The emergence of the definite article: 
*ille* in competition with *ipse* in Late Latin

Anne Carlier and Walter De Mulder

1. The restructuring of the demonstrative system according to Harris and the role of subjectification

It is generally accepted that the demonstratives in Classical Latin are structured in a person-oriented system: *hic*, *iste* and *ille* are linked respectively to the first, the second and the third person. Their meaning has been traditionally defined in terms of physical location: *hic* denotes a referent located near the speaker, *iste* marks its location close to the hearer, whereas *ille* situates it in a location remote from both speaker and hearer. *Is* is unmarked with respect to this distinction and has mainly an anaphoric role.

According to Harris (1978:69), from Classical to Vulgar Latin, the demonstrative system underwent the shift represented in Table 1, whereby the identity marker *ipse* was recruited to fulfil a typically demonstrative function.

<table>
<thead>
<tr>
<th>DEMONSTRATIVES</th>
<th>IDENTITIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>“1” “2” “3”</td>
<td>“self” “same”</td>
</tr>
<tr>
<td>Classical Latin</td>
<td><em>hic</em> <em>iste</em> <em>ille</em> <em>is</em> <em>ipse</em> <em>Idem</em></td>
</tr>
<tr>
<td>Vulgar Latin</td>
<td><em>iste</em> <em>ipse</em> <em>ille</em> <em>met-ips-i-num</em></td>
</tr>
</tbody>
</table>

Harris’s proposal is essentially based upon a comparison between the Classical Latin system and the result of the evolution in different Romance languages, rather than upon an empirical analysis of Late Latin texts. For instance, *ille* is analyzed as inheriting the value of the distance-unmarked
is, because this hypothesis explains why it can be a forerunner of the definite article and of the personal pronoun in most Romance languages. The elimination of the demonstratives hic and is\(^2\) and the identity marker idem is based on the fact that none of them were retained in the Romance languages. As for iste, ipse and ille, their position in the Vulgar Latin demonstrative system is inferred from the fact that they gave rise in Spanish, Portuguese and Catalan to a three-term deictic system, represented in Table 2.

Table 2: Demonstratives derived from iste, ipse and ille in Spanish, Portuguese and Catalan (Harris, 1978:71-72)

<table>
<thead>
<tr>
<th>Vulgar Latin</th>
<th>(accu-)iste</th>
<th>(accu-)ipse</th>
<th>(accu-)ille</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>este</td>
<td>ese</td>
<td>aquel</td>
</tr>
<tr>
<td>Portuguese</td>
<td>este</td>
<td>ese</td>
<td>aquele</td>
</tr>
<tr>
<td>Catalan</td>
<td>aquest</td>
<td>aqueix</td>
<td>aquell</td>
</tr>
</tbody>
</table>

Traugott puts forward the hypothesis of a cline of intersubjectification (see Traugott, this volume, and Traugott and Dasher 2002:225):

Non-/ less-subjective - Subjective - Intersubjective
(Ideational / Contentful Interpersonal / Procedural)

The restructuring of the demonstrative system as it has been analyzed by Harris (1978) is mentioned by Traugott and Dasher (2002) as an example of subjectification: iste and ipse shift in the direction of the deictic center and become respectively the first and the second person demonstrative. Ipse is reinforced in its function of identity marker by met\(^3\) and by the superlative suffix –imum (Harris 1980:69).

Harris’s representation of the restructuring of the demonstrative system, from Classical to Late Latin, raises at least three questions. Firstly, it could give the impression that the demonstratives in Classical Latin have a non-subjective meaning. However, iste and ille already exhibit a subjectified meaning in Classical Latin: the distal demonstrative ille expresses a respectful attitude of the speaker / writer towards the referent (Socrates ille “the illustrious Socrates”), whereas the proximal demonstrative iste is deprecative (iste homo “this dishonorable man”). In this sense, ille and iste partake of social deixis, at least when they refer to persons (Traugott and Dasher 2002:226 ff.).
The second question is raised by Harris’s analysis of the identity marker *ipse*. According to Harris (1978), the three-term deictic demonstrative system in Late Latin was rebuilt by the recruitment of the identity marker *ipse* for the role of the second person demonstrative. However, this representation of the demonstrative system in Late Latin cannot be invoked as the common source of the demonstrative system in all Romance languages, since, as is also noted by Harris (1978), several Romance languages do not have a descendant of *ipse* within their demonstrative system. In Old French, for instance, there is no ternary but rather a binary opposition between the proximal demonstrative *cist* (< reinforced form *ecce istu(m)*) and the distal demonstrative *cil* (< reinforced form *ecce illu(m)*). (Standard) Italian does not have a demonstrative derived from *ipse* either. Moreover, *ipse* provides the source of the definite article in Romance languages such as the Balearic variety of Catalan and Sardinian (Aebisher 1948; Selig 1992; Bauer 2007). Yet the textual function characteristic of the definite article is not directly related to the second person.

Thirdly, the semantic shift postulated for *ille* and its relationship to the disappearance of *is* deserve some further analysis. According to Harris (1980:78), “had *is* survived, it would surely have been the source of both the definite articles and the third person pronouns in Romance”. The remote demonstrative *ille* inherited this double destiny. This link is however not straightforward. Even though the Classical Latin *is*, which is unmarked as to distance, is almost restricted to anaphora, it does not have the monopoly on the marking of anaphoric relationships: the distance-marked demonstratives *hic*, *iste* and *ille* and the identity expressions *ipse* and *idem* are also used in this role (Menge, 2000:§§ 68-81; Pinkster 1996). Hence, the decrease in frequency of *is* as an anaphoric device, from Classical to Late Latin, cannot be assumed to automatically and immediately result in its replacement by *ille*.

This paper will be devoted to the emergence of the definite article in the transition from Latin to French, and will reexamine the semantic nature of *ille* and its relationship to *ipse* on their pathway to the status of definite article, both from a theoretical and an empirical viewpoint. On the theoretical side, we will investigate whether the notion of (inter)subjectification proposed by Traugott and Dasher (2002) and Traugott (1995; 2003; this volume) is relevant to explain the emergence of the definite article out of the demonstrative in Late Latin. The empirical basis of this study is a historical narrative entitled the *Chronicle of the Merovingian Times*, written in Gaul during the 7th Century by a certain Fredegarius and continued during
the 8th Century. We opted for a narrative because in this type of text discourse participants are introduced and tracked as they are involved in a sequence of events. A narrative is thus appropriate to illustrate a variety of anaphoric relationships that could be expressed by the incipient definite article. The very choice of Fredegarius’ *Chronicle* and its continuations as our corpus is motivated by the period of its composition: it goes back to a pivotal period preceding the divorce between Latin and Romance languages. As has been noted by Banniard (1992) and Renzi (2000:187), from 650 to 750, learned written style and oral speech are growing more distant without however being perceived as different languages. Hence, the often despised “barbaric” Latin of the Fredegarius’ *Chronicle* written during this period and its numerous “errors” against the classical standards should be viewed as an attempt to represent the new sounds, morphology and syntax of the emerging Romance language and to keep in this way the text accessible for an illiterate public. These very strong interferences with oral speech will diminish from 750 on, when the Carolingian reform recommends a readjustment to the Classical norms and leads in this way to an autonomous transcription of the oral speech.

From a statistical viewpoint, our corpus does not confirm the evolutions sketched in Table 1. Adnominal *is* is indeed less frequent than in Classical Latin. But contrary to what might be suggested by Table 1, *hic* is still often used, whereas *iste* is exceptional. The most striking feature, however, is the high frequency of *ipse*. Given the historiographical nature of the text, the meaning of *ipse* cannot be analyzed as the second person demonstrative, resulting from a process of subjectification. Rather, *ipse* extends its frequency as a marker of anaphora and competes in this role with *ille*.

### 2. The development from the demonstrative to the definite article in Late Latin

In this section, we will examine several hypotheses that have been proposed to account for the grammaticalization process transforming the third person demonstrative *ille* into the definite article. Firstly, the semantic shift from demonstrative to definite article has been described as a process of semantic weakening or “bleaching” (§ 2.1). According to a second hypothesis, the demonstrative becomes a definite article when it is selected not to identify the referent but to express the speaker’s subjective attitude towards the referent (§ 2.2). The viewpoint that will be defended here is
that the driving force that triggers the grammaticalization process leading to the definite article is intersubjectification (§ 2.3).

2.1. Semantic weakening or “bleaching”

2.1.1. The hypothesis

According to Harris (1978:67-68), a demonstrative determiner serves two interrelated functions: “to specify one particular item (as opposed to an indefinite number of alternatives) and to mark the proximity of that item, whether that proximity be spatial, temporal, or simply psychological […]; a definite article, on the other hand, when fulfilling its original semantic function […],7 has only the first of these two functions, namely to specify […] one particular referent as opposed to all others”. In this perspective, the demonstrative turns into a definite article when it loses the information relating to the proximity of the referent of the nominal expression, i.e. its deictic feature (quoted from Harris 1978:70; also see Hawkins 2004:84).

Lyons (1999:161, 332) argues that the shift from demonstrative to definite article is initiated in those contexts of use where the referent is either directly perceived in the physical surroundings or straightforwardly recalled from the preceding discourse. Deictic information is redundant in these cases because the referent is immediately accessible and so semantic weakening or “bleaching” can occur. The new definite article is assumed to be restricted to this range of uses in the initial stage and to expand to other contexts of use in a later stage.

2.1.2. Evaluation of the hypothesis

It is generally agreed that the definite article, in comparison with its demonstrative source, lacks the deictic meaning component. Several facts remain nevertheless unexplained in the framework of Lyons’s hypothesis.

2.1.2.1. In the contexts of use where the shift from demonstrative to definite article is initiated according to Lyons’s (1999) hypothesis, deictic information is redundant, so that proximal and distal demonstratives alternate without semantic distinction. This free alternation can be illustrated for the
situational use by the French examples (1) and for the textual use by the Latin examples (2).

(1)  a. *Ne renverse pas ce seau-ci!*
    b. *Ne renverse pas ce seau-là!*
    Don’t knock over *this / that bucket!*

(2)  a. *Lucca castrum dirigunt, atque funditus subvertunt, custodes illius castri capiunt* (Continuations § 25)
    They go to the *fort* of Loches, they raze it to the ground and take prisoner the guardians of *that fort*.
    b. *Radulfus haec cernens, castrum lignis monitum in quodam montem super Vnestrude fluvio in Toringia construens, exercitum undique quantum plus potuit collegens, cum uxorem et liberis in hunc castrum ad se definsandum stabilibit.* (Fredegarius § 87)
    Observing this, Radulf put up a *fort* protected with a wooden stockade on a rise above the banks of the Unstrut, in Thuringia, and when he had assembled from everywhere an army as big as he could, he established himself with his wife and children within *this fort* to withstand a siege.

As pointed out by Greenberg (1978:61) and Himmelmann (1997:96-98) from a typological perspective, the most common source of the definite article is the distal demonstrative or the third person demonstrative. If there is free alternation between proximal and distal demonstratives in the contexts of use where the shift from demonstrative is initiated, how can we account for this widespread tendency?

2.1.2.2. As has been observed by several empirical studies of Late Latin corpora (Renzi 1976:29; Orlandini 1981; Calboli 1990; Selig 1992:165; Fruyt 2003:109), there is one context where *ille* occurs frequently, without any serious competition of other determiners*: *ille* is used when the nominal expression refers to a referent not previously mentioned in the text, whose identity is established by a relative clause. This definite first mention use of *ille* is illustrated by (3).

(3)  *[…] homines illos quos Waiofarius ad defendendam ipsam ciuitatem dimiserat clementiam sue pietatis absoluit dimissisque reuersi sunt ad propria.* (Continuaciones, § 43)
Of his goodness he showed mercy to the (litt. that) men that Waiofar left there to defend the city, and dismissed them to go off home.

The association of *ille* with a relative clause was so frequent in Late Latin that it led according to Herman (2000) to a morphological contamination of *ille* by the relative pronouns *qui / cui* (nominative singular: *illi*; dative: *illui, lui, ilaei, lei*).

Selig (1992:169) observes pertinently that the use of *ille* in this type of context favors the weakening of the “demonstrative force” and that it foreshadows its later article function. Contrary to what can be expected in the framework of Lyons’s hypothesis, there is however in this context of use no highly accessible referent available, neither in the immediate situation nor in the preceding context.

2.1.2.3. It is widely accepted that the characteristic context of use of the definite article is the associative anaphor (Kleiber 1992, Himmelmann 1996, 1997), where the definite article gives the instruction to retrieve the referent indirectly, by activating a frame of accessible knowledge which the referent is associated with (Kleiber 1992, Epstein 2002). The demonstrative, by contrast, typically focuses the hearer’s attention on the referent directly, without considering its structural links with a frame of accessible knowledge. Neither Lyons’s hypothesis nor the hypothesis deriving the definite article from an anaphoric demonstrative referring to a previously mentioned discourse participant do account for this meaning shift from direct reference to indirect reference, which is crucial in the development of the definite article.

2.1.3. Pragmatic strengthening and semantic weakening

Although it is true that the definite article, in comparison with its demonstrative source, has lost the deictic meaning component, more attention should be paid to the factors that set off the grammaticalization process to understand the evolutionary pathway from distal demonstrative to definite article. With respect to the early stages of grammaticalization, Hopper and Traugott (2003:75-76) point out that “meaning changes and the cognitive strategies that motivate them […] are crucially linked to expressivity.” According to the same authors (2003:94), it is only at later stages, as gram-
maticalization continues and forms become more routinized, that meaning loss or “bleaching” typically occurs.

2.2. Subjectification

Traugott (this volume) puts forward the view that there is subjectification when the strengthening of the speaker’s pragmatic viewpoint is encoded as a semantic feature. A study of the first stages of the development of the definite article exploiting the notion of subjectivity is conducted by Epstein (1993, 1994a-b, 1995, 2001, 2002). Adopting the hypothesis of Traugott and König (1991:191) according to which “especially in the early stages of grammaticalization, new meanings may be added to an element in the form of a “strengthening of the expression of speaker involvement” (Epstein 1994b:77), he argues that the foregrounding of this speaker involvement and of his subjective attitude towards the referent is the crucial factor in the emergence and evolution of the definite article.

2.2.1. The hypothesis

Considering that earlier studies on the definite article privilege a purely referential analysis, Epstein (1993, 1994a-b, 1995, 2001, 2002) claims that the use of the definite article is often motivated by expressive factors. The opposition between the referential and the expressive use of the definite article corresponds according to him to a difference of orientation: the referential use is hearer-oriented insofar as the definite article guides the hearer to the intended referent, whereas the expressive use is speaker-oriented and “is motivated by the speaker’s desire to portray the referent as discourse prominent” (Epstein 2001:186).

The expressive value of the definite article, in particular in the early stages of its development, arises according to Epstein (1993:129) from its demonstrative origin: the demonstrative being the linguistic equivalent of the act of pointing, it identifies an entity in the immediate situation of its utterance and acts at the same time as a marker of attention11.

From a diachronic perspective, Epstein (2001:186) argues that expressivity is responsible for the shift from demonstrative to definite article: the development of the demonstrative *ille* into a definite article is initiated when speakers begin to choose – for expressive purposes – a demonstrative
in contexts where the identification of the referent normally does not re-
require it. This type of use thus allows the demonstrative to lose its referential
properties. An example is given in (4), where the content of the nominal
expression is sufficient to identify the bishop and where the demonstrative
is used purely for emphasis.

(4) Ductus itaque sanctus Eugenius ad regem, cum illo Arrianorum episcopo pro fide catholica decertavit. (Gregory of Tours, quoted by Trager 1932:173, cf. Epstein 2001)
And so the holy Eugenius was led before the king, and disputed with the bishop [lit. “that bishop”] of the Arians in defense of the catholic faith.

2.2.2. Evaluation of the hypothesis

2.2.2.1. The Fredegarius’ corpus confirms the role of ille as a marker of
discourse prominence in Late Latin. Ille appears especially for highlighting
a protagonist or an important object at the moment when it is involved in an
event that marks a turning point in the narration. Witness the use of ille in
the following excerpt, where the author tells how, after repeated military
successes of the Persians against the Greeks and their territorial expansion,
one moment of distraction of a Persian patrician causes the decisive defeat
of the Persians and the annexation of their Empire. The mention of this
patrician at the very moment when he is involved in this fatal event is pre-
ceded by ille, while the other anaphoric expressions referring to the same
referent are not preceded by any determiner.

(5) Aeraclius imperatur arma sumens, telam priliae et falange a suis
postergum preparatum relinquens, singolare certamen, ut nouos
Dauit procedit ad bellum. Emperator Persarum Cosdroes
patricium guidam ex suis quem fortissimum in prelio cernere
potuerat huius conuenenciae ad instar pro se contra Aeraglio ad
priliandum direxit. Cumque uterque cum aequitis hy duo
congressione priliae in inuicem propinquarint, Aeraglius ait ad
patricium, quem emperatore Persarum Cosdroae stembat, dixit:
“Sic conuenerat, ut singulare certamen priliandum debuissimus
configliere: quare postergum tuom alii secuntur?” Patricius ille
girans capud conspecere qui postergum eius uenerit, Aeraglius
The emperor Heraclius armed himself, left behind him his army drawn in fighting array, and advanced to the fray like a new David. But the Persian Emperor Chrosroes honored their pact by sending one of his patricians (litt.: a certain patrician), whose great value he knew, to fight in his place against Heraclius. Advancing on horseback to each other, Heraclius said to the patrician, whom he took for the Emperor Chrosroes: “Since we have agreed to single combat, why are those other warriors following behind you?” The patrician turned his head to see who was following him, whereupon in a flash Heraclius spurred his horse forward, drew his short sword and cut off the patrician’s head.

2.2.2.2. Although *ille*, in accordance with Epstein’s hypothesis, can have the role of discourse prominence marker, one could ask if this use has to be considered as *speaker-oriented* i.e. expressing a purely subjective judgment, whereas the referential use is assumed to be *hearer-oriented*. In fact, the speaker also plays an active role in acts of reference and seldom uses determiners only to express subjectivity. He rather tries to give the hearer the key to a correct and coherent interpretation of the discourse, not only by signalling the textual importance of a referent at a certain point of the discourse, but also by guiding the hearer to the intended referent. Hence, referential value and expressive value are not disjoint but complementary and most often closely linked. In our view, the emergence of the definite article has to be conceived from the same perspective: what motivates the development of the articles is not the expression of the speaker’s viewpoint on the referent as such but rather an interactional strategy between speaker and hearer to explicitly mark the way the referent has to be identified as well as the discourse status of the referent. A strict separation between referential and expressive use is therefore artificial.

2.2.2.3. According to Epstein (1993:127), *ille*’s role of discourse prominence marker originates in the deictic force of its demonstrative source. However, his analysis of the development of the definite article does not mention *ipse*, an important competitor in the early grammaticalization process leading to the definite article. Though not a demonstrative, *ipse* can also mark discourse prominence, on the basis of its meaning as identity marker. Hence, in the following example, *ipse* can be glossed as “himself” or “in person”.
In this year Caesarea, wife of Anauld, the Persian emperor, left her husband and came with four male and four female servants to the blessed John, bishop of Constantinople. She said that she was a private person and besought the blessed John to baptize her. She was baptized by the bishop in person, and the illustrious wife of the Emperor Maurice stood godmother to her.

A similar observation can be made for the early uses of the indefinite singular article *uns* in Old French: although devoid of deictic meaning, the incipient article *uns* only occurs in noun phrases denoting referents that have some prominence in the text, typically protagonists, as witnessed in example (7), whereas no determination is used in noun phrases referring to background entities or referents whose individuality is less important than their role or class membership (8).

(7) *Jadis avient que un lu pramist que char ne mangeriet, ceo dist, les quarante jurs de quareme;*  
It once happened that a wolf made the vow not to eat meat – as he said – during the forty days of Lent.  
(Marie de France, *Fables*, “Le loup et le mouton”, ed. C. Bruckner)

(8) *Il avint chose que un pere managier, qui n’avoit nul fiz, mais feme espouse soulement, si escrit ensi en son testament: “Se un fiz m’est engendré, ou plusors, cil me soit hoir.”* (J. d’Antioche, *La Rettori-que de Marc Tulle Cyerceron*, LXV)  
It happened that a head of the family, childless although married (lit.: who didn’t have a son, but only a wife), wrote the following words in his testament: “If one or several sons are born to me, they should be my heirs.”

In view of these facts, it is difficult to maintain that *ille* owes its role of prominence marker to its demonstrative origin alone. A different hypo-
thesis will be put forward here: in a language stage where zero determination is the rule for a certain type of nominal reference, in this case definite reference, the use of a demonstrative or another linguistic expression in an article-like role is a conscious strategy by which the speaker intends to externalize for the hearer the discourse status of a referent. It is normal that these explicit marks of text cohesion only bear on nominal expressions whose referent is important enough to have a certain textual permanence. Nevertheless, we agree with Epstein that in the case of *ille*, its demonstrative origin did contribute to its discourse-based function. The pathway from demonstrative to definite article is, however, conceived differently in the two approaches.

Epstein’s approach, definite articles evolve from demonstratives as markers of the speaker involvement. Invoking the following analysis of Trager (1932),

> It was an emotional, affective usage, which grew upon the language unconsciously until it became firmly fixed as an article … there is nothing at all of a logical analysis; on the contrary, they are wholly non-logical and subjective (Trager, 1932:54-55, quoted by Epstein, 2001)

Epstein (2001:186) maintains that the transition from the demonstrative *ille* to the definite article boils down to a process in which the referential value of *ille* resolves into an expressive value, initiated by the fact that some speakers begin to use *ille* for expressive purposes when the identification of the referent does not require it.

In our view, on the contrary, definite articles develop as explicit markers regulating speaker–hearer interaction. We will argue in section 2.3 that the discourse-based function assumed by *ipse* and *ille* in Late Latin does not eliminate their referential instruction but develops from this referential instruction by pragmatic strengthening and tends to be superposed on it rather than to replace it. This assumption not only allows us to explain the constraints imposed upon the use of *ille* and *ipse* in their new role in the speaker-hearer interaction (§ 2.3.1), but also provides a basis for understanding why *ille*, though less frequent than *ipse*, was eventually selected as the source of the definite article in most Romance languages (§ 2.3.2). We will finally show how the hypothesis of a cline of intersubjectification sheds a new light on the grammaticalization process leading from the third person demonstrative *ille* to the definite article (§ 2.3.3).
2.3. Intersubjectification

Late Latin innovates with respect to Classical Latin by the increasing use of referential markers to express definiteness and anaphoric relationships. Consider for example the following passage, where the emphasis on the identity of the referent is not motivated by a risk of ambiguity, since no other king is mentioned in the immediate textual environment.

(9)  
Dum hec ageretur, Remistanius, filius Eudone quondam, fidem suam quod praedicto regi Pippino promiserat feellit, et ad Waiofarum iterum ueniens, dictioni sue faciens. Quod Waiofarius cum magno gaudio eum recepit, et adiuorem sibi contra Francos uel praedicto rege eum instituit. Superscriptus Remistanius contra praedicto rege et Francos seu custodias quas ipse rex in ipsas ciuitates dimiserat, nimium infestus accessit, et Bitoriuo seu et Limestone quod ipse rex adquisierat, praedando nimium uastauit, ita ut nullus colonus terre ad laborandum tam agris quam uineis colere non audebant. Praedictus rex Pippinus in Betoricas per hiemem totum cum regina sua Bertradane in palatium resedit. (Cont. § 50)

Meanwhile Remistanius, son of the late Eudo, broke the oath of fealty that he had sworn to *aforementioned King* Pippin. He went back to Waiofar and became his man. Waiofar was delighted to receive him and to make use of his help against the Franks and the *aforementioned King*. The aforementioned Remistianus attacked the *aforementioned King* and the Franks and the garrisons which *this same king* had left in the cities, and he destroyed the districts of Berry and also the Limousin, that *this same king* had conquered and he did this so effectively that not a peasant dared work in the fields and vineyards. *The aforementioned King Pippin* spent the whole winter with Queen Bertrada in his residence at Bourges.

It is in the context of this global increase of the marking of anaphoric relationships that we should consider the extended use of *ipse* and *ille* that will eventually lead to their status of definite article in the emerging Romance languages.

In fact, *ipse* and *ille* are already attested as markers of anaphora in Classical Latin. When the antecedent and the anaphoric expression do not have the same head noun, *ipse* and *ille* alternate in this role with the demonstra-
tives *hic* and *iste* and with the identity marker *idem*. As pointed out by Pinkster (1996:148), some form of determination is even obligatory with “empty” nouns like *tempus*, *locus*, *regio*. In Late Latin, *ipse* and *ille* become common in more types of contexts and they even extend their use to contexts where the head noun of the anaphoric expression is identical to the head noun of the antecedent, at the expense of zero determination. However, *ipse* and *ille* do not alternate freely: their distribution as markers of definiteness and anaphor is a function of their specific meaning.

2.3.1. *Competition between ille* and *ipse*

2.3.1.1. As has been pointed out by Bertocchi (2000), *ipse* in Classical Latin can be interpreted in two divergent directions: presupposing a set of other possible referents, *ipse* can identify a referent by including those other referents (10a: “other things than virtue can be despised, but from all of them virtue is the less expected”) or by excluding them (10b: “from him and from no one else, even this is unexpected”).

(10)  

a. *A multis uirtus ipsa contemnitur*

   *Even virtue* is despised by a lot of people.

b. *Milonem occisum et ex ipso Clodio audirent*

   They heard the death of Milo from the mouth of *Claudius himself*.

It is the exclusive interpretation of *ipse* that will be exploited in Late Latin to express anaphoric relationships.

*Ipse*, in its exclusive interpretation, is a strong marker of identification. It reinforces the continuity of the referential chain when this chain is endangered because the inserted mention of another referent creates a risk of ambiguity (Selig 1992:159) or because it is for some other reason unexpected that the referent is identical to the previously mentioned one. It is in this way that the use of *ipse* can be motivated in the following example:

(11) *Inde egressus Crisceco uilla ueniens in Pontio, Leudesio sub dolo fيدem promitti se simulans sefellit, facto placito ut coniuracione facta cum pacaе dixerent. Sed Ebroinus fallaciter agens ut solebat, conpatri suo insidias praeparans ipsum Leudesium interficit; […]* (Continuationes, § 2)
Then Ebroin left Baizieux and reached the domain of Crécy in Ponthieu. He there deceived Leudesius by making a false promise that they should arrange a meeting and, after having exchanged loyalty oaths to each other, should part friends. But, as usual, Ebroin acted treacherously. He laid an ambush for his godfather and slew this same Leudesius.

It is indeed contrary to the expectations that the main character, after having given his word to his godfather Leudesius, kills this very same person.

2.3.1.2. Being a deictic expression, *ille* conveys the instruction to identify the referent by means of indications present in the context of its occurrence, whether it is the situational or the textual context. Hence, *ille* implies that the referent is not presented as it was given before or as already salient in the context. The demonstrative *ille* can nevertheless refer to a previously mentioned referent, but in this use, it brings about a new identification of the referent and creates thus a discontinuity in the referential chain, without however breaking it down (see the analysis of the French demonstratives determiners by Kleiber (1986a/b; 1987a; 1990b) and De Mulder (1997)). As was illustrated by example (5), it is for instance typically used when the referent is implied in an event that constitutes the turning point of the narration.

The deictic force of *ille* that induces a new identification of the referent by means of the context of utterance also explains why *ille*, and not *ipse*, is used when the anaphoric expression contains new information about the previously mentioned referent by means of a recategorization.

(12)  *Quando Deo conplacuit, Aubedo ligatarius directus a Chlodoueo regi causam legationes usque ad Chrotharium regem Langobardorum, Papia coinomento Ticino ciuitatem Aetaliae peruenisset, cernens regina, quam sepius in legationem ueniens uiderat et ab ipsa benigne semper susceptus fuere fuisse retrusam, quasi iniunctum habens exinde inter citera Chrothario regi suggessit quod illam parentem Francorum quam reginam habuerat, per quem etiam regnum adsumserat, non dibuisset umiliare; [...] (Fredegarius § 71)*

In God’s good time, Aubedo, sent by King Clovis on an embassy to the Lombard King Rothari, reached the Italian city of Pavia, or Ticinum. Realizing that the Queen, whom he had often seen dur-
ing his missions and who had always received him well, was in-
carcerated, he put it in the course of the conversation to King
Rothari, as if on instructions, that it would have been better not to
humiliate that relative of the Franks who had been given to him
as Queen and who had been instrumental in obtaining the throne
for him.

The anaphoric use is not the only textual role of ille. As has been
pointed out by Selig (1992:165), ille occurs also in first mention when the
existence of the referent is established by a relative clause. This use is co-
herent with the instruction conveyed by ille, inducing a new identification
of the referent. The relative clause allows in this case the hearer to identify
the referent. Ipse is practically not used in this configuration (cf. § 2.1.2.2.
and footnote 8, but also Herman 2000:101; Fruyt 2003:104). The first men-
tion use of ille is illustrated by the example (3), repeated here for conven-
ience.

(3) […] homines illos quos Waiofarius ad defendendam ipsam ciui-
tatem dimiserat clementiam sue pietatis absoluit dimissisque
reversi sunt ad propria. (Continuationes, § 43)
Of his goodness he showed mercy to the (litt. that) men that Waio-
far left there to defend the city, and dismissed them to go off home.

2.3.1.3. The extension of the conditions of use of ipse and ille can be
described as a result of pragmatic strengthening: the speaker/writer puts the
original meaning of ipse and ille into service to guide the hearer/reader in
his interpretation process of identifying the referent and to underline its
importance from the point of view of the textual structure. Although ipse
and ille are largely conventionalized in Late Latin as markers of this
speaker–hearer interaction, there is no divorce from their original mean-
ing. In this context, it has to be mentioned that the Fredegarius corpus does
not contain reinforced forms of the type ecce ille or met-ipsimus.13

There are however some occurrences where ipse and ille no longer have
their full original meaning, inaugurating the desemantization process that
will lead to the creation of the definite article.

(13) Burgaris superatis, noue milia uerorum cum uxoris et liberis de
Pannonias expulsi, ad Dagoberto expetint, petentes ut eos in terra
Francorum manendum receperit. Dagobertus iobit eos ad iae-
mandum Badowarius recipere, dummodo pertractabat cum Francis quid exinde fierit. Cumque dispersi per domus Baioariorum ad hyemandum fissent, consilium Francorum Dagobertus Baioariis iobet ut Bulgarus illus (var. Bulgaros illos) cum uxoris et liberis unusquisque in domum suam una nocte Baiuarae inter-ficerint. Quod protinus a Baiouaries est impletum; nec quisquam ex illis remansit Bulgaris nisi tantum Alciocus cum septinientis uiris et uxoris cum liberis, qui in marca Vinedorum saluatus est. (Fredegarius, § 72)

The Bulgars were beaten and nine thousand of them were chased out of Pannonia with their wives and their children. They sought asylum of Dagobert, begging him to take them in and give them a home in Frankish territory. Dagobert gave instructions that they might winter among the Bavarians, and in the meantime he would deliberate with the Franks about their future. When they were dispersed among the Bavarian homesteads for the winter, Dagobert took the advice of his Franks and ordered the Bavarians to kill those Bulgars with their wives and families during the night in their homes. The order was at once carried out. Nobody of those Bulgars survived, at the exception of Alzeco, with 700 men and their wives and their families: they found safety in the Wendish March.

Since the Bulgarians constitute the principal theme of this anecdote, their identification should not be an issue at the end of the excerpt and the instruction conveyed by ille, requiring a new identification of the referent by means of the context of utterance, is too strong. The process of semantic bleaching is even in a more advanced stage for ipse, especially in the more recent part of the corpus. As witnessed in example (14), the identity marker ipse is used as an anaphoric device although the referential chain is not endangered: the referent is resumed not only under the same denomination, i.e. its proper name, but moreover as involved in the same event.

(14) ipse praedictus rex cum reliquis Francis et obtimatibus suis persequendum Waiofarium ire perrexit. Cumque praedictus rex ipsum Waiofarium persequente non reperreret, iam tempus hiem erat, cum omni exercito ad Betoricas, […], reuersus est. (Continuationes, § 49)

The King himself set out the remaining Franks and nobles in pur-
suit of Waiofar. But as he persued the same Waiofar without finding him and it was already winter, he returned with his army to Bourges, where he had left Queen Betrada.

2.3.2. Selection of ille over ipse

Ipse and ille initiate concurrently the grammaticalization process leading to the definite article: *ipse* conveys the instruction to resume the referent identically and strengthens the continuity of the referential chain, whereas *ille* signals that the referent should be identified anew by means of indications in the context of its occurrence and introduces in this way a discontinuity in the referential chain even in its anaphoric uses.

The existence of two competing forms in the beginning stages of the grammaticalization process is not at all exceptional, as is shown by the archetypal example of negation in French (Old French: *ne ... pas / point / mie / goutte / ...* > Modern French *ne ... pas* (Meillet 1912)). It is nevertheless paradoxical that *ipse*, though more frequent in Late Latin, gave way to *ille* and was eventually not retained as a definite article in most of the Romance languages (cf. Aebischer 1948, Selig 1992:133, and Bauer 2007 for an overview).

Several hypotheses, phonetic (Banniard 1998), socio-linguistic (Aebischer 1948) and structural (Selig 1992), have been put forward to resolve this paradox. Our hypothesis takes into account the specific referential instruction conveyed by each of these two competing expressions.

In comparison with the demonstrative *ille*, mobilizing the context of utterance for a new identification of the referent, the identity marker *ipse* conveys a simpler and more precise referential instruction as an anaphoric device. As long as *ipse* and *ille* coexisted, this precision of its referential instruction turned out to the advantage of *ipse* and explains why it was more frequent. The coexistence of several linguistic expressions as incipient definite articles was not to persist, however, especially when a paradigm of articles started to develop, also containing the indefinite singular article, derived from the unity numeral *unus*. Indeed, when such a tightening of paradigms or “paradigmatization” (Lehmann 1992) takes place, the resulting highly grammaticalized paradigms tend to be structured in terms of binary or two-term oppositions (Lehmann 1992:136). The development of a paradigm of articles thus led to the selection of one expression as the definite article, in binary opposition with the indefinite article *unus*. In this
process, most Romance languages finally preferred *ille* over *ipse*, because *ille* had several assets to be selected in the grammaticalization process leading to the definite article.\(^{15}\)

Firstly, *ille* is more flexible as an anaphoric device than *ipse*. Contrary to *ipse*, *ille* does not require that the referent be strictly identical to an aforementioned one; it allows, in virtue of its deictic force, that new elements in the context of its occurrence be taken into account for the identification of the referent. This is the reason why *ille* is not restricted to anaphora, but is also used in the case of first mentions (cf. example (3)).

Moreover, *ille* takes advantage of its status of demonstrative associated with the third person. Whereas *hic* and *iste*, linked respectively to the first and the second person, i.e. the speech participants, denote a referent that is wholly identified in the context of their occurrence, *ille* signals that the identification of the referent is not saturated by the immediate context of its occurrence, and thus suggests that information not available in this context should be activated to retrieve the referent. In this way, *ille* can be understood as an invitation addressed to the hearer to mobilize previous knowledge in order to retrieve the referent.\(^{16}\) This intersubjective dimension of the distal demonstrative is foregrounded in its so-called “recognitional” use, illustrated by the following French example:

\[(15)\] **Cette personne**, tu sais, Mlle E... dont je t’ai parlé, avec qui je devais prendre le thé dans la quinzaine chez Mme Chesneau qui la connaît, eh bien, j’ai appris hier qu’elle a un amant, un grand banquier de Paris, qui ne veut pas l’épouser. (Villiers de L’Isle-Adam, *Correspondance générale*)

**That person**, you know, Ms E… I talked you about her, with whom I had to have tea within two weeks in the house of Mrs Chesneau who knows her, well, I heard yesterday that she has a lover, an important banker of Paris, who doesn’t want to marry her.

As is observed by Himmelmann (1996, 1997:61, 2001) and Diessel (1999), the recognitional use of demonstratives has a distinctive feature with respect to their other uses: the referent is not mentioned in the preceding context or present in the current discourse situation. The demonstrative points to “specific, presumably shared” knowledge (Himmelmann 2001:833) – often based upon a common experience – that has to be activated by the hearer in order to identify the referent. Moreover, the speaker
is not sure that the hearer will be able to activate this knowledge and to identify the referent – witness the insertion of *tu sais* “you know” in example (15) – and can provide supplementary information that should allow the addressee to find the referent, for instance by adding a relative clause. Examples of the recognitional use of *ille* in Late Latin are cited by Selig (1992:166-167). The following example is written on the back of a royal charter of the Saint Denis’ Abbey:


Here are the documents of the toll of *the* (litt. *that*) *market* (of our market, of the well-known market organized on the day of the Patron Saint)

In its recognitional use, the demonstrative *ille* is similar to the definite article to the extent that it does not require the referent to be available in the context or the immediate situation. However, there also is an important difference: in the recognitional use of the demonstrative, the referent is identified on the basis of specific knowledge that is shared by speaker and hearer; in characteristic uses of the definite article, those where it cannot be replaced by a demonstrative determiner, the identification of the referent is based on stereotypical knowledge, or knowledge that is presumably shared by all members of the relevant speech community. As pointed out by Himmelmann (1997; 2001), the recognitional use of the distal demonstrative allows us to understand the transition between the distal demonstrative and the definite article: the distal demonstrative becomes a definite article when the presumably shared knowledge necessary to identify the referent is no longer presented as specific to the speaker and the hearer, but as consisting of information presumably shared by all members of the speech community.

In sum, the evolution of the Latin demonstrative system, which eventually led to the creation of the definite article, has been shown to be far more intricate than what could be concluded from the schema presented by Harris (1978) that was used by Dasher and Traugott (2002) as an illustration of the process of subjectification. Nevertheless, the hypothesis of a cline of intersubjectification is still relevant to describe the evolution from demonstrative to definite article.
2.3.3. The cline of intersubjectication

Traugott (this volume) presents the cline of intersubjectification as follows:

| Non-/less-subjective | Subjective | Intersubjective |

As to the identification of the different stages of this evolution, it is an oversimplification to identify the Classical Latin use of demonstratives with the first stage of this evolution: the inclusion or exclusion with respect to the deictic center is very easily subject to an evaluative interpretation and the first stage should perhaps be conceived rather as a theoretical starting point. In Classical Latin, some demonstratives can already be used as social deictics and encode the subjective attitude of the speaker towards the referent (cf. §1).

A new evolution is observed in Late Latin, when *ipse* and *ille* become involved in a grammaticalization process that will lead to the creation of a definite article. The force that triggers this grammaticalization process is strategic interaction between the writer and the reader: the writer aims to offer to the reader the clues needed for a clear and coherent interpretation of the text and for the identification of the referents participating in the narrative. In Late Latin, definite reference to the participants in the narrative is normally marked by zero determination. However, when the writer wants to attract the reader’s attention to the referential act, either because the precise and exact identification of the intended referent is not guaranteed, or because the referent has topical or focal status, he develops explicit markers for this purpose. He makes an extended use of *ille* and, even more markedly, of *ipse* by exploiting their original meanings: by using the identity marker *ipse*, he gives the instruction to resume a previously mentioned referent identically; by means of the demonstrative *ille*, he asks the reader to make a new identification of the referent. Since the use of these expressions is marked with respect to zero determination, it can have as such the effect of attracting attention to their referent and of conveying it topical or focal value. In the case of *ille*, the role of “attention-getter” relies moreover on its deictic meaning, in virtue of which *ille* requires the hearer to take into account the specific occurrence of the demonstrative and its context.

Traugott (2003:126) defines subjectification as “the mechanism whereby meanings come over time to encode or externalise the speaker / writer’s perspectives and attitudes as constrained by the communicative
world of the speech event, rather than by the so-called “real-world” characteristics of the event or situation referred to”. According to this definition, subjectification is involved in the first stage of the grammaticalization process that leads to the creation of the definite article: by the extended use of *ipse* and *ille*, the writer verbalizes his perspective on the referent, its discourse status and its role in creating text connectedness. Even if *ipse* as an identity marker and *ille* as the distance demonstrative already had this meaning component in Classical Latin, their extended use in Late Latin leads to the foregrounding of this meaning component. In line with this subjectification process, it is not surprising that in these uses, the demonstrative *ille* implies an imbalance between writer and reader, since it signals that the writer is striving to ensure unambiguous reference, by guiding the reader to the referent. As pointed out above, since the writer is not sure that the reader will be able to identify the referent, he frequently adds, in recognitional uses of the demonstrative, relative clauses that should give the reader more identifying information. In this way, it can truly be said that subjectification involves a centering of meaning on the speaker or writer (Traugott 2003:129).

However, if the use of a demonstrative *ille* signals that the writer attracts the reader’s attention to the reference act – amongst other things because he is not sure that the hearer will be able to identify the referent –, this implies that the writer has a representation of the reader’s text model and takes into account this representation of the reader’s text model when choosing the referential expression. In as far as the demonstrative is an expression of the writer’s awareness of the beliefs of the reader, it is also an expression of intersubjectivity. The question might be raised whether the evolution from demonstrative to definite article can be described in terms of a strengthening of this intersubjective dimension.

Intersubjectification can be defined as the mechanism whereby the meaning of an expression, instead of being centered on the speaker as a result of a preceding subjectification, becomes “more deeply centered on the addressee” (Traugott 2003:129). Traugott (2003:130) cites the three-step evolution of *let’s* as a typical example of the cline of intersubjectification:

1) *Let us go, will you*: “allow us (imp.)” >
2) *Let’s go, shall we*: “I propose (hortative)” >
3) *Let’s take our pills now, Roger*: “mitigator / marker of ‘care-giver register’”.


The transition from the first to the second step involves subjunctification; the transition from the second to the third step represents an instance of intersubjectification, since the speaker “in the newer use positions him- or herself as empathetic to the addressee’s possible objection to the projected activity (in this case, pill-taking) in the here and now of the ongoing discourse” (Traugott 2003:130). The meaning change between the two last stages can actually also be described as a shift in perspective: whereas in the second stage, the event is presented from the speaker’s perspective and reflects his point of view, in the third stage, the speaker adopts the perspective of the hearer.

The transition from demonstrative to definite article shows up a similar shift in perspective. For the demonstrative, the role of the writer / speaker is dominant: in the initial stage of the grammaticalization process, demonstratives are used when the writer / speaker feels it is necessary to attract the reader’s / hearer’s attention, because he thinks the hearer / reader will have to make some effort to identify the referent or because the referent has a prominent role in the text or discourse. The definite article signals in its typical uses (cf. endnote 14) that the hearer can identify the referent because the knowledge involved in the identification is stereotypical or presumably shared by all members of the speech community. In other terms, whereas the demonstrative signals that the referent is presented dominantly from the speaker’s perspective, the emergence of the definite article implies a shift to a less speaker-centered and more intersubjectively shared perspective.

The analysis presented above might seem surprising in the light of recent papers by Traugott, where she focuses more exclusively on attitudinal aspects rather than on text coherence as she distinguishes between subjunctification and intersubjectification: subjunctification is the development of meanings that express the speaker attitude or viewpoint, while intersubjectification is the development of meanings that express the speaker attention to addressee self-image (Traugott 2003:126, 130; this volume). In this perspective, honorifics and other items that convey social attitudes of politeness are prototypical examples of intersubjectification. However, even in these articles, some examples allow for a broader conception of subjunctification and intersubjectification, which was also present in earlier work, such as Traugott (1995:47), for example, where subjunctification is defined as “the tendency to recruit lexical material for purposes of creating text and indicating attitudes in discourse situations”. In this view, connectives, ex-
pressing coherence relations between propositions, or anaphoric or cataplectic pronouns can be considered as instances of subjectification (Traugott 1982:250), because it is the speaker who decides how to link the propositions in the text to each other and how to link propositions to the context. More generally put, subjectification also includes “the development of meanings with which the speaker ‘creates text’” (Breban 2006b:260). As far as referential expressions are concerned, then, we agree with Breban (2006a:104) that “attitudinal and textual subjectivity are […] two different manifestations of the same phenomenon neither of which can be excluded at the expense of the other”; subjectification as well as intersubjectification have according to us both an attitudinal and a textual dimension.

3. Conclusion

This paper offers an account of the emergence of the definite article in the transition from Late Latin to French. It tries to elucidate the enigma of the selection of *ille* as the source of the definite article in French and in most Romance languages, in spite of its low frequency in comparison with its competitor *ipse* in Late Latin. It is argued that the grammaticalization from the distal demonstrative *ille* into the definite article cannot be reduced to a linear process of desemantization, consisting only in the loss of its deictic meaning, but that it also involves a stage of pragmatic enrichment: by using an identity marker such as *ipse* or a demonstrative such as *ille*, the speaker not only tells the hearer how the referent of the noun phrase has to be identified; in a language stage where zero determination is the rule, the use of these expressions is also a strategy by which the speaker conveys to the hearer an instruction to pay attention to the identification of the referent, thus highlights it and verbalizes its discourse status.

In her production based approach of change, Traugott (this volume) argues that subjectification should precede intersubjectification and that it is likely to occur in the early stages of the grammaticalization process. The current study shows how grammaticalization and subjectification are correlated in the first stage of the process leading to the definite article: this process is initiated when speakers increasingly use existing expressions like *ipse* and *ille* in order to facilitate the identification of the referent and indicate its discourse status, strengthening thereby their respective subjective meaning component. In this initial stage the role of the speaker is dominant.
Our analysis also brings to light a shift in balance between speaker and hearer as grammaticalization proceeds: the transition to a full-fledged definite article goes along with a more symmetric relationship between speaker and hearer and involves in this way a strengthening of the intersubjective dimension.

Intersubjectification as it is conceived here is not limited to the encoding of attitudinal aspects; it also concerns more globally items that materialize the strategic interaction between speaker and hearer and reflect the active role of the speaker to orient and to guide the hearer in his interpretational tasks. These reflections raise the question how intersubjectification should be defined, a need for precision that is also hinted at at the end of Traugott’s article.

References


26

Breban, T.

Calboli, G.

1990 Les pronoms démonstratifs latins et la formation de l’article roman. RPh. 64: 71-88
Carlier, A.

Carlier, A. and W. De Mulder

Charolles, M.

Christol, A.

Diessel, H.


Epstein, R.


1994a Discourse and Definiteness: Synchronic and Diachronic Perspectives. Ph.D., San Diego, Univ. of California.
Epstein, R.


Himmelman, N.
1996 Demonstratives in narrative discourse: a taxonomy of universal uses.

Himmelmann, N.

Hopper, P.J. and E.C. Traugott

Kleiber, G.

Kleiber, G.
1986b Pour une explication du paradoxe de la reprise immédiate UN Ni→LE Ni / UN Ni→CE Ni. Langue française 72: 54-79.

Kleiber, G.

Kleiber, G.
1987b L’opposition cist / cil en ancien français ou comment analyser les démonstratifs?. Revue de linguistique romane 51, 5-35.

Kleiber, G.

Kleiber, G.

Kleiber, G.

Lehmann, C.

Lehmann C.

Löfstedt, E.
1911 Philologische Kommentar zur Peregrinatio Aetheriae. Uppsala.

Lyons, C.
1999 Definiteness. Cambridge: CUP.
Meillet, A.  

Menge H.  

Nocentini, A.  
1987  L’uso dei demonstrativi nella Peregrinatio Egeriae e la genesi dell’articolo romanzo. Atti del convegno internazionale sulla Peregratio Egeriae, 137-158.

Orlandini, A.  

Pei, M.  
1936  Old French Demonstratives. Language 12, 47-51

Pinkster, H.  

Renzi, L.  

Renzi, L.  

Selig, M.  

Serbat, G.  

Trager, G.  
1932  The Use of the Latin Demonstratives (Especially Ille and Ipse) up to 600 A.D., as the Source of the Romance Article. New York: Publications of the Institute of French Studies.

Traugott, E.C.  

Traugott, E.C.  

Traugott, E.C.

Traugott, E.C. this volume From ideational to interpersonal: a reassessment.


Notes

1 In spite of the register-based terminology of ‘Classical’ versus ‘Vulgar Latin’, Harris’s perspective is diachronic. In accordance with this diachronic perspective, instead of ‘Vulgar Latin’, we will use the term ‘Late Latin’ in this article.

2 The disappearance of *hic* and *is* is generally explained in terms of phonetic weakness (Harris 1978, Vincent 1997). As to *idem*, its destiny is linked to that of *is*, as it contains *is* as one of its constituents.

3 In Classical Latin, *-met* is a reinforcing suffix combining with personal pronouns, often followed by *ipse* (e.g. *egomet ipse, memet ipsum* cf. Menge 2000: § 57). The formation of the Vulgar Latin form *metipsimum*, giving rise to the forms *medismo* in Italian, *mismo* in Spanish, *même* in French and so on, involves a reanalysis (cf. Vaänänen 1967: § 279; Harris 1978: 69).

According to Serbat (1975: 124), we find for instance in the Chronicle of Fredegarius the first occurrence of the Romance future composed of the infinitive and the auxiliary habere agglutinated to the infinitive:

Et ille respondebat: Non *dabo*. Justinianus dicebat: *Daras* (daras = dare + -as < habes)

He answered: I *will not give*. Justinianus said: you *will give*

These tendencies are not idiosyncratic. The numeric importance of hic and/or the relative paucity of iste have also been observed in other corpora by Fisher (1908), Bonnet (1968: 301, cited by Harris 1978), Selig (1992), Abel (1971), Flobert (1974: 152) and by the numerous studies on the Peregrinatio (inter alii Löfstedt 1911, Nocentini 1987, Christol 1994, Vincent 1997, Fruyt 2003). They are actually inherited from Classical Latin (Fisher 1908). As is noted by Abel (1971) and Bauer (2007:117), the persistence of the high frequency of hic and of the sporadic use of iste is problematic in the light of their ultimate fate in the Romance languages. According to Abel (1971), the high frequency of hic in the Vetus Latina corpus could be ascribed to its frequent use in fixed expressions, which could be interpreted as a phenomenon of fossilization (Lehmann 1991). This interpretation of the statistical data biased by the results in Romance is however contested by Flobert (1974: 152). The long-lasting high frequency of hic in Late Latin is a subject of further study.

According to Greenberg (1978), the definite article can lose its specifying function and become merely a sign of nominality. This hypothesis has been exploited by Harris (1978: 76-78) to analyze the Modern French definite article, on the basis of two observations:

- In Modern French, nouns have lost their suffixal endings of number and gender. The large expansion of the articles during the history of the French language can be explained in the light of this evolution, since they inherit the grammatical function of marking number and gender.
- According to Harris, under the pressure of the generalization of an obligatory prenominal determiner in French, the definite article even extended its use to contexts where it cannot be motivated by its fundamental semantic value, which consists in specifying a particular entity. Harris (1978: 76) mentions in this context the generic use of the definite article *le* (*J’aime les livres; la haine provoque la guerre*).

It has to be noted that these two evolutions are at least partially independent of each other: in other Romance languages, like Spanish, Portuguese or Italian, where zero determination is still widespread, genericity is nevertheless conceptualized in terms of definiteness. Similarly, in Modern French, the use of the definite article in generic interpretation is semantically motivated and is not merely due to the regression of zero determination (cf. Kleiber 1990a). For a
general evaluation of Harris’s analysis of the modern French definite article as a pure marker of nominality, see De Mulder and Carlier (2006).

8 On the basis of the analysis of her corpus, Selig (1992: 165) affirms: « Der Bereich der definite Erstnennungen ist die eigentliche Funktionsdomäne von ille in den hier analysierten Texten » (‘The domain of definite first mentions is the proper functional domain of ille in this corpus’). The use of ipse in combination with a relative clause establishing the identity of the referent is very exceptional and limited to certain types of documents (Selig 1992: 158).

9 The relevance of this first mention use for the emergence of the definite article also calls into question the idea that the definite article originates from anaphoric uses of demonstratives, as proposed e.g. by Diessel (2006: 477).

10 The opposition between both determiners can be illustrated by the following classical examples (cf. Charolles 1990: 128):

Nous arrivâmes dans un village. L’église était sur une colline.
We arrived in a town. The church was on a hill

Nous arrivâmes dans un village. Cette église, tout de même, quelle horreur!
We arrived in a town. That church, really, how awful!

As Charolles points out, in the first example, the church is identified in virtue of the stereotypical relation between churches and villages (‘each village has a church’) and is presented as ‘the church-of-the-village’. In this example, the definite article cannot readily be replaced by a demonstrative determiner. The demonstrative determiner is, however, appropriately used in the exclamative utterance that is part of the second example. It presents the referent as a recollection of a past experience of the speaker, in which he was directly confronted with the church in question. Thus, the use of the demonstrative reflects a more subjective construal of the referent.

11 In the same vein, Diessel (2006:469) holds that demonstratives serve two basic functions: “1. First, they indicate the location of a referent relative to the deictic centre. 2. Second, they serve to coordinate the interlocutors’ joint attentional focus.” In his view, the fact that demonstratives coordinate the interlocutors’ joint focus of attention, which is a fundamental element of human communication, makes demonstratives a basic, but also a particular word class.

12 This highlighting effect of explicit determination, in contrast with the default case of zero determination, is also stressed by Selig (1992:117-119). She holds, though, that the focalizing effect of determination disappears in the transition from (Late) Latin to Romance, in the context of the grammaticalization of the (in)definiteness opposition. However, as pointed out above, the incipient indefinite singular article un(s) in Old French also bears on nominal expressions referring to protagonists or other important referents, whereas zero determina-
tion, being the default case for indefinite reference, is used when the nominal expression refers to a background entity whose precise identity is less important than the category to which it belongs.

13 Pei (1936: 49) observes that the reinforced forms *ecce ille* and *ecce iste*, which appear in the (preclassical) theatre texts of Plautus, are globally not in evidence in the Late Latin texts. See however the often quoted occurrences in the *Peregrinatio*. As is noted by Renzi (2000: 185) on the basis of a study on Late Latin charts, in the same way, *ipse* does not alternate with the reinforced form, but cumulates its classical meaning and its new anaphoric role. The reinforced form is however attested sporadically in other types of texts.

14 *Cf.* Carlier & De Mulder (2007: § 2.5) for a critical overview of these hypotheses.

15 As has been shown by Vincent (1997), this evolution goes along with the selection of *ille* as the third person pronoun. In Late Latin, the distinction between pronoun and determiner does not correspond to a difference of morphosyntactic category, but to a difference of syntactic construction. For certain demonstratives, some specialization is observed (e.g. *is* cf. § 1), prefiguring the emergence of separate paradigms.

16 An analysis along these lines has been proposed by Vuillaume (1980) for the German distal demonstrative and by Kleiber (1987b) for the Old French distal demonstrative.

17 Himmelmann (1997; 2001) considers the defining uses of the definite article to be its larger situation uses, where the article refers to entities that are considered to be unique in a given speech community (e.g. *the sun, the Queen, the pub*), and its associative anaphoric uses, where the uniqueness expressed by the definite article is justified by stereotypical knowledge.

18 As pointed out by Himmelmann (1997: 94-101), this hypothesis is less problematic than the more traditional idea that the definite article develops from the demonstrative’s anaphoric uses (which has recently been resumed by Diessel 2006: 477).

19 Traugott and Dasher (2002) use the terms *speaker / writer* and *addressee / reader*. Since our corpus is a written text, we will use *writer* and *reader*. In so doing, we do not take into account subtle distinctions between writer, author, narrator etc., which do not seem essential to the point we try to make here.

20 Epstein also considers subjectification to be involved in the emergence stage of the definite article. In his conception, however, subjectification is restricted to expressive uses, as opposed to referential uses (*cf.* § 2.2).

21 The term *text model* is used here as a shorthand for the representation of the meaning of the text that is elaborated by the reader when reading the text. The crucial importance of man’s capacity to understand the other’s mental or inten-
tional states, as well as its close relationship to pointing and to the use of demonstratives, has been pointed out by Diessel (2006: 468).

Indeed, “intersubjectivity in my view refers to the way in which natural languages, in their structure and their normal manner of operation, provide for the locutionary agent’s expression of his or her awareness of the addressee’s attitudes and beliefs, most especially their face or self-image” (Traugott 2003; this volume).