix <DEM	*det<*tód* +suffi<DEM	suffix <DEM

[47] Cf. footnote no. 25 in [Section 4.2.2].

Conjunction (ARG, CONS)	*(that)*, so *that* <*tód	*(dass)*, so*dass* <*tód	*(dat)*, om-*dat*/ zo*dat* <*tód	*(at)*, så *at* <*tód	*(at)*, Slik *at*/ Så *at* <*tód	*(att)*, Så *att* <*tód	*(að)*, svo *að* <*tód
Relative	*that* <*tód	*das* <*tód	*dat* <*tód	RELATIVE PARTICLE	RELATIVE PARTICLE	RELATIVE PARTICLE	RELATIVE PARTICLE

Table 20. Present-Day relevant forms per function and derivation.[48]

Abbreviations

ACC	Accusative	ME	Middle English
ADJ	Adjective	N	Neuter
ADV	Adverb	NL	(Present-Day) Dutch
ARG	Argumental Complementiser	NO	(Present-Day) Norwegian
A-S	Anglo-Saxon	NOM	Nominative
CONJ	Conjunction	OE	Old English
CONS	Consecutive Conjunction	OF	Old Frisian
DA	(Present-Day) Danish	OHG	Old High German
DAT	Dative	OLF	Old Low Franconian
DE	(Present-Day) German	ON	Old Norse
DEM	Demonstrative Pronoun	OS	Old Saxon
DET	Determiner	PERS	Personal Pronoun
EN	(Present-Day) English	PG	Proto-Germanic
F	Feminine	PIE	Proto-Indo-European
GEN	Genitive	PL	Plural
GR	Ancient Greek	PP	Past Paticiple
IND	Indefinite Pronoun	PRSP	Present Participle
INF	Infinitive Conjunction	REL	Relative (Pronoun/Particle)
INSTR	Instrumental	SG	Singular
IS	(Present-Day) Icelandic	SV	(Present-Day) Swedish
IT	(Present-Day) Italian	SG	Singular
LAT	Latin	SV	(Present-Day) Swedish
M	Masculine		

[48] This table provides an overview of the Present-Day reflexes of **tód* and, where relevant, of the alternative forms used by the Scandinavian languages.

Bibliography

Primary Sources

(A) "RUNES: Forschungsprojekt der Akademie der Wissenschaften zu Göttingen", 2021, *RuneS*.

(B) Krause, Todd B. & Slocum, Jonathan, 2005, "Early Indo-European Online: Old Norse Online", *The University of Texas at Austin - Linguistics Research Center*.

(C) Krause, Todd B. & Slocum, Jonathan, 2006, "Early Indo-European Online: Gothic Online", *The University of Texas at Austin - Linguistics Research Center*.

(D) Robinson, Orrin W., 1992, *Old English And Its Closest Relatives*, Stanford, California, Stanford University Press.

(E) Slocum, Jonathan & Lehmann, Winfred P., 2007, "Early Indo-European Online: Old English Online". *The University of Texas at Austin - Linguistics Research Center*.

References

Allen, Cynthia L., 2019, "The Definite Article In Old English: Evidence From Ælfric'S Grammar", in Yáñez-Bouza N., Moore E., van Bergen, L. & Hollman, W. B. (eds.). *Categories, Constructions And Change In English Syntax*, Cambridge University Press, 130-146.

Axel-Tober, Katrin, 2017, "The Development of the Declarative Complementizer in German", *Language* 93(2), e29-e65.

Bacskai-Atkari, Julia, 2020, "English Relative Clauses In A Cross-Germanic Perspective", *Nordlyd* 44(1), 93-115.

Biberauer, Theresa & Roberts, Ian, 2017, "Parameter Setting", *The Cambridge Handbook of Hystorical Syntax*, 134-162.

Bisang, Walter, & Malchukov, Andrej, 2020, *Volume 1: Grammaticalization Scenarios From Europe And Asia*, Berlin, Walter de Gruyter & Co.

Bybee, Joan L., 2006, "Language Change and Universals", In Mairal, R. (ed), *Linguistic Universals*, Cambridge, Cambridge University Press, 179–194.

Campbell, Lyle, 1998, *Historical Linguistics – An Introduction*, Cambridge, Massachussets, The MIT Press.

Chantraine, Pierre, 1958, *Grammaire Homérique – Tome II: Syntaxe*, Paris, Librairie C. Klincksiek, 165.

Chomsky, Noam, 1988, *Current Issues in Linguistic Theory*, Berlin, De Gruyter, 14-15.

Clackson, James, 2007, *Indo-European Linguistics: An Introduction*, Cambridge, Cambridge University Press.

Cotticelli-Kurras, Paola, 2020, "Relative Clauses And Wh-Words In Ancient IE Languages: Etymology, Function, Distribution, Classification", Lecture, Università degli Studi di Verona.

Crisma, Paola, 2020, *Language Contact And The Early English Noun Phrase*, Canterano (RM), Aracne editrice.

Dahl, Östen, 2018, "Grammaticalization in the Languages of Europe", In In Narrog, H. & Heine, B. (eds.), *Grammaticalization From A Typological Perspective*, Oxford, Oxford University Press.

De Roberto, Elisa, 2010, "Correlative, Strutture", *Enciclopedia Dell'italiano*, Treccani.

Dunkel, George, 2014, *Lexikon der indogermanischen Partikeln und Pronominalstämme*, Heidelberg, Universitätsverlag Winter.

Fontana, Alessandro, 2022, *The Proto-Indo-European Anaphoric and Relative Pronouns: Analysis and Considerations.*

Fuß, Eric & Grewendorf, Günther, 2014, "Freie Relativsätze Mit D-Pronomen", *Zeitschrift Für Sprachwissenschaft* 33 (2) 165-214.

Harbert, Wayne Eugene, 2007, The Germanic Languages, Cambridge, Cambridge University Press.

Haspelmath, Martin, 2001, "The European Linguistic Area: Standard Average European". *Language Typology And Language Universals - Volume 2*, Berlin/New York, Walter De Gruyter.

Haudry, Jean, 1979, "L'antéposition de la Relative en Indo-européen", *La Linguistique* 15 (1), 101-10.

Heine, Bernd & Kuteva, Tania, 2010, "Grammaticalisation Theory as a Tool for Reconstructing Language Evolution", *The Oxford Handbook Of Language Evolution*, 512-527.

Hendery, Rachel, 2012, *Relative Clauses in Time and Space - A Case Study in the Methods of Diachronic Typology*, Amsterdam/Philadelphia, John Benjamins Publishing Company.

Keretchashvili, George, 2018, "Pronoun Paradigms". *Chrestomathy Of Gothic And Anglo-Saxon Written Records.*

Krause, Todd B., & Slocum, Jonathan, 2005, "Early Indo-European Online: Old Norse Online", *The University of Texas at Austin - Linguistics Research Center*.

Krause, Todd B., & Slocum, Jonathan, 2006, "Early Indo-European Online: Gothic Online", *The University of Texas at Austin - Linguistics Research Center*.

Kroch, Anthony, 2001, Syntactic Change, In Baltin, M. & Collins, C. (eds.), Hanbook of Syntax, Blackwell.

Lee, Soojin, 2006, "That Or Which?: The That's That Of Which Is Which".

Lehmann, Christian, 2002, "Thoughts on Grammaticalisation", *Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt*, Volume 9.

Lyons, Christopher, 1999, *Definiteness*, Cambridge, Cambridge University Press.

Minard, Armand, 1936, *La Subordination dans la Prose Védique*, Paris, Les Belles Lettres.

Monteil, Pierre, 1963, *La phrase relative en grec ancien*, Paris, Klincksieck.

Narrog, Heiko & Heine, Bernd, 2018, "Introduction: Typology and Grammaticalisation", In Narrog, H. & Heine, B. (eds.), *Grammaticalization From A Typological Perspective*, Oxford, Oxford University Press.

Pat-El, Na'ama, 2021, "Typological Approaches And Historical Linguistics", *The Handbook Of Historical Linguistics – Volume II*, John Wiley & Sons, Inc., 183-195.

Quiles, Carlos & López-Menchero, Fernando, 2011, A *Grammar Of Modern Indo-European*, Badajoz: Indo-European Language Association.

Radford, Andrew, 2004, Minimalist Syntax, Cambridge, Cambridge University Press, 9.

Roberts, Ian & Roussou, Anna, 2007, "Syntactic Change – A Minimalist Approach to Grammaticalization", *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 129 (3), 469-476.

Robinson, Orrin W., 1992, *Old English And Its Closest Relatives*. Stanford, California, Stanford University Press.

Rydberg-Cox, Jeff, 2021, *A Digital Tutorial For Ancient Greek*, Accessed 23/09/2021.

Sijs, Nicoline van der, 2010, "Dat (Voornaamwoord, Voegwoord)". *Etymologiebank.nl*, Instituut voor de Nederlandse Taal.

Skrzypek, Dominika, Piotrowska, Alicja & Jaworski, Rafał, n.d., *The Diachrony Of Definiteness In North Germanic*, Leiden, Brill.

Stroh-Wollin, Ulla, 2020, "Hinn And Hinn: Early Icelandic As The Clue To The History And Etymology Of Two Old Scandinavian Words", *Nordic Journal Of Linguistics* 43 (2), 205-228.

Velupillai, Viveka, 2012, *An Introduction To Linguistic Typology*, John Benjamins Publishing Company.

Vigfusson, Gudbrand, 1874, "AT", *An Icelandic-English Dictionary*, Oxford, The Clarendon Press.

Viti, Carlotta, 2013, "Forms and Functions of Subordination in Indo-European", *Historische Sprachforschung / Historical Linguistics* 126, 89-117.

Viti, Carlotta, 2015, *Variation und Wandel in der Syntax der alten indogermanischen Sprachen*, Tübinger Beiträge zur Linguistik (TBL).

Zimmermann, Richard, 2012, "Grammar Competition In Old English Relative Clauses", Lecture, Université de Genève.

Appendix – Corpus selection
Key:
No = Progressive number of entry
L = Language
Or. = Origin
SF = Syntactic function
Inf. = Inflection
So. = Source [cf. Primary Sources in Bibliography]
bf. = before

No	L	Entry	Text	Or.	SF	Inf.	Translation	So.
1	GOT	*Warþ þan in dagans* **jainans***, urrann gagrefts fram kaisara Agustau, gameljan allana midjungard.*	*Gospels* (4th cent.)	*jáins*	DET	ACC. M.PL	And it came to pass in those days, that there went out a decree from Caesar Augustus, that all the world should be taxed.	C
2	GOT	*Warþ þan in dagans jainans, urrann gagrefts fram kaisara Agustau, gameljan allana midjungard.* **Soh** *þan gilstrameleins frumista warþ at [wisandin kindina Swriais] raginondin Saurim Kwreinaiau.*	*Gospels* (4th cent.)	*sa, so, þata*	DET	NOM. F.SG	And this taxing was first made when Cyrenius was governor of Syria.	C
3	GOT	**sa** *ist sunus meins sa liuba, þamma hausjaiþ*	*Gospels* (4th cent.)	*sa, so, þata*	DEM	NOM. M.SG	This is my beloved Son: hear him.	C
4	GOT	*jah þatei gasaƕ jag gahausida* **þata** *weitwodeiþ*	*Skeireins* (5th cent.)	*sa, so, þata*	DEM	ACC. N.SG	and what he saw and heard, that he testifies	C
5	OE	*Næfde se here, Godes þonces, Angelcyn ealles forswiðe gebrocod, ac hīe wǣron micle swīþor gebrocede on* **þǣm** *þrim gēarum mid cēapes cwilde ond monna;*	*Anglo-Saxon Chronicles* (Year 897)	*sē, sēo, ðæt*	DET	DAT. PL	The enemy had not, by the mercy of God, entirely crushed the English altogether, but they were afflicted much more in those three years by pestilence of cattle and of men;	E
6	OE	*ealles swīþost mid* **þǣm** *þæt manige þāra sēlestena cynges þēna þe þǣr on londe wǣron forðfērdon on þǣm þrym gēarum.*	*Anglo-Saxon Chronicles* (Year 897)	*sē, sēo, ðæt*	PERS	NOM. N.SG	most of all among them many of the best of the king's thanes who were there in the land died within those three years.	E

NO	L	ENTRY	TEXT	OR.	INF.	TRANSLATION	SO.
7	ON	*En þar galzk stundum meira, en stundum minna, unz Óláfr inn Digri gørði skýrt at hverr maðr skyldi gjalda konungi hálfa mǫrk, sá er fœri á miðli Norvegs ok Íslands, nema konur eða **þeir** menn er hann næmi frá.*	*On the Settling of Iceland* (Ari Þorgilsson 1067-1148)	*hann, hon, þat*	DET	NOM. M.PL	Sometimes more were paid, sometimes less, until Olaf the Thick made definite that each man should pay the king a half mark, whoever would travel between Norway and Iceland, except women or those men whom he should exempt.	B
8	ON	*Fyrir **þá** sǫk*	*Snorra Edda* (Snorri Sturluson 1179-1241)	*sá, sú, þat*	DET	ACC. F.SG	For that reason	B
9	OLF	*ginathi in uuarheide sina uue sal **thia** suocan?*	*Psalm 60* (n.d.)	–	DEM	–	Who shall question his grace and truth?	D

Table 21. Syntactic function: Demonstrative

NO	L	ENTRY	TEXT	OR.	INF.	TRANSLATION	SO.
10	ON	*watehali**hino**hor(na) | hahaskaþihaþuligi*	*Older Fuþark* (2nd–8th cent.)	*hinn*	–	may the horn wet this stone. cut the aftermath. may the mowing lie.	A
11	GOT	*wasuh þan neƕa pasxa, **so** dulþs Iudaie.*	*Gospels* (4th cent.)	*sa, sō, **þata***	NOM. F.SG	was then nigh the passover, the feast of the Jews	C
12	OE	*Þā hē þæt þā sumre tīde dyde, þæt hē forlēt **þæt** hūs þæs gebēorscipes, and ūt wæs gongende tō nēata scypene, þāra heord him wæs þǣre nihte beboden,*	Bede's Account of the Poet Caedmon (671-735)	*sē, sēo, ðæt*	ACC. N.SG	Then one time he did this, so that he left the house of the feast and was going out to the cattle shed (their care was entrusted to him for the night).	E

13	ON	*Þá mǽllti Haraldr Ænghla konongr viðr Norðmenn þá er með hanum váro, «Kenndo þér þenn hinn myckla meðr **þæim** blá kyrtli oc hin faghra hialm, er þer skaut sér af hestinum frem?*	*Battle of Stamford Bridge* (1066)	*sá, sú, þat*	DAT. M.SG	Then Harald, king of the Angles, spoke with the Norsemen who were with him: 'Do you know that noble man with the blue kirtle and impressive helmet, who launched himself off his horse?'	B
14	ME	*þe wrecce men of **þe** land*	*Peterborough Chronicle* (1137)	*þe*	-	the wretched men of the land	-
15	ME	***þe** wrecce men of þe land*	*Peterborough Chronicle* (1137)	*þe*	-	the wretched men of the land	-

Table 22. Syntactic function: Article

No	L	Entry	Text	Or.	Inf.	Translation	So.
16	GOT	*usgeisnodedun þan allai **þai** hausjandans is ana frodein jah andawaurdjam is.*	*Gospels* (4th cent.)	*sa, so, þata*	NOM. M.PL	And all that heard him were astonished at his understanding and answers.	C
17	GOT	*Iesu sokeiþ Nazoraiu **þana** ushramidan;*	*Gospels* (4th cent.)	*sa, so, þata*	ACC. M.SG.	Jesus of Nazareth which was crucified;	C
18	GOT	*unte jabai afletiþ mannam missadedins ize, afletiþ jah izwis atta izwar **sa** ufar himinam*	*Gospels* (4th cent.)	*sa, so, þata*	NOM. M.SG	For if ye forgive men their trespasses, your heavenly Father will also forgive you:	C
19	GOT	*jah qaþ im: izwis atgiban ist kunnan runa þiudangardjos gudis, iþ jainaim **þaim** uta in gajukom allata wairþiþ*	*Gospels* (4th cent.)	*sa, so, þata*	DAT. M.PL	And he said unto them: 'Unto you it is given to know the mystery of the kingdom of God, but unto them that (are) without in parables all these things are done'	C
20	GOT	*Atsaiƕiþ armaion izwara ni taujan in andwairþja manne du saiƕan im; aiþþau laun ni habaiþ fram attin izwaramma **þamma** in himinam.*	*Gospels* (4th cent.)	*sa, so, þata*	DAT. M.SG	Take heed that ye do not your alms before men, to be seen of them: otherwise ye have no reward of your Father which is in heaven.	C

21	GOT	*iþ þu þan bidjais, gagg in heþjon þeina jah galukands haurdai þeinai bidei du attin þeinamma **þamma** in fulhsnja, jah atta þeins saei saiƕiþ in fulhsnja, usgibiþ þus in bairhtein.*	*Gospels* (4th cent.)	*sa, so, þata*	DAT. M.SG	But thou, when thou prayest, enter into thy closet, and when thou hast shut thy door, pray to thy Father which is in secret; and thy Father which seeth in secret shall reward thee openly.	C
22	GOT	*ei sijai so arma-hairtiþa þeina in fulhsnja, jah atta þeins **saei** saiƕiþ in fulhsnja, usgibiþ þus in bairhtein.*	*Gospels* (4th cent.)	*sa+ei*	NOM. M.SG	But thou, when thou prayest, enter into thy closet, and when thou hast shut thy door, pray to thy Father which is in secret; and thy Father which seeth in secret shall reward thee openly.	C
23	OE	*alle būtan ānum, sē wæs þæs aldormonnes godsunu; ond hē his feorh generede, ond þēah hē wæs oft gewundad*	*Anglo-Saxon Chronicles* (Year 755)	*sē, sēo, ðæt*	NOM. M.SG	all but one, who was his nobleman's godson; and he saved his life, though he was much wounded.	E
24	OE	*Þonne is ān port on sūðeweardum þǣm lande, þone man hǣt Scīringes hēal.*	*Voyages of Ohthere and Wulfstan* (bf. 899)	*sē, sēo, ðæt*	ACC. M.SG.	There is a port in the south of that land, which one calls Skiringssal.	E
25	OE	*Eal þæt his man āþer oððe ettan oððe erian mæg, þæt liþ wið ðā sǣ;*	*Voyages of Ohthere and Wulfstan* (bf. 899)	*sē, sēo, ðæt*	NOM. N.SG.	All that a man can either graze or plough extends alongside the sea;	E
26	OE	*Geworhton ðā | Wedra lēode hlǣw on hliðe, | **sē** wæs hēah ond brād, wæglīðendum | wīde gesȳne,*	*Beowulf* (1000)	*sē, sēo, ðæt*	NOM. PL.	the people of the weders built upon the hill a mound, which was high and wide, visible to sea-farers from afar,	E
27	OE	*and eall þæt to fǣsle | frea ælmihtig habban wolde | under hrof gefor to heora ætgifan, | swa him ælmihtig weroda drihten | þurh his word abead.*	*Genesis* (1000)	*sē, sēo, ðæt*	ACC. N.SG.	and all that the Lord Almighty would have for progeny. And he went under the roof as their provider, as the Almighty, the Lord of hosts, bade him by His word.	E

| 28 | OE | *he wǣs milde þam godum mannum þe god lufedon* | Peterborough Chronicle (1087) | *sē, sēo, ðæt* | DAT. PL. | he was gentle with those good men that love god | E |
| 29 | OF | *ther hi on eskriuin hede tha tian bodo, tha skolde hi lera tha Israheliska folke* | *Asegabook* (1300) | – | – | where he had written the ten commandments, which he should teach to the israelite folk | D |

<p align="center">Table 23. Syntactic function: Relative</p>

No	L	Entry	Text	Or.	SF	Translation	So.
30	OE	*Þā hē þæt þā sumre tīde dyde, þæt hē forlēt þæt hūs þæs gebēorscipes, and ūt wæs gongende tō nēata scypene, þāra heord him wæs þǣre nihte beboden,*	Bede's Account of the Poet Caedmon (671-735)	*þæt*	CONS	Then one time he did this, so that he left the house of the feast and was going out to the cattle shed (their care was entrusted to him for the night).	E
31	OE	*For þon cnyssað nū heortan geþōhtas, | þæt ic hēan strēamas, sealtȳþa gelāc | sylf cunnige;*	The Seafarer (10th cent.)	*þæt*	ARG	therefore it <u>strikes</u> now the thoughts of the heart, that i the humble streams, the tumult of sea waves myself should test	E
32	ON	*En þar galzk stundum meira, en stundum minna, unz Óláfr inn Digri gørði skýrt at hverr maðr skyldi gjalda konungi hálfa mǫrk, sá er fœri á miðli Norvegs ok Íslands, nema konur eða þeir menn er hann næmi frá.*	On the Settling of Iceland (Ari Þorgilsson 1067-1148)	*at*	ARG	Sometimes more were paid, sometimes less, until Olaf the Thick <u>made definite</u> that each man should pay the king a half mark, whoever would travel between Norway and Iceland, except women or those men whom he should exempt.	B
33	ON	*ok svá mikit gerðisk af því at þeir vildu eigi nefna guð.*	Snorra Edda (Snorri Sturluson 1179-1241)	*at*	CONS	and this went <u>so far</u> that they no longer desired to speak the name of god.	B

<p align="center">Table 24. Syntactic function: Conjunction</p>