

Some strategies for coding sentential subjects in English: from exaptation to grammaticalization¹

Teresa Fanego

University of Santiago de Compostela

This study examines two types of *-ing* subjects in English: the Late Modern English pattern *the deceiving him was easy* and constructions such as *by trying to make her mother happy proved unlucky for Paul*, which are becoming frequent among American undergraduates. It is argued that the presence of *the* and *by* in both structures is prompted by the desire to provide explicit grammatical marking for clauses that, because of their role as pre-verbal subjects, are cognitively very complex. The paper discusses the ongoing grammaticalization of *by Xing* sequences and outlines the historical developments leading to the emergence of the pattern with an introductory *the*, for which it proposes an analysis in terms of Lass's (1990, 1997) concept of *exaptation*.

1. Introduction

Fauconnier (1997:176) and Hogan/Hogan (1998) draw attention to the increasing use in American undergraduate writing of the sentence type illustrated in (1)-(2), where a gerundive clause introduced by the preposition *by* serves as pre-verbal subject:

- (1) *By trying to make his mother happy* proved unlucky for the character Paul in this story.
- (2) *By interviewing buyers of instant coffee and buyers of regular coffee* is akin to asking people whether they prefer the taste of Coke or Pepsi.

In this paper, I will argue that this “impertinent” *by* —to use Hogan/Hogan's terminology— has evolved from its ordinary prepositional use into a clause linker or complementizer whose function is to provide explicit marking for the following subject clause. I will further argue that its role is analogous to that of the initial *the* in the complement type exemplified in (3):

- (3) COPC 1709 Berkeley *New Theory of Vision* 073/030-P23: *The not observing what has been delivered in the two last sections* seems to have occasioned no small part of the difficulty that occurs in the business of erect appearances.

This kind of subject clause was frequently met with in earlier stages of English, but became obsolete from about the middle of the nineteenth century (Fanego forthcoming

b). Both in this type of clause and in the *by*-clauses in (1)-(2), the presence of an introductory element —*the* and *by* respectively— appears to be a means of avoiding the processing difficulties of bare *-ing* clauses when these occur in subject position (e.g. “*trying to make his mother happy* proved unlucky for Paul”). A corollary of this need for increased marking is that, as noted above, *the* and *by* have developed uses which approach those of so-called complementizers; in this way two elements, a determiner and a preposition respectively, which are already part of the grammar have come to acquire a new grammatical function.

The remainder of this paper will be devoted to examining this development from the viewpoint of grammaticalization theory. In Section 2 I briefly review the concepts “complementation” and “complementizer”. Sections 3 and 4 summarize the earlier history of English gerundive subject clauses and propose to account for the use of *the* exemplified above in terms of Lass’s (1990, 1997) notion of *exaptation*. Section 5 discusses the grammatical status of impertinent *by* in light of the data adduced by Fauconnier (1997) and Hogan/Hogan (1998), and Section 6 offers some concluding remarks.

2. Complement types and complementizers

Complement clauses are commonly defined as subordinate clauses functioning as an argument of a predicate (Noonan 1985:42), as in examples (1)-(3) above. They may serve as subject, object² or predicative of the superordinate clause and may include a finite or nonfinite verb form. As noted by Noonan (*ibid.*), they are often “associated with a word, particle, clitic, or affix whose function it is to identify the entity as a complement. Such forms are known as *complementizers*”.

Noonan’s oft-quoted definition is indicative of the heterogeneity of the items subsumed under the category “complementizer”, whose usefulness in grammatical description has sometimes been questioned.³ In the case of English, listings of complementizers usually include items so diverse in origin and syntactic behaviour as *that*, *if*, *whether* and *for* (“I am waiting *for* it to stop raining”), but infinitival *to* has also been considered a complementizer at times (see, for instance, Noonan 1985, Van Valin/LaPolla 1997:469ff), and in the early days of generative grammar one could often read of the “complementizers” *-ing* and *Poss*, as in “*John’s coming tomorrow*” (cf.

Rosenbaum 1967, Huddleston 1971:164ff). In the main, I will here adopt the position of Huddleston/Pullum (2002:22, 954ff, 1011ff), who recognize a category of *subordinators* —*that*, *if*, *whether*, and infinitival *to* and *for*— functioning as mere “syntactic markers of a particular syntactic construction” (p.1183) and not heading their own clauses. They are hereby distinguished from a much larger class of prepositions which are the heads of their own constructions and can, in some cases, govern either clauses or noun phrases (p.669):

- (4) a. I left *before the lecture ended*
 b. I left *before signing the contract*
 c. I left *before the end of the lecture*

For Huddleston/Pullum, one of the criteria for the distinction between the subordinator and prepositional classes is the fact that the items in the prepositional class “have evident semantic content, and this content is clearly the major factor in determining the function and distribution of the construction they introduce” (p.1012). Thus the italicized sequences in (4a-c) above are adjuncts of time, and “are obviously construed in this way by virtue of the meaning of *before*” (ibid.). By contrast, items like *that* (“it is important *that detailed records be kept*”) or infinitival *for* (“it is important *for detailed records to be kept*”) have “no identifiable meaning of [their] own” (p.1183).

Some complement types, for instance English gerundive clauses (“*writing letters is easy*”), do not have a complementizer or subordinator associated with them. Likewise, finite declarative complements may be introduced in English by *that* or zero:

- (5) *That they were lying* is now quite obvious.
 (6) *That he really intended to cheat us* I still can’t believe.
 (7) I think [*that*] *it is a good idea*.

With object clauses, the choice between *that* and zero is determined by complex grammatical and stylistic factors which have often been discussed in the literature (cf. McDavid 1964, Elsness 1984, Fanego 1990, Finegan/Biber 1995, Rohdenburg 1995, 1996, Huddleston/Pullum 2002:953ff, Thompson 2002) and need not detain us here. When the clause is a subject ((5)) or otherwise precedes the matrix predicate ((6)), *that* is obligatory, as it has then the essential role of signalling the start of a subordinate clause rather than a main clause. For much the same reason, bare infinitivals cannot occur as subjects, except in informal style in the reversed version of the specifying *be*

construction when the internal complement contains the verb *do* (cf. Declerck 1991:491-492, Huddleston/Pullum 2002:1254):⁴

- (8) *Plead mitigating circumstances* is all you can do.
 (9) *To refuse her request* /**Refuse her request* would be unthinkable.

Gerundive clauses are in fact the only English complement type which does not require increased marking in order to be used as pre-verbal subject:

- (10) *Inviting your uncle* was a bad mistake.
 (11) *There being no handle to the suitcase* makes it difficult to carry.

Yet, interestingly enough, all descriptions of Present-day English usage (Quirk et al. 1985:1063-1064, Declerck 1991:499-500, Huddleston/Pullum 2002:1192-1193) note that, at least in written English and in formal styles, if a subject *-ing* clause has an overt subject, this is preferably marked with the genitive case —the historically older form—, despite the fact that in all other environments accusative pronouns and plain case NPs have largely displaced genitives:

- (12) *His* /**him* *being too credulous* often gets him into trouble. [quoted from Declerck 1991:500]
 (13) It involved *the Minister of Transport* / *the Minister of Transport's losing face*. [quoted from Huddleston/Pullum 2002:1192]
 (14) I don't mind *John* / *him* *knowing the truth*.

Among other possible explanations for these facts of English grammar, it could be argued that the genitive form is favoured in (12) because it serves as a more efficient signal of the nominal role (i.e. subject) of its own clause in the larger construction. In (13)-(14), *It involved* and *I don't mind* prepare the ground for the constituent which is to follow, and this, irrespective of case marking, can only be an object of some kind. Later in this paper I will return to these issues in greater detail.

3. English verbal gerunds:⁵ origins and development

The ancestor of the gerundive clauses exemplified in the preceding sections was an Old English⁶ abstract noun of action formed through the addition of the derivational suffix *-ing* to a verb stem, as in *spilling* “destruction” (< *spillan* “destroy”) and *wending* “turning” (< *wendan* “turn”); see Visser (1963-1973:§1001) for details.⁷ In Old English and Early Middle English these nouns behaved like any other noun in all relevant respects, and could therefore take nominal dependents of various kinds. The following

examples illustrate their use with determiners (*the, his*) and with *of*-phrases serving as their notional objects:

- (15) a1359 *Midland Prose Psalter* 10.9 [Tajima 1985:63]:
 þe biginnyng of wysdome is *dredyng of our Lord*
 “the beginning of wisdom is dreading our Lord”
- (16) 1472-1488 *Cely Letters* 94/5 [Tajima 1985:68]:
 at *the makynge of thys lettyr*
 “when writing this letter”
- (17) c1385 Chaucer *Troilus and Criseyde* V 1833 [Tajima 1985:70]:
 And thus began *his loving of Criseyde*

In the course of the Middle English period, *-ing* nominals began to acquire verbal properties. According to Tajima’s analysis (1985), which is based on a very large sample of ME writings covering the span 1100-1500, the verbalization of the gerund proceeded as follows. Around 1300 appeared the first instances with direct objects ((18)), and from the fifteenth century or later were found other verbal features, such as the ability to express distinctions of voice (1417 “without *being stolen*”; cf. Tajima 1985:113ff) and tense/aspect (1580-81 “after *having failed*”; cf. Tajima 1985:111-113, Fanego 1996b:129ff). Subject arguments in non-genitive form ((19)) occurred sporadically from Late Middle English, but remained very rare for a long time afterwards;⁸ in pre-verbal position, for instance, gerunds with an accusative pronoun subject, as in (20), are not recorded until the twentieth century:

- (18) c1300 (MS a1400) *English Metrical Homilies* 112/2-4 [Tajima 1985:76]:
 Sain Jon was [...] bisi In ordaining of priestes, and clerkes, And *in casting kirc werkes*
 “Saint John was [...] busy ordaining priests and clerics, and *in planning church works*”
- (19) c1400 *Laud Troy Book* 6317-18 [Tajima 1996:574]:
 he was war of *hem comyng* and of here malice
 “he was informed of them coming and of their wickedness”
- (20) 1932 D. Hammett *The Thin Man* 252 [quoted from Jespersen *MEG* V:§9.8.4]:
him hanging around like this, just messing things up, don’t fit in anywheres that I can see

Two features of ME *-ing* nominals are particularly relevant to the present research. The first concerns their close association with prepositional use,⁹ which can be traced back to at least Early Middle English (cf. Houston 1989). The data adduced by Expósito (1996:173-180) on the distribution of *-ing* nouns in Chancery English c1400-1450 are revealing in this regard: of the 135 nominal or partly nominal gerunds occurring in her 48,000-word corpus 81.50 per cent were found after a preposition, 12.60 per cent were

objects and a further 5.90 per cent subjects. These figures are in agreement with my own findings for the Early Modern period: in a sample of 317,621 words in the EModE section of the Helsinki Corpus of English Texts (Kytö 1996), I recorded 1,286 gerunds (= 79.50%) functioning as prepositional complements and 332 (= 20.50%) in other clause functions (Fanego 1996b:122-123).

Secondly, *-ing* nominals, like abstract nouns in general, could often be used without a preceding determiner in cases where one might be expected today; example (15) above and the first gerund in (18) illustrate this point (for details see Kisbye 1972:8ff, Fischer 1992:219-220, Fanego forthcoming a). Since in Middle English prepositional phrases, complement clauses and particles of various kinds could occur readily in both NP and VP structure, and, in addition, there was often no morphological distinction between adjectives and adverbs, some gerundial sequences lacking a determiner, such as those quoted in (21)-(24), could prove ambiguous between a verbal and a nominal reading:

- (21) ?a1200 *Old English Homilies* 151/17-18 [Tajima 1985:107]:
 þe teares þe man wepeð *for longenge to heuene* ben cleped rein water oðer deu water
 “the tears that man weeps in longing for heaven are called rain water or dew water”
- (22) ?a1300 (MS c1330) *Arthour and Merlin* 1301-02 [Tajima 1985:76]:
 þe messenger made anon asking *Whi he made swich leiþeing*
 “the messenger at once asked why he lied / asked the question why he lied”
- (23) 1340 *Ayenbite of Inwyt* 261/32 [Tajima 1985:102]:
at uerste guoinge in
 “when first entering”
- (24) c1280 *Southern Passion* 1874 [Tajima 1985:101]:
 Vnder þe Monument 3eo stod *wiþoute wepyng sore*
 “she stood close by the sepulchre without weeping bitterly / without bitter weeping”

I have argued elsewhere (Fanego forthcoming a) that constructions of this kind served as the locus for the reinterpretation of the ME *-ing* nominal as a form that was sufficiently verb-like as to govern direct objects and develop other verbal features. The actualization (Timberlake 1977, Harris/Campbell 1995:77ff, Harris 2003:532, 536-537) of the reanalysis, that is, the gradual manifestation of the innovated underlying structure at the observable level of language use, was a very complex process which lasted several centuries and was initially restricted only to those gerunds without an initial

determiner (henceforth: Type I gerunds) that were dependent on a preposition,¹⁰ as in example (18) repeated here for easier reference:

- (18) c1300 (MS a1400) *English Metrical Homilies* 112/2-4 [Tajima 1985:76]:
 Sain Jon was [...] bisi In ordaining of priestes, and clerkes, And *in casting kirc werkes*
 “Saint John was [...] busy ordaining priests and clerics, and *in planning church works*”

With this kind of gerund, direct objects and other verbal traits became increasingly more frequent in the course of time. By Late Middle English, direct objects were still rare: their proportion with respect to *of*-phrases was then less than 10 per cent (cf. Donner 1986). In the period 1570-1640 it had risen to 75 per cent and by 1700 it reached 95 per cent (see Fanego 1996b:128).

By contrast, gerunds with initial determiners (henceforth: Type II gerunds) remained prevalingly nominal until well into the Early Modern period. Thus, in my Helsinki Corpus data for the years 1570-1640 direct objects with Type II gerunds reached a mere 10 per cent (11 ex. out of 113; see Fanego 1998:106), though by 1700 they had increased to 52 per cent (61 ex. out of 117). The following are a few examples illustrating the gradual verbalization of this type:¹¹

- (25) HC 1602-1603 *Statutes* IV, 1028: And *for the better restrayninge the saide Offenses, and more severe punishinge the same*, be it further enacted [...]
 (26) HC 1702 R. Haddock, Sr. *Haddock Correspondence* 43: I do not find by your letter that you were with your Comodore *at the takeing and destroying the French shippes*
 (27) HC 1685 *Lisle Trial* IV, 123C1: And for that white-headed Man that speaks *of my denying them*, as I said before, he was one of them that rifled and plunder'd my House,

As can be observed in these quotations, the extension of direct objects and other verbal features to Type II gerunds gave rise to the emergence of constructions with a mixture of nominal and verbal properties; some of these (e.g. a possessive determiner occurring in combination with a direct object, as in (27)) are still possible in Present-day English, as pointed out in Section 2 of this paper.

From about the second half of the seventeenth century Type I verbal gerunds (e.g. *writing letters*) at last began to spread outside their original prepositional environment. (28)-(31) illustrate their use as object, subject and predicative respectively:

- (28) HC 1699 Langford *Plain and Full Instructions to Raise all Sorts of Fruit-Trees* Sample 2, 122: Many neglect *keeping the Heads of Stocks clay'd after the first time when they are grafted,*
- (29) HC 1699 Langford *ibid.* Sample 2, 114: *Slitting the bark* is an excellent additional help to most of the aforesaid evils, and also for bark-binding,
- (30) COPC 1778 Walpole *Life of Mr. Thomas Baker* 130/077-P19: The sufferings of Charles I whose crimes were not of the magnitude of his son's, had raised a spirit of enthusiasm in his partisans, and conjured up in their minds a profane idolatry of kings, [...] as if *being born of a certain race* could entitle any family to a right of violating with impunity all laws, both divine and human.
- (31) HC 1697 Vanbrugh *The Relapse* I, 59: Your Lordship does me too much honour, it was *exposing your Person to too much Fatigue and Danger*, I protest it was;

Occasional instances of this kind can be found in my data for Early and Late Modern English (Fanego 1996a-b, 1998, forthcoming b), but they remain rare even as late as the nineteenth century, as can be seen in Table 1. Further, the figures in that table reveal that the greatest restrictions on their use applied to the subject function. Despite the fact that, overall, sentential subjects in English are statistically more frequent than predicatives (cf. Elsness 1981; Fanego 1992:12), in my Late Modern English material Type I verbal gerunds in predicative function consistently outnumber subjects.

This restricted distribution of Type I verbal gerunds explains the use of constructions such as those in (32)-(38); these can be found especially in the last decades of the seventeenth century and in the eighteenth century, that is, coinciding with the years when the system of English gerundial complementation was being completely restructured as a result of the reanalysis that had taken place in Middle English.

- (32) COPC 1706 Kennett *History of England* 3G00 0012/102-P0: [...] on April the second, the lord marquis of Winchester reported to the house, that the committee had gone yesterday to Newgate; and had sent for sir William Perkins and sir John Friend severally, [...] and according thereunto did examine them: and that sir William Perkins owned *his being privy to the intended assassination, and of being in company when it was discoursed of at two or three meetings*; and thinks it was a fault that he did approve of it.
- (33) ECF 1766 S. Robinson Scott *The History of Sir George Ellison* 18: Slavery was so abhorrent to his nature [...] that he had hitherto avoided *the keeping any negroes* [...]. But the case was now altered; he had with his wife married a considerable plantation, cultivated by a numerous race of slaves, nor could his affairs go on without them. This much embittered his possession; and perhaps few have more severely lamented *their being themselves enslaved by marriage*, than he did *his being thus become the enslaver of others*.
- (34) COPC 1769 Granger *Biographical History* 0051/019-P1: Henry I, youngest son of William the conqueror, gained the crown by usurpation, and defended it with vigor and dexterity. His engaging person and address, his courage, learning, and

eloquence, have been much celebrated. The greatest blemish of his reign was *his putting out the eyes of his elder brother, and confining him twenty=eight years in Cardiffe castle in Glamorganshire.*

- (35) COPC 1768 Priestley *Principles of Government* 9G32 0043/055-P0: On the authority of the ancients, Dr Brown might as well contend for another institution of the famed Egyptians; viz, *their obliging all persons to follow the occupations of their fathers*;
- (36) HC 1664 J. Strype *Letters* 181: and therefore have endeavoured, and shall, the time providence hath allotted for my stay here, continue in the same endeavour: to wit, *of redeeming the hastening hours, and improving them*, so as that it may be for my advantage and credit.
- (37) 1766 J. Burn *A practical grammar of the English language* 154: All the precautions of prudence, moderation, and condescension which Eumenes employed, were incapable *of molifying the hearts of those barbarians or extinguish their jealousy.* [quoted from Sundby et al. 1991:395]
- (38) ECF 1780 Reeve *The Old English Baron* 41: One day, as they were walking in a wood near the castle, Edmund asked the father, what meant those preparations for building, *the cutting down trees, and burning of bricks?* —What, said Oswald, have you not heard that my Lord is going to build a new apartment on the west side of the castle?

What several of these examples have in common is that their authors resort to different strategies in order to provide the verbal gerunds with an introductory element of some kind. Thus in (32)-(35) the possessive determiners underlined (*his, their*) are pleonastic and add no information; their main role appears to be serving as licensers of the following *-ing* clauses. In (32) and (36), in turn, it is the preposition *of* which is used with this function.¹² Example (32) deserves special mention: despite the fact that the two coordinated gerunds share the same subject argument (*his*), the distance intervening between the genitive “licenser” *his* and the second gerund leads Kennett to insert an anacoluthic *of* in front of the latter. Finally, as shown in (37) and (38), the distance factor could at times bring about a complete change of construction (an infinitive clause in (37), a nominal gerund in (38)), rather than the introduction of a second licensing element.

Throughout the eighteenth century, the definite article was the element most commonly employed in this function of licenser of a verbal *-ing* clause when this was a subject, object, predicative or appositive; outside these syntactic environments, its use was very infrequent.¹³ Witness (33) (*the keeping any negroes*) and (38) above, and (39)-(45) here:

- (39) ECF 1749 Cleland *Memoirs of a Woman of Pleasure* 114: And at the same time kept my thighs so fast lock'd, that it was not for a strength like his to force them open, or do any good. Finding thus my advantages, and that I had both my own

and his motions at command, *the deceiving him* became so easy, that it was perfect playing upon velvet:

- (40) ECF 1725 Haywood *Fatal Secret* 217: believing that if he were refused there, he should never get an Opportunity of applying to the young Lady, he chose to make the first Declaration to herself; *the gaining her Affections* being the material Point, he considered all others of little Consequence.
- (41) COPC 1742 Fielding *Joseph Andrews* 066/105-P23: [...] a consideration which I take the liberty to recommend a little to the reader: for however swift his capacity may be, I would not advise him to travel through these pages too fast: for if he does, he may probably miss *the seeing some curious productions of nature* which will be observed by the slower and more accurate reader.
- (42) ECF 1743 Fielding *A Journey from this World to the Next* 59: [...] what should such Fellows as you do, where there are no Cities to be burnt, nor People to be destroy'd? But let me advise you to have a stricter Regard to Truth for the future, and not call *the depopulating other Countries* the Service of your own.
- (43) ECF 1744 S. Fielding *David Simple* 71: She certainly would soon have broke her Heart, had she known that all this Misery [...] was her own Fault; but as she thought it his Inconstancy, to his Generosity, in not telling her the Truth, she owed *the avoiding that painful Reflection*.
- (44) COPC 1711 *Spectator* Addison 114/040-P06: There is one piece of sophistry practiced by both sides, and that is *the taking any scandalous story that has been ever whispered or invented of a private man, for a known undoubted truth*, and raising suitable speculations about it.
- (45) NCF 1794 Godwin *Things As They Are* 147: Fraught with these prepossessions, the civilities that had once or twice occurred in the bustle of a public circle, *the restoring her fan which she had dropped, or the discommoding her of an empty tea-cup*, made her heart palpitate,

Clauses nominalized by attaching the definite article are attested in both Indo-European and non-Indo-European languages, as reported by Comrie/Thompson (1985:391-393), Koptjevskaja-Tamm (1993:49-50) and Lyons (1999:60-61). However, the aspect which I find of particular interest, and which brings me back to my initial remarks in Section 1 of this paper, is that in eighteenth-century English the nominalizing use of *the* illustrated in the above examples was associated in particular with pre-verbal subject clauses. In a corpus of 800,000 words covering the span 1700-1872, I recorded 21 subject clauses introduced by *the*, as in (39)-(40) above; 12 examples were objects (cf. (41)-(43)), 6 predicatives ((44)), 5 appositives ((38) and (45)), and 2 occurred in “absolute” use.¹⁴ If we compare the double figures for subjects with the much smaller number (6 ex.) of unintroduced subject gerunds reflected in Table 1,¹⁵ it seems clear that, for a while, the presence of an initial *the* was an important strategy to mark *-ing* subject clauses in English, as I argued at the beginning of this paper.

As I have shown elsewhere (Fanego forthcoming b), from the end of the eighteenth century normative pressures, coupled with complex language internal developments, led

to the disuse of all hybrid gerunds introduced by *the*. Some instances were simply reworded as verbal gerunds of Type I (i.e. structures like “he may probably miss *the seeing some curious productions of nature*” gave way to “he may probably miss *seeing some curious productions of nature*”), others were recast entirely and came to be expressed by means of *that*-clauses or *to*-infinitives, and yet others survived into Present-day English as purely nominal structures (e.g. *the signing of the contract*). The details of this change are irrelevant to the present discussion. Instead, my concern in the next sections will be examining the “meaning” of *the* and its grammatical status when used in the way exemplified above.

Table 1. Syntactic distribution in Late Modern English of verbal gerunds of Type I [based on four text samples ranging from 100,000 to 120,000 words. In brackets: normalized frequencies per 10,000 words].

	Subperiod 1 (1700-1726)	Subperiod 2 (1732-1757)	Subperiod 3 (1761-1797)	Subperiod 4 (1849-1872)
Pre-verbal Subject	1 (0.08)	2 (0.2)	3 (0.25)	0
Extraposited or dislocated Subject	1 (0.08)	1 (0.1)	1 (0.08)	4 (0.4)
in Object position	18 (1.5)	19 (1.9)	28 (2.33)	32 (3.2)
Subjective Predicative	5 (0.4)	7 (0.7)	4 (0.33)	5 (0.5)
Appositive	2 (0.17)	0	0	1 (0.1)
Absolute	0	0	0	4 (0.4)
governed by a preposition	519 (43.2)	507 (50.7)	529 (44.08)	254 (25.4)
TOTAL	546 (45.5)	536 (53.6)	565 (47.1)	300 (30.0)

Note: The relative frequencies of non-prepositional use increase across time, from 4.9% in subperiod 1 to 5.4%, 6.4%, and 15.3% in subperiods 2, 3 and 4 respectively.

4. The meaning and grammatical status of *the*

4.1 The meaning of *the*

Definiteness in English (Givón 1984:387ff, Foley/Van Valin 1985:283-287, Declerck 1991:320ff; Lyons 1999:2ff) has to do with contextual factors such as the speaker’s assessment of what the hearer knows and hence can be expected to identify, and also with the status of a NP as referring to a participant being introduced into the discourse (new information) or to a participant already established in the discourse (given information). Depending on these factors, a NP will be marked with *the*, *a* or zero.

While the above statements apply to most NPs, it has been observed that nominalizations in English are usually definite. As Nichols (1988:404) puts it,

definiteness is favored when the nominalization has a subject or object phrase with *of* [...]. Indefiniteness is favored only by nominalizations that are highly lexicalized (such as *explanation, goal, study*). These principles, and other doubtless involved, must follow from regular rules of English grammar which act together to create a statistical propensity for definiteness on nominalizations. This means that the speaker has little freedom to choose between definiteness and indefiniteness [...]

As I pointed out in the previous section, Middle English NPs headed by *-ing* nominals were often not marked for definiteness (cf. (15) above: *dredyng of our Lord*), possibly because the specifying *of*-phrases following them were enough to restrict their reference (cf. Declerck 1991:327, Fischer 1992:220). By the eighteenth century, however, usage had changed and the majority of the nominal gerunds occurring in my Late Modern English material take the definite article, whether they encode given information or not, as in (46)-(48):¹⁶

(46) COPC 1716 Addison *Freeholder* 080/030-P03: There are many reasons why the women of Great Britain should be on the side of the freeholder, and enemies to the person who would bring in arbitrary government and popery. As there are several of our ladies who amuse themselves *in the reading of travels*, they cannot but take notice what uncomfortable lives those of their own sex lead, where passive obedience is professed and practiced in its utmost perfection. In those countries, the men have no property but in their wives, who are the slaves to slaves:

(47) COPC 1710 Steele *Tatler* 087/035-P01: I went on Friday last to the opera, and was surprised to find a thin house at so noble an entertainment, [...] For my own part, I was fully satisfied with the sight of [...] signior Nicolini, who sets off the character he bears in an opera by his action, as much as he does the words of it by his voice. [...] He performs the most ordinary action in a manner suitable to the greatness of his character, and shows the prince even *in the giving of a letter or the dispatching of a message*.

(48) ECF 1725 Davys *Familiar Letters* 306: To ARTANDER. —I'll swear, Artander, I was never so merry in my Life, as *at the reading of your last Letter*;

Let us now turn for a moment to the hybrid gerunds introduced by *the* which are the concern of this section. It could of course be argued that structures like *the seeing some curious productions* (cf. (41)) and others of the same nature arise in English when the direct objects that had been available for some time with Type I gerunds (*by writing letters*) became extended to the nominal gerund type illustrated in (46)-(48).¹⁷ From this perspective, the article preceding the hybrid gerunds adduced in (39)-(45) above would thus be the functional continuation of the article required by nominal structures like

(46)-(48). This account is substantially correct and it is not my intention to contradict it: it is clear that the use of *the* in (39)-(45) can only be explained by reference to its use in (46) and similar examples. Yet the syntactic evidence discussed in the previous section suggests, I believe, that during the complex process of verbalization undergone by English gerundive nominals *the* temporarily took on a new function and became a grammatical marker that was in complementary distribution with the prepositions which so often introduced verbal gerunds (e.g. *by writing letters*). We may recall here that the distribution of *the* was by no means random: it was clearly associated with gerundive clauses functioning as core complements (subject, object, or predicative) of a higher predicate. If the complement in question was a subject, then the occurrence of *the* was highly predictable: at least in eighteenth-century English, the article was the default strategy for marking such clauses.

In its new function of grammatical marker or complementizer—the specific label we apply to it is irrelevant—, what was the semantic contribution of *the* to the overall construction? I pointed out earlier in this paper that the use of articles before complement clauses is attested in languages other than English. A case in point is Spanish, where some verbal infinitives can be introduced by the definite article *el*:

- (49) *El haber hecho el examen* no significa que estés aprobada.
 the to-have done the exam not implies that you-are passed
 “The fact that you have taken the exam does not imply that you have passed /
 Having taken the exam does not imply that you have passed.”
- (50) *Julia deplora el ser una carga para sus hermanos.*
 Julia regrets the to-be a burden for her brothers
 “Julia regrets being a burden on her brothers.”

As is clear from my translations, such clauses are always factive and presupposed (de Miguel 1996, Hernanz 1999:2274, 2279), whether they function as subject ((49)) or object ((50)). The use of the article, therefore, depends on the presence in the matrix clause of a predicate of the appropriate kind (i.e. factive), and the semantic contribution of *el* to the complement clause can thus be related to the contextual variable “givenness”, which, as pointed out at the beginning of this section, is a property of definite NPs.

It would be tempting to ascribe the same function to the definite article used in the English hybrid gerunds. Quite a few of the examples occurring in my eighteenth-century data are indeed factive; witness (3), repeated here for convenience, and (51):

- (3) COPC 1709 Berkeley *New Theory of Vision* 073/030-P23: *The not observing what has been delivered in the two last sections* seems to have occasioned no small part of the difficulty that occurs in the business of erect appearances.
- (51) ECF 1744 S. Fielding *David Simple* 105: He was soon weary of her, and then left her in a Condition the most unable to bear Afflictions, to suffer more than can be expressed. *The being forsaken by the Man she loved*, and the Horror of being discovered by her Father, made her almost distracted;

However, non-factive instances can easily be found. Several of the *-ing* clauses in (39)-(45) do not refer to “facts”, but to actions which, in some cases ((40)-(41)), have not yet been fulfilled. The gerund clause in (38) (“those preparations for building, *the cutting down trees*, and burning of bricks”) clearly refers to an activity in progress, as does, by definition, the nominal gerund coordinated with it. On the other hand, gerunds not introduced by *the* could be factive, as can be seen from (30) (“as if *being born of a certain race* could entitle...”). All this confirms, I believe, the status of *the* as an empty grammatical element serving to mark the boundary of its own clause, as I suggested at the beginning of this paper.

Late Modern English normative grammarians sometimes objected to un-introduced *-ing* clauses because they could prove ambiguous.¹⁸ Thus the American Gould Brown recommended avoiding their use and replacing them with an infinitive:

[p]articiples that have become nouns, may be used as such with and without the article, as, *spelling, writing, drawing*. But we sometimes find those which retain the government and the adjuncts of participles, used as nouns before or after verbs; as “*Exciting* such disturbances, is unlawful.” — “Rebellion is *rising* against government.” This mongrel construction is liable to ambiguity and ought to be avoided. (1845 *The Institution of English Grammar* p.162; quoted from Visser 1963-1973:§1039)

A participle used after *be, is, was, &c.* produces a construction which is more naturally understood to be a compound form of the verb; and which is therefore not well adapted to the sense intended, when one tells what something is, was, or may be. Example: ‘Whose business *is shoeing* animals’. Say, ‘Whose business it is to shoe animals’. (1851 *The Grammar of English Grammars* p.620; quoted from Visser 1963-1973:§1051)

Even if one opts for retaining the *-ing* clause, the resort to more explicit grammatical marking may help to ensure the speedy recognition of the immediate constituent structure, as shown in recent research by Rohdenburg (1995, 1996, 1999). This “transparency principle”, to use Rohdenburg’s label (1996:151), seems to be the factor prompting the use of the article *the*, as also of possessive determiners and prepositions, in many of the examples adduced earlier in this section.

4.2 The grammatical status of *the*

I have argued so far that during the complex process of verbalization undergone by the English gerund, the article *the*, which frequently occurred in NPs headed by *-ing* nominals, developed into an empty grammatical marker or subordinator. Since such shifts in function are characteristic of grammaticalization processes (Heine/Reh 1984, Heine/Claudi/Hünnemeyer 1991, Hopper/Traugott 2003[1993], Lehmann 1995[1982], Heine 2003), it is worth considering whether we can apply a grammaticalization analysis to the new use of *the* referred to above.

Prototypical instances of grammaticalization involve the change from open lexical categories such as nouns and verbs to closed relational categories such as adpositions, conjunctions and case markers. However, the development of adpositions or pronouns into complementizers, and likewise changes such as adposition > case affix, demonstrative > article, emphatic personal pronoun > clitic pronoun are also widely recognized as exemplifying the increased grammaticalization of already grammatical items, according to a cline of grammaticality which Hopper/Traugott (2003[1993]:7) represent as follows:

(52) content item > grammatical word > clitic > inflectional affix

It is also generally accepted that the grammaticalization of linguistic expressions involves several interrelated mechanisms (see Heine 2003:578-592), namely:

- i) *desemanticization*: loss in meaning content;
- ii) *extension* or *context generalization*: use in new contexts;
- iii) *decategorialization*: loss in morphosyntactic properties characteristic of the source forms, leading to structural effects such as *paradigmatization* (the tendency for grammaticalized forms to be arranged into increasingly small paradigms), *fixation* (the fact of free linear ordering becoming fixed), and loss of independent word status (cliticization, affixation);
- iv) *erosion* or *phonetic reduction*, that is, loss in phonetic substance.

As Heine (ibid.:579) notes, each of these mechanisms is concerned with a different aspect of language structure or language use: (i) relates to semantics, (ii) to pragmatics, (iii) to morphosyntax, and (iv) to phonetics. In the course of time, elements in grammaticalizing constructions tend to undergo all four processes to a greater or lesser

extent, but in the early stages of grammaticalization there may be desemanticization “although there are as yet no noticeable pragmatic, morphosyntactic, or phonetic changes” (ibid.:580).

The developments taking place under grammaticalization are largely unidirectional; the shifts in meaning (Heine 1997:6, 2003:579, 583, 591-592), for instance, evolve from concrete, or less abstract meanings, to more abstract, grammatical meanings. Thus, Traugott/König (1991:199ff) trace the development of the OE accusative noun *hwile* “time” into a temporal conjunction (*they arrived while I was sunbathing*) and eventually into a concessive marker (*while you like peaches, I like nectarines*), and point out the overall shift from reference to a relatively concrete state of affairs (a particular time) to expression of the speaker’s assessment of the relevance of simultaneity in describing events, to assessment of contrast between propositions. In the same way, the development into a complementizer of the Ancient Greek pronominal correlative *ὅτι*, “which referred anaphorically to a coreferential topical pronoun in the preceding clause and introduced a focalized specification about it” (Cristofaro 1998:67), involves a shift from an item used to refer “to individual entities on a deictic or textual level [to a] conjunction only signify[ing] abstract logical relations” (ibid.:82).

Finally, also distinctive of the changes taking place under grammaticalization is that they proceed gradually, and that the resulting constructions tend to have stronger internal dependencies;¹⁹ as Haspelmath (1998:318) puts it,

grammaticalization is the gradual drift in all parts of the grammar toward tighter structures, toward less freedom in the use of linguistic expressions at all levels. [...] constructions become subject to stronger constraints and come to show greater cohesion.

Coming back to the use of *the* which is the concern of this section, there is no evidence that the transition of *the* from article to complementizer was gradual; rather, the shift in function appears to have been sudden and abrupt. As a result of this dramatic switch in class *the* ceased to be a member of the paradigm of articles and joined the class of complementizers, but it is doubtful that we can speak here of *paradigmaticization*: as already noted, the most evident aspect of this parameter for measuring grammaticality is the sheer size of the paradigm which is joined by the grammaticalizing element (cf. Lehmann 1985:307-309), and both articles and subordinators belong to very small, closed classes. In this respect, therefore, the

changes affecting *the* do not exhibit one of the most distinctive properties of grammaticalization, namely the shift from a more open to a more closed class.

The second of the mechanisms mentioned above, namely, context generalization, does not seem to apply to subordinating *the*, either. Though in verbal gerunds (*the deceiving him*) the syntactic relationship between *the* and the gerund clause is different from the relationship holding between a determiner and the following phrase, the sentence slots (i.e. subject, object, predicative, etc.) where the new verbal construction occurs are the same as those where the nominal gerunds could be found. It is therefore only on the semantic level that one might perhaps say that the article *the* is less grammatical than subordinating *the*: definite articles are anaphoric items whose referents are established earlier in the discourse, and hence have a textual function which might be considered less abstract than the function of subordinators or complementizers as mere indicators of grammatical relations.

There is, however, one major problem for the above analysis. I pointed out in Section 4.1 of this paper that in the eighteenth century most *-ing* nominals were definite; this means that the speaker, as aptly noted by Nichols (1988:204) for Present-day English, had “little freedom to choose between definiteness and indefiniteness on nominalizations”. In the absence of widespread functional contrasts between *-ing* nominals with and without the definite article, the presence of this before an *-ing* noun had, therefore, virtually no discriminatory semantic value or communicative function. It is thus not clear in what sense one could say that the article *the* occurring with nominal gerunds (*the reading of travels*) had greater semantic content than the subordinating *the* introducing hybrids (*the deceiving him*). This suggests that a grammaticalization analysis for subordinating *the* is inadequate²⁰ and that we have to account for it in some other form. My proposal in the next section is to view it as an instance of exaptation, in the sense of Lass (1990, 1997).

4.3 *The* and exaptation

The term *exaptation* was borrowed by Roger Lass (1990, 1997:316ff) from evolutionary biology. Exaptation, as Lass (1997:316) notes,²¹

is opportunistic: it is a kind of conceptual renovation, as it were, of material that is already there, but either serving some other purpose, or serving no purpose at all. Thus perfectly ‘good’ structures can be exapted, as can junk of various kinds.

Basic to the definition of exaptation is linguistic novelty: exaptation is “conceptual invention” (ibid.:319). In a typical case, the material exapted is at the point of exaptation doing “something else (which it may continue to do); but it is still capable of being remanufactured or restructured, and still exapted, in a sense, *as part of a different kind of coexisting structure*” (p.320) [emphasis added].

Lass (1990:99) mentions as an example of potentially exaptive material the *-s* ending of the English 3rd person singular present tense. Since English word-order is now fixed and subject pronouns are obligatorily expressed, this relic inflection has “none of the informative concordial function that verb suffixes had in earlier times... it is a systemic excrescence... But the exaptive impulse is strong”. Croft (2000:129), who discusses exaptation under the name of *hypoanalysis*, shows that, in fact, the exaptation of the *-s* suffix has already occurred in some nonstandard varieties of English where *-s* is now used for specific or single events (e.g. *I sees the doctor tomorrow*), and thus contrasts with other verb forms denoting iterative or habitual actions (e.g. *I do see him every day*).

A well-known exaptive change is the development of the obligatory definite article *k-* of some Nilo-Saharan languages into a nominalizer deriving verbal nouns, such as Ngambay Mundu *k-usa* “act of eating” from *usa* “eat” (cf. Greenberg 1991). Articles in Nilo-Saharan have reached the stage of development that Greenberg (1978:69) identifies as Stage III: they occur with virtually all nouns and “no longer [have] any synchronic connection with definiteness or specificity”. Since they are a mere sign of nominality on the large majority of common nouns, there exists “the possibility of reinterpretation of the articulated form and the attribution to it of new functions” (Greenberg 1991:305).²² The *k-* prefix is thus “attributed the contextual property of nominality that is inherent to the nouns it is invariable found with, and then comes to be used as a nominalizer” (Croft 2000:129).

The exaptive development of Nilo-Saharan *k-* has clearly much in common with the changes affecting the definite article in eighteenth-century English, except for the fact that the lack of communicative function and discriminatory semantic value of *the* did not obtain with all classes of English nouns, but only with the specific subset of *-ing* nominals. This, I believe, enabled *the* to be reinterpreted as a grammatical marker of subordination which came to be employed chiefly with gerunds functioning as core

complements (i.e. subjects, objects or predicatives) and was thus largely in complementary distribution with the prepositions introducing all other verbal gerunds.

Norde (2001, 2002) draws attention to some cases of morphological exaptation in Swedish. The Old Swedish plural suffix *-on*, originally the marker of the nominative and accusative plural of weak neuter nouns, as still in Modern Swedish *ögon* “eyes”, was exapted for derivation and became a formative of berry-names such as *hallon* “raspberry” < *hall* “stony ground”. Swedish has undergone a process of deflexion that covered more than three centuries, and has developed from a comparatively heavily inflected language to a language with very little inflection. Norde (2002:49) suggests that exaptation and other grammatical changes²³ she examines may be facilitated by favourable language-internal circumstances such as a “disruption of the system” (= *Systemstörung*, a term borrowed from Plank 1995). Swedish deflexion can appropriately be identified as a kind of disruption of the system, and so can, of course, the process whereby in English the nominal system of gerundial complementation was entirely restructured and replaced by a new, chiefly verbal, system.²⁴

5. Impertinent *by* and grammaticalization

5.1 Impertinent *by*-sentences characterized

As noted in the opening lines of this paper, Hogan/Hogan (1998) document the increasing occurrence in American undergraduate writing of the sentence type illustrated in (1)-(2) above, repeated here for easier reference:

- (1) *By trying to make his mother happy* proved unlucky for the character Paul in this story.
- (2) *By interviewing buyers of instant coffee and buyers of regular coffee* is akin to asking people whether they prefer the taste of Coke or Pepsi.

The construction is rare in the writing of that segment of the undergraduate population “who are accomplished, careful writers” (p.183), but often appears in the prose of more inexperienced writers “who seem to have no other major confusion about the conventions of sentence structure in standard written English” (ibid.).²⁵

On average, fifteen impertinent *by*-sentences can easily be found “in any batch of 400-500 pages of undergraduate student writing” (ibid.). Four or five hundred pages of handwriting,²⁶ at about 250 words per page, gives a figure of some 100,000/125,000 words. This confirms that impertinent *by*-sentences cannot be dismissed as sporadic

errors, but should be considered a productive syntactic tool among many American undergraduates.

It would have been interesting to know, too, the average frequency of “normal” *-ing* subjects (e.g. “*trying to make his mother happy* proved unlucky for Paul”) in Hogan/Hogan’s written material, but they say nothing on this point. To judge from Huddleston (1971:194-195), gerundial subjects are not very common in Contemporary English: in a 135,000-word corpus of written scientific English dating from the mid nineteen sixties, he met only seven instances.²⁷ We can conclude, therefore, that for those speakers or writers who have the impertinent *by* as part of their grammars, this is an important strategy—maybe the *main* strategy—to ensure placement of an *-ing* clause in subject position.

Throughout their study, Hogan/Hogan quote a total of 21 *by*-sentences²⁸ which they do not classify into further subgroups. It is possible, however, to discern a number of different structural patterns, as follows:

- a) *By* + *-ing* clause as subject of the matrix predicate (13 ex.). Two subtypes:
 - i. the *-ing* clause has no expressed subject (10 ex.); cf. (1)-(2) above and (53);
 - ii. the subject of the *-ing* clause is a plain case NP (2 ex.) or is in the possessive form (1 ex.). Cf. (54) below.
- b) *By* + *-ing* clause dislocated to the left; an anaphorically linked pronoun functions as subject of the higher predicate (6 ex.). Cf. (55)-(56) below.
- c) *By* + factive *that*-clause as matrix subject (1 ex.). Cf. (57).
- d) *By* + NP as matrix subject (1 ex.). Cf. (58).

(53) *By simply taking an average of the survey responses on why customers purchased the product*, would be to disregard the great extremes customers indicated.

(54) *By the company not addressing the problem* would tend to make customers take their business elsewhere.

(55) *By banning Welsh*, it was like taking away part of their culture and ancestry.

(56) Finally I realized that teachers are people too and *by me trying to huff and puff at them* that just made matters worse.

(57) *By the fact that he is cool under fire* shows that the author intended him to have heroic qualities.

(58) *By the way elections are sequenced every four years* prevents one party from gaining permanent control.

Finally, with respect to the matrix predicates that “can elicit the *by*-construction” (p.185), Hogan/Hogan identify two main classes. One is a group of causatives illustrated by *make*, *affect*, *change* and *prevent* (see (54), (56) and (58) above), and the other is a group of “semiotic verbs” (p.186) such as *show*, *mean* or *indicate* (cf. (57)). In addition, some *by*-sentences contain a comparative expression (e.g. *be akin*, *be like*; cf. (2) and (55)), and these are considered by Hogan/Hogan as a subset of the semiotic type.

5.2 Impertinent *by*-sentences as grammatical blends

To account for the construction under discussion, Hogan/Hogan (1998) resort to the cognitive process of *conceptual blending*, as defined by Fauconnier/Turner (1996, 1998). These analyse constructions such as the Caused-Motion Construction (e.g. *they prayed the boys home*; cf. Goldberg 1995) as *grammatical blends* created to express some new conceptual integration of events into one complex event. In blending, “structure from two input [mental] spaces is projected to a separate space, the ‘blend’. The blend inherits partial structure from the input spaces, and has emergent structure of its own” not provided by the inputs (Fauconnier/Turner 1996:113). Thus, in the case of example (1) above, the two input structures from which the blend is assumed to derive would be roughly of this form (Hogan/Hogan 1998:196):

- (1) a. *Trying to make his mother happy* proved unlucky for Paul.
 b. *By trying to make his mother happy* Paul became unlucky.

For Hogan/Hogan (1998:181), their data on the resulting grammatical blends

support the conclusion that Americans are reorganizing their intellectual concepts of causality to accommodate the concept that generic situations or circumstantial processes are causal agents, more so than human individuals or groups.

Similarly, Fauconnier (1997:176) points out that “conceptually, the new construction [with *by*] is able to carry the additional information that the subject NP is a means for somebody. Syntactically, we have a novel structure with a *by*-phrase in subject position”.

Though, as will be seen below, *by*-sentences could appropriately be considered as resulting from the merger of the two constructions in (1a-b), it is doubtful that they can be analysed as “grammatical blends”, in the specialized sense which is attached to this term within Fauconnier/Turner’s framework. Basic to the definition of grammatical

blending is the idea that the blend expresses “complex *novel* events” (Mandelblit 2000:199); in other words, the grammatical form that signals integration can be used to express the previously unintegrated events and “new relations [are made] available that did not exist in the separate inputs” (Fauconnier 1997:150). However, none of this applies to impertinent *by*-sentences; though these are syntactically novel structures — English grammar does not license *by*-clauses as subjects—, they do not represent a semantic innovation or a new integration of events, since exactly the same conceptual information is expressed in Standard English by the corresponding constructions without an initial preposition:

- (1a) *Trying to make his mother happy* proved unlucky for Paul.
 (54a) *The company not addressing the problem* would tend to make customers take their business elsewhere.

In the literature on grammaticalization it is widely recognized that grammatical changes often replicate or parallel earlier changes: “there are indeed many examples suggesting that, once a given grammatical form declines and/or disappears, a new form tends to be recruited on the same conceptual pattern as the old one, with the result that a kind of morphological cycle emerges” (Heine/Claudi/Hünemeyer 1991:246; see also Hopper/Traugott 2003[1993]:21). As I see it, the factor prompting the use of impertinent *by* in undergraduate American writing has much to do with the one discussed earlier in this paper in connection with hybrid *the*-gerunds, namely the desire to provide an explicit marker of subordination for clauses which, because of their role as pre-verbal subjects, are cognitively very complex. In the process, it is clear that the semantic content of the erstwhile preposition *by* has shifted, as is shown by its occurrence in contexts (e.g. (2), (57), etc.) where its original meaning of means is, at best, tenuous. At the same time the *by*-clause itself has evolved from an adjunct of means functioning in the clause periphery to a complement serving as subject argument. These developments, which can be considered paradigm examples of grammaticalization, are discussed in the next section.

5.3 The grammaticalization of impertinent *by*-sentences

Recent work on grammaticalization emphasizes that lexical or grammatical items develop grammatical functions only in specific constructional contexts: “grammaticalization [is] the process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts is assigned grammatical function” (Traugott 2003:645). As Croft (2000:156) notes, grammaticalization is thus essentially syntagmatic.

With this in mind, let us consider in the first place the basic semantic features of syntactic constructions involving gerundial subjects. Since no major changes seem to have taken place in this regard for quite some time, throughout the following discussion I will make use of data from both Late Modern and Present-day English.

Not rarely, *-ing* subjects form part of a larger construction with *be*, or a related copular verb, in either its ascriptive or specifying functions. Two eighteenth-century examples illustrating this point are (39) (“*the deceiving him* became so easy...”) and (40) (“*the gaining her Affections* being the material Point...”). Far more commonly, however, the matrix predicate belongs to one of the two semantic classes identified by Hogan/Hogan (1998), namely:

- a. causatives like *make*, *affect*, *change*, *occasion* or *prevent*. See examples (3), (45), (51), (54), (56) and (58) above, among others.
- b. pseudo-specifying verbs like *be*, *show*, *mean* or *indicate*, or related periphrastic expressions like *be akin*, *be like*. See (2), (38),²⁹ (55) and (57).

Causatives need little comment. Pseudo-specifying predicates (= semiotic predicates, in Hogan/Hogan’s terminology) or, more properly, pseudo-specifying sentences, are sentences in which both the subject and the internal complement are clausal.³⁰ Where the verb is *be*, as in the following grammatical versions of (53) and (55) above, they look like specifying constructions, but differ from these in that they are not reversible (cf. **Taking away part of their culture and ancestry was banning Welsh*):

(53a) *Simply taking an average of the survey responses on why customers purchased the product* would be to disregard the great extremes customers indicated.

(55a) *Banning Welsh* was taking away part of their culture and ancestry.

As already noted by Jespersen (1924:154) with respect to the idiomatic *seeing is believing*

we may be inclined to take *is* as implying identity [...] but the identity is more apparent than real. It would be impossible to invert the terms, and the logical purport of the saying is merely this: seeing immediately leads to, or causes, belief.

Pseudo-specifying sentences indicate, therefore, “what the situation described in the subject entails or necessarily involves” (Huddleston/Pullum 2002:1256), and can generally be paraphrased by means of an *if... then* construction: “if you ban Welsh, you take away...”.

It will be clear from the above characterization of sentences with an *-ing* subject that they have much in common with constructions containing *by*-adjuncts of means, since in both cases there is an implication that a situation or event is contingent upon another event; compare, for instance, *his leaking the report to the media will sabotage the plan* with *he will sabotage the plan by leaking the report to the media*. These similarities explain, of course, the recruitment of *by* to introduce gerundial subject clauses in American college writing; as pointed out by Hopper/Traugott (2003[1993]:186), “the grammaticalization of items [...] is constrained by the grammatical function to be expressed, and by the appropriateness of the inferences from the source items for the function in question”.

In order to trace the possible stages in the development of *by* from its role of preposition introducing an adjunct of means to its new role as clause subordinator or complementizer, I conducted a computer search of the matching BROWN and FROWN corpora. These are one-million-word samples of written American English dating back, respectively, to the years 1961 and 1992 and containing texts from fifteen different register categories (for details, see Mair 2002). My search confirmed that, as expected, impertinent *by*-sentences have not yet made their way into Standard English: neither BROWN nor FROWN yielded any examples. Both corpora, however, showed that *-ing* adjuncts introduced by the preposition *by* are very frequent in written English: each contained well over 300 instances.³¹ This finding was predictable, in view of the fact that in Huddleston’s corpus of written scientific English (see Section 5.1 above), *by*, with 121 occurrences, was by far the most common preposition introducing *-ing* clauses; the two next most common items, *of* and *in*, numbered 60 and 50 instances respectively (Huddleston 1971:209).

BROWN and FROWN also revealed two other interesting facts. First, *by Xing* adjuncts are often presupposed —as is also the case with many subject clauses—, and hence topicalized (43 ex. in BROWN, 39 ex. in FROWN), as in (59)-(60):

- (59) FROWN [Press Editorial] B17 87: By every index, poverty, high crime and broken families continue to plague white communities, too, but, *by transforming all of these problems into black ‘pathologies’*, it is easier for Americans to think of blacks as a ‘problem people,’
- (60) FROWN [Popular Lore] F03 15: A sextant is made so that you can determine the angle between the horizon, a star and the point where you’re located. *By doing this with two different stars*, you get two different angles.

Secondly, the *by Xing* string can at times function as a complement tightly integrated in clause structure, rather than as a peripheral adjunct. Two different sets of constructions can be distinguished here. One involves passive sentences containing a *by*-gerundial clause:

- (61) FROWN [Popular Lore] F07 88: A detector outside the incinerator filters out background radiation from the 2,200 degree F flame and picks up the distinctive PAC ‘signatures’. Under real operating conditions, a surge of these hardy compounds could be destroyed *by releasing a blast of oxygen or hydrogen peroxide* [...] Or the incinerator gases could be diverted through special scrubbers.
- (62) FROWN [Religion] D07 101: I assumed that like other works of art, the name ‘conversion’ contained elements of the process by which it was made and that they could be unpacked *by analyzing the artifact*. As a technical word in the language of manufacture, conversion denoted a variety of processes.

As noted by Croft (1991:237ff) and others, English uses the same marker —the preposition *by*— to encode means adjuncts (e.g. *she travels to work by bus*) and the type of VP-internal complement usually known as passive agent (e.g. *she was hated by her boss*). In most cases, the distinction between them is straightforward, both semantically and syntactically. However, in sequences such as (61)-(62) the differences are neutralized; in (61), for instance, the *by*-phrase can be interpreted as a means adjunct matching the means adjunct in the active sentence quoted as (61a). But it could also represent the passive agent corresponding to the active subject of (61b):

- (61a) Someone could destroy a surge of these hardy compounds *by releasing a blast of oxygen or hydrogen peroxide*.
- (61b) *Releasing a blast of oxygen or hydrogen peroxide* could destroy a surge of these hardy compounds.

In the other type of structure referred to above the *by*-phrases serve as predicative complements within a specifying construction; in each case, the subject of the sentence

contains head nouns such as *method*, *treatment* or, especially, *way* (cf. also example (58) in section 5.1). Since these nouns themselves denote means of doing something, the semantic contribution of *by* to the overall construction is minimal; in some instances it could in fact be argued that *by* has undergone desemanticization (cf. section 4.2), its chief role being to mark the clause boundary,³² as was the case with the article *the* occurring in an eighteenth-century example such as (44) above (“one piece of sophistry practiced by both sides is *the* taking any scandalous story...”):

- (63) The usual method of obtaining TO(T), the transverse optical mode at the zone center, is *by analyzing measurements of the residual ray reflectivity band with the aid of classical dispersion theory*. [quoted from Huddleston 1971:203]
- (64) BROWN [Skills and Hobbies] E08 1630: Tire size can be determined in several ways but the one that is the easiest and as accurate as any is *by measuring the effective radius of a wheel and tire assembly*.
- (65) FROWN [Miscellaneous] H25 216: According to Mary Smith of NCC, one way her organization has dealt with this issue was *by having a strong community action group* that worked hard to let elected officials know what they thought was best for their community.

As I see it, the three uses of *by*-phrases which I have just exemplified may have contributed, each in its own way, to the rise of impertinent *by*-constructions. The choice of *by* as an overt marker of gerundial subject clauses among American undergraduates could be receiving support from the close semantic relationship existing between the apparent passive agents in structures such as (61)-(62) and the corresponding active subjects. Secondly, it is worth considering examples (63)-(65) in the light of Cristofaro’s (1998:68-72) observations regarding the development of Ancient Greek *ὡς* from a relative pronominal in the instrumental form meaning “by means of which, the way in which” into a complementizer. According to Cristofaro, this use originated in a limited number of contexts through the conventionalization of some pragmatic implicatures: the Greek equivalents of clauses such as *he sang in which way the Achaeans destroyed the fortress*, where the communicative focus was on how the subordinate event took place, rather than on the fact that it actually took place, could be reinterpreted as asserting the fact that the event had actually taken place, thus leading to a neutralization of the semantic distinction between “how, in which way” and “that”, and to the eventual spread of *ὡς* to complement constructions.

In much the same way, since in (63)-(65) above the *by*-phrase functions as an obligatory constituent in clause structure, and *by* itself is largely redundant because of the means component inherent in the nouns *method* and *way*, the preposition might be

inferred to have complement properties. Such contexts, therefore, could represent the ground for generalizing the use of *by* to other classes of complements displaying similar semantic features, in particular to *-ing* subjects, where a causal or means component, as repeatedly noted, is often implicit.

Finally, turning now to (59) and similar examples, only minor syntactic adjustments are needed to convert them into impertinent *by*-sentences:

(59a) FROWN [Press Editorial] B17 87: By every index, poverty, high crime and broken families continue to plague white communities, too, but, *by transforming all of these problems into black ‘pathologies’*, this makes it easier for Americans to think of blacks as a ‘problem people,’

As I pointed out in Section 5.1, some of the examples of impertinent *by* adduced by Hogan/Hogan (1998) are of this type: the *by*-clause is disjoined from the grammar of the sentence by left-dislocation and is recapitulated by an anaphoric pronoun such as *it*, *that* or *this* functioning as subject of the matrix predicate (cf. (55)-(56)). It seems reasonable to assume that these structures, in which the *by*-clause is a “satellite” (cf. Koster 1978:57ff) rather than a constituent fully integrated in clause structure, represent a transitional stage towards the type illustrated in (1) (*by trying to make his mother happy proved unlucky for Paul*) and similar examples. Further research on American college writing would of course be needed in order to clarify this and other points, including, for instance, the apparent extension of *by* to factive *that*-clauses (cf. (57)) and even to NPs with the appropriate semantic characteristics (cf. (58)).

Before concluding this section, let me just point out that for those writers who have the impertinent *by* as part of their grammars, the new construction offers the obvious advantage not only of explicitly marking the start of the *-ing* clause, but also of providing a syntactic governor for the *-ing* clause subject when this is overt, as in (56) above (“*by me trying to huff and puff at them that just made matters worse*”), especially if the subject in question is a pronoun in the objective case. In other words, there can be little doubt that the difficulties of language users to handle structures where an apparent object comes at the head of the clause must also have been instrumental in promoting the use of *by*-constructions; in (56) and similar examples, *by* is thus serving a role exactly analogous to that of the erstwhile preposition *for* in infinitival clauses such as “*for him to refuse Jane’s request would be unthinkable*”.

To sum up, the grammaticalization of the construction which has been the concern of this section involved the following changes:

- a) The *by*-clause has lost the positional mobility which characterized it when it was a prepositional adjunct, and can only appear sentence initially; cf. section 4.2 above and Lehmann's (1995[1982]:164) parameter of *fixation*.
- b) Its syntactic category has changed, as shown by the fact that, like NPs and other nominal constituents, but unlike prepositional phrases, it can: i) function as matrix subject; ii) undergo left-dislocation and be subsequently referred to by an anaphoric pronoun.³³ Both the subject slot and the left-detached position in the sentence are contexts where the *by*-clause could not be used before.
- c) When used in its new role as subject, the *by*-clause is more tightly integrated in the structure of the sentence than when it was a means adjunct (cf. Haspelmath's definition of grammaticalization in section 4.2). In terms of hierarchies of clause linkage such as the Interclausal Syntactic Relations Hierarchy (Foley/Van Valin 1984:238ff, Van Valin/LaPolla 1997:442ff, 670; see also Lehmann 1988), the level of juncture of the *by*-subject is the clause core ("core juncture"), that of the original *by*-adjunct the clause itself ("clausal juncture").
- d) The semantic content of the preposition *by* has also changed. From being an element coding circumstantial meanings (i.e. "by means of"), which are part of the ideational dimension (Halliday 1994:151, 237) of language, it has come to serve mainly a textual role as marker of the following *-ing* clause. As is typical of grammaticalization processes, this desemanticization of *by* appears to have taken place gradually, as suggested by structures such as (63)-(65), where the means component of the preposition *by* is weakened.
- e) *By* has shifted from a relatively open set, that of prepositions, to a much smaller set of subordinators or complementizers; cf. Lehmann's (1995[1982]:132ff) parameter of *paradigmaticization*.

6. Concluding remarks

Sentences containing complement clauses are often described as structures involving the integration of two events, one coded in the main clause, the other in the complement clause, to yield a complex event. "Not surprisingly", as Givón (1993:2) notes, "this complex event is in turn coded by a complex syntactic structure".

The specific type of complex syntactic structure which has been the concern of this research —sentences with a pre-verbal *-ing* subject— differs from all other related constructions in English in the absence of a marker to signal the start of the subject clause. I have argued in this paper that two different syntactic developments, one taking place in Late Modern English, the other in recent American English, may be seen as manifestations of a cognitive principle of “transparency” (Rohdenburg 1995, 1996) prompting the resort to more explicit grammatical marking for such clauses.

I have also proposed that the frequent use of the article *the* as a subordinator introducing *-ing* subjects and other core complements during the Late Modern English period can be considered as an instance of exaptation, in the sense of Lass (1990, 1997). Previous research on exaptive changes suggests that these may be facilitated by favourable language-internal circumstances such as, for instance, a disruption of the system (cf. Plank 1995, Norde 2002), a label that may quite appropriately be applied to the replacement of the nominal system of gerundial complementation inherited from Old English by a new verbal system.

Exaptive *the* went out of use by the mid nineteenth century, but in certain varieties of English another grammatical form, the preposition *by*, has recently been recruited as a marker of pre-verbal *-ing* subjects, thus replicating the earlier change and giving rise to “a kind of morphological cycle” (Heine/Claudi/Hünemeyer 1991:246) for which many analogues can be found in the literature on grammaticalization. The rise of subordinating *by* parallels earlier changes taking place in the history of English in yet another respect, namely, in that the subordinators *to* and *for* which mark English infinitival complements have likewise developed out of the corresponding prepositions.

As noted in Section 3 above, subordinating *the* became obsolete from the mid nineteenth century as a result of both normative pressures and language internal developments. Prescriptive norms continue to be very much in force today, so college English handbooks have recently started to earnestly entreat undergraduate students to avoid impertinent *by*-sentences (Hogan/Hogan 1998:184-185). This suggests that these may never become entrenched and grammatical in Present-day English, and might ultimately suffer the same fate as their predecessors.

References:

- Aarts, Bas and Charles F. Meyer, (eds.) 1995. *The verb in contemporary English: theory and description*. Cambridge: Cambridge University Press.
- Blake, Norman, (ed.) 1992. *The Cambridge history of the English language*, Vol. II: 1066-1476. Cambridge: Cambridge University Press.
- Comrie, Bernard and Sandra A. Thompson. 1985. "Lexical nominalization". In: Timothy Shopen, (ed.), *Language typology and syntactic description*, Vol. III: *Grammatical categories and the lexicon*, 349-398. Cambridge: Cambridge University Press.
- Cristofaro, Sonia. 1998. "Grammaticalization and clause linkage strategies. A typological approach with particular reference to Ancient Greek". In: Anna Giacalone Ramat and Paul J. Hopper, (eds.), 59-88.
- Cristofaro, Sonia. 2003. *Subordination. A typological study*. Oxford: Oxford University Press.
- Croft, William. 1991. *Syntactic categories and grammatical relations. The cognitive organization of information*. Chicago and London: The University of Chicago Press.
- Croft, William. 2000. *Explaining language change. An evolutionary approach*. Harlow, Essex: Longman/Pearson.
- Croft, William. 2001. *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Declerck, Renaat. 1991. *A comprehensive descriptive grammar of English*. Tokyo: Kaitakusha.
- Diessel, Holger and Michael Tomasello. 2001. "The acquisition of finite complement clauses in English: a corpus-based analysis". *Cognitive Linguistics* 12:97-141.
- Disterheft, Dorothy. 1980. *The syntactic development of the infinitive in IE*. Columbus, OH: Slavica.
- Disterheft, Dorothy. 1990. "The role of adaptive rules in language change". *Diachronica* VII:181-198.
- Donner, Morton. 1986. "The gerund in Middle English". *English Studies* 67: 394-400.
- Elsness, Johan. 1981. "On the syntactic and semantic functions of *that*-clauses". In: Stig Johansson and B. Tysdahl, (eds.), *Papers from the First Nordic Conference for English Studies*, 281-303. Oslo: University of Oslo.
- Elsness, Johan. 1984. "*That* or zero? A look at the choice of object clause connective in a corpus of American English." *English Studies* 65:519-533.

- Expósito González, M^a de la Cruz. 1996. *La estructura del sintagma nominal en el inglés de la cancillería: 1400-1450*. Barcelona: Kadle Books.
- Fanego, Teresa. 1990. "Finite complement clauses in Shakespeare's English, Parts I and II", *Studia Neophilologica* 62/1:3-21, 62/2:129-149.
- Fanego, Teresa. 1992. *Infinitive complements in Shakespeare's English*. Santiago de Compostela: Universidad de Santiago de Compostela.
- Fanego, Teresa. 1996a. "The development of gerunds as objects of subject-control verbs in English (1400-1760)". *Diachronica* XIII:29-62.
- Fanego, Teresa. 1996b. "The gerund in Early Modern English: evidence from the Helsinki Corpus". *Folia Linguistica Historica* 17:97-152.
- Fanego, Teresa. 1998. "Developments in argument linking in Early Modern English gerund phrases". *English Language and Linguistics* 2:87-119.
- Fanego, Teresa. Forthcoming a. "On reanalysis and actualization in syntactic change: the rise and development of English verbal gerunds". *Diachronica* XXI.
- Fanego, Teresa. Forthcoming b. *Gerundial complementation in English: historical development and present-day usage*.
- Fauconnier, Gilles. 1997. *Mappings in thought and language*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles and Mark Turner. 1996. "Blending as a central process of grammar". In: Adele E. Goldberg, (ed.), *Conceptual structure, discourse and language*, 113-130. Stanford, California: CSLI Publications.
- Fauconnier, Gilles and Mark Turner. 1998. "Conceptual integration networks". *Cognitive Science* 22:133-187.
- Finegan, Edward and Douglas Biber. 1995. "That and zero complementisers in Late Modern English: exploring ARCHER from 1650-1990". In: Bas Aarts and Charles F. Meyer, (eds.), 241-257.
- Fischer, Olga. 1992. "Syntax". In: Norman Blake, (ed.), 207-408.
- Foley, William A. and Robert D. Van Valin, Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Foley, William A. and Robert D. Van Valin, Jr. 1985. "Information packaging in the clause." In: Timothy Shopen, (ed.), *Language typology and syntactic description*, Vol. I: *Clause structure*, 282-364. Cambridge: Cambridge University Press.

- Giacalone Ramat, Anna, and Paul J. Hopper, (eds.) 1998. *The limits of grammaticalization*. Amsterdam and Philadelphia: John Benjamins.
- Givón, T. 1984. *Syntax. A functional typological introduction*, Vol. I. Amsterdam and Philadelphia: John Benjamins.
- Givón, T. 1993. *English grammar. A function-based introduction*, Vol. II. Amsterdam and Philadelphia: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions. A construction grammar approach to argument structure*. Chicago and London: The University of Chicago Press.
- Greenberg, Joseph H. 1978. "How does a language acquire gender markers?" In: Joseph H. Greenberg, Charles A. Ferguson and Edith A. Moravcsik, (eds.), *Universals of human language*, Vol. 3: *Word structure*, 47-82. Stanford: Stanford University Press.
- Greenberg, Joseph H. 1991. "The last stages of grammatical elements: contractive and expansive desemanticization". In: Elizabeth Closs Traugott and Bernd Heine, (eds.), 301-314.
- Haiman, John and Sandra A. Thompson, (eds.) 1988. *Clause combining in grammar and discourse*. Amsterdam and Philadelphia: John Benjamins.
- Halliday, M.A.K. 1994 [1985]. *An introduction to functional grammar*. London: Edward Arnold.
- Haspelmath, Martin. 1998. "Does grammaticalization need reanalysis?" *Studies in Language* 22:315-351.
- Haspelmath, Martin. 2002. "On directionality in language change, with particular reference to unidirectionality in grammaticalization". Paper presented at the International Conference "New Reflections on Grammaticalization 2", Amsterdam, 4-6 April 2002.
- Harris, Alice C. 2003. "Cross-linguistic perspectives on syntactic change". In: Brian D. Joseph and Richard D. Janda, (eds.), 529-551.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Heine, Bernd. 2003. "Grammaticalization". In: Brian D. Joseph and Richard D. Janda, (eds.), 575-601.

- Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: The University of Chicago Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hernanz, M. Lluïsa. 1994. "Argumentos implícitos, operadores nulos e interpretación arbitraria: el caso de los infinitivos pseudoecuativos". In: Violeta Demonte, (ed.), *Gramática del español*, 315-362. México, D.F.: El Colegio de México.
- Hernanz, M. Lluïsa. 1999. "El infinitivo". In: Ignacio Bosque and Violeta Demonte, (eds.), *Gramática descriptiva de la lengua española*, 2197-2356. Madrid: Espasa Calpe para la Real Academia Española.
- Hogan, Michael J. and Lisa A. Hogan. 1998. "'Impertinent' by sentences in American writing: a cognitive motivation". *Word* 49:181-203.
- Hopper, Paul J. and Elizabeth Closs Traugott. 2003 [1993]. *Grammaticalization*. Cambridge: Cambridge University Press.
- Houston, Anne. 1989. "The English gerund: syntactic change and discourse function". In: Ralph W. Fasold and Deborah Schrifin, (eds.), *Language change and variation*, 173-196. Amsterdam and Philadelphia: John Benjamins.
- Huddleston, Rodney. 1971. *The sentence in written English. A syntactic study based on an analysis of scientific texts*. Cambridge: Cambridge University Press.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hudson, Richard. 1995. "Competence without Comp?". In: Bas Aarts and Charles F. Meyer, (eds.), 40-53.
- Jack, George B. 1988. "The origins of the English gerund". *NOWELE* 12:15-75.
- Jespersen, Otto. 1909-1949. *A Modern English grammar on historical principles*. 7 vols. Copenhagen: Ejnar Munksgaard. [Reprinted, London: George Allen & Unwin, 1961, 1965, 1970.]
- Jespersen, Otto. 1958 [1924]. *The philosophy of grammar*. London: George Allen & Unwin.
- Joseph, Brian D. and Richard D. Janda, (eds.) 2003. *The handbook of historical linguistics*. Malden (MA), Oxford, Melbourne and Berlin: Blackwell Publishing.

- Kisbye, Torben. 1971-1972. *An historical outline of English syntax. Parts I and II*. Aarhus: Akademisk Boghandel.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London and New York: Routledge.
- Koster, Jan. 1978. "Why subject sentences don't exist". In: Samuel J. Keyser, (ed.), *Recent transformational studies in European languages*, 53-64. Cambridge, MA: The MIT Press.
- Kytö, Merja. 1996 [1991]. *Manual to the diachronic part of the Helsinki Corpus of English Texts: coding conventions and lists of source texts*. 3rd ed. Helsinki: Department of English, University of Helsinki.
- Lass, Roger. 1990. "How to do things with junk: exaptation in language evolution". *Journal of Linguistics* 26:79-102.
- Lass, Roger. 1992. "Phonology and morphology". In: Norman Blake, (ed.), 23-155.
- Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lehmann, Christian. 1985. "Grammaticalization: synchronic variation and diachronic change". *Lingua e Stile* XX:303-318.
- Lehmann, Christian. 1988. "Towards a typology of clause linkage". In: John Haiman and Sandra A. Thompson, (eds.), 181-225.
- Lehmann, Christian. 1995 [1982]. *Thoughts on grammaticalization*. München/Newcastle: LINCOM Europa.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Mair, Christian. 1990. *Infinitival complement clauses in English. A study of syntax in discourse*. Cambridge: Cambridge University Press.
- Mair, Christian. 2002. "Three changing patterns of verb complementation in Late Modern English: a real-time study based on matching text corpora". *English Language and Linguistics* 6:105-131.
- Mandelblit, Nili. 2000. "The grammatical marking of conceptual integration: from syntax to morphology". *Cognitive Linguistics* 11:197-251.
- McDavid, Virginia. 1964. "The alternation of 'that' and zero in noun clauses". *American Speech* 39:102-113.
- Miguel, Elena de. 1996. "Nominal infinitives in Spanish: an aspectual constraint". *Canadian Journal of Linguistics* 41:29-53.

- Miller, D. Gary. 2002. *Nonfinite structures in theory and change*. Oxford: Oxford University Press.
- Nichols, Johanna. 1988. "Nominalization and assertion in scientific Russian prose." In: John Haiman and Sandra A. Thompson, (eds.), 399-428.
- Noonan, Michael. 1985. "Complementation". In: Timothy Shopen, (ed.), *Language typology and syntactic Description*, Vol. II: *Complex constructions*, 42-140. Cambridge: Cambridge University Press.
- Norde, Muriel. 2001. "Deflexion as a counterdirectional factor in grammatical change". In: Lyle Campbell, (ed.), *Grammaticalization: a critical assessment*, 231-264. Special issue of *Language Sciences* 23.
- Norde, Muriel. 2002. "The final stages of grammaticalization: affixhood and beyond". In: Ilse Wischer and Gabriele Diewald, (eds.), *New reflections on grammaticalization*, 45-65. Amsterdam and Philadelphia: John Benjamins.
- Plank, Frans. 1995. "Entgrammatisierung – Spiegelbild der Grammatisierung?" In: Norbert Boretzky et al., (eds.), *Natürlichkeitstheorie und Sprachwandel*, 199-219. Bochum: Universitätsverlag Brockmeyer.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rohdenburg, Günter. 1995. "On the replacement of finite complement clauses by infinitives in English". *English Studies* 76:367-388.
- Rohdenburg, Günter. 1996. "Cognitive complexity and increased grammatical explicitness in English". *Cognitive Linguistics* 7:149-182.
- Rohdenburg, Günter. 1999. "Clausal complementation and cognitive complexity in English". In: Fritz-Wilhelm Neumann and Sabine Schülting, (eds.), *Anglistentag 1998 Erfurt Proceedings*, 101-111. Trier: Wissenschaftlicher Verlag.
- Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: The MIT Press.
- Sundby, Bertil, Anne Kari Bjørge and Kari E. Haugland. 1991. *A dictionary of English normative grammar 1700-1800*. Amsterdam and Philadelphia: John Benjamins.
- Tabor, Whitney, and Elizabeth Closs Traugott. 1998. "Structural scope expansion and grammaticalization". In: Anna Giacalone Ramat and Paul J. Hopper, (eds.), 229-272.
- Tajima, Matsuji. 1985. *The syntactic development of the gerund in Middle English*. Tokyo: Nan'un-do.

- Thompson, Sandra A. 2002. "Object complements' and conversation: towards a realistic account". *Studies in Language* 26:125-164.
- Timberlake, Alan. 1977. "Reanalysis and actualization in syntactic change". In: Charles N. Li, (ed.), *Mechanisms of syntactic change*, 141-177. Austin and London: University of Texas Press.
- Traugott, Elizabeth Closs. 2003. "Constructions in grammaticalization". In: Brian D. Joseph and Richard D. Janda, (eds.), 624-647.
- Traugott, Elizabeth Closs and Bernd Heine, (eds.) 1991. *Approaches to grammaticalization*, Vol. I. Amsterdam and Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs and Ekkehard König. 1991. "The semantics-pragmatics of grammaticalization revisited". In: Elizabeth Closs Traugott and Bernd Heine, (eds.), 189-218.
- Van Valin, Robert D. and Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
- Visser, Frederikus Theodorus. 1963-1973. *An Historical syntax of the English language*. 3 parts in 4 vols. Leiden: E.J. Brill.

¹ I would like to thank Bernard Comrie and two anonymous reviewers for valuable comments and suggestions. I am also indebted to the audience at the presentation of an earlier version of this paper at the 8th International Cognitive Linguistics Conference (Logroño, 20-25 July 2003). Last but not least, I am grateful to the Autonomous Government of Galicia (grant no. PGIDT01PXI20404PR) and the Spanish Ministry of Science and Technology (grant no. BFF2001-2914) for generous financial support, and to Sonia Cristofaro for letting me have access to a pre-publication copy of her monograph on subordination.

² Since this research is concerned only with gerundive subject clauses and these have enough distinctive subject properties to make their analysis as subject unproblematic I will continue to refer to all complement clauses with the familiar labels "subject" and "object". However, as noted by Huddleston/Pullum (2002:958, 1017ff, 1206ff) and others, the term "object" is inadequate for many of the clauses functioning as internal complement of the verb. In addition, the assumption itself that sequences such as *I think they're reightable* illustrate the use of a finite complement clause serving as the argument of a higher predicate has been challenged by Thompson (2002; see also Diessel/Tomasello 2001), who argues that actual conversational English provides no evidence that the clauses in question are in any sense subordinate. Instead, what the data show is the frequent use of "a schema consisting of an epistemic/evidential/evaluative phrasal fragment and a clause" (p.155), with the fragment expressing speaker stance toward the content of the clause.

³ Cf. Hudson (1995:42): “this class has never been designed to include all and only the words which link a subordinate clause to the verb of a higher clause. In fact, it looks suspiciously like a dustbin for the words which can introduce a complement clause but which do not seem to belong to any other word-class such as ‘pronoun’ or ‘preposition’, rather than a word-class which plays an essential role in the grammar.”

⁴ My terminology for clause functions is taken from Huddleston/Pullum (2002:216ff), who distinguish between *external* (i.e. the subject) and *internal* complements. Internal complements are further subdivided into *core* and *non-core* (or *oblique*) complements. For related terms used in other frameworks, see, among others, Foley/Van Valin (1984:Chapter 3) and Van Valin/LaPolla (1997:25ff).

⁵ Here as elsewhere (Fanego 1996a-b, 1998, forthcoming a-b) I apply the label “gerund” to any *-ing* form having the same distribution as nouns or noun phrases. In the relevant literature, a distinction is sometimes made between the term *verbal noun*, which is used for forms in *-ing* that show only the syntactic characteristics of a noun, as in (15)-(17) below, and the term *gerund*, which is reserved for those forms in *-ing* displaying verbal features. As will become apparent later in this paper, such distinction is of little use when tracing the historical development of the gerund, in view of the categorially hybrid nature of most gerundial constructions throughout the history of English.

⁶ Though there is no complete agreement among scholars as to when exactly one period in the history of English ends and another begins, I will here adopt the more or less traditional divisions of Old English (OE; up to about 1100), Middle English (ME; 1100-1500), Early Modern English (EModE; 1500-1700), Late Modern English (LModE; 1700-1900), and Present-day English (PDE), with a further subdivision of Middle English into Early Middle English (EME; 1100-1300) and Late Middle English (LME; 1300-1500).

⁷ The *-ing* suffix of the action noun was distinct from the ending of the present participle (*-ende*). In the course of Middle English the variant *-ing(e)* became available also for the participle, first in the south of England and then in the rest of the country (Lass 1992:146). However, as shown by Jack (1988) and others on the basis of chronology and dialect evidence, the verbalization of the *-ing* nominal which is described in the paragraphs below took place quite independently of its morphological merger with the present participle, at least during the initial stages of the change.

⁸ My own data on the gerund (Fanego 1998:90-92, forthcoming a) confirms that direct objects were acquired much earlier than subjects in non-genitive form. As regards perfect and passive gerunds, further research is needed in order to determine whether they developed later than direct objects, too, as is suggested by Tajima (1985:135ff), or whether their non-occurrence until the fifteenth/sixteenth centuries is simply a reflection of their low frequency at all stages of English, which makes them much less likely to appear in texts, as aptly pointed out by Tabor/Traugott (1998).

⁹ This was probably a result of the resistance of the Middle English infinitive to be used after prepositions other than *to* and *for to*. As I have discussed elsewhere (1996b, forthcoming a; see also Jack 1988, Miller 2002:315-350), it seems likely that the great expansion in the use of prepositions in the course of the Middle English period, as a consequence of the decay of the Old English inflectional system, gave rise to

a situation in which a form of the verb capable of being used prepositionally was often called for; the gerund may have served to fill this gap.

¹⁰ An anonymous reviewer observes that he/she finds implausible the reanalysis-plus-actualization model (Harris/Campbell 1995) which I propose for the development of the verbal gerund, since it makes it difficult to explain why “it took so long for the verbal properties to be actualized”. In accordance with Tajima’s data (1985), he/she argues instead that the change was gradual (a view with which I fully agree; see below) and was moreover governed by universals of grammar, as follows: in theories of parts of speech such as that put forward in Croft (1991:36-98, 2001:63-107; see also Cristofaro 2003:267-282), nouns and verbs are the typologically unmarked prototypes for reference to an object and for predication of an action. Nominalizations, by contrast, require “overt structural coding” (Croft 2001:88ff) in order to be used as referring expressions, that is, they require suffixes such as *-ion* in “the *destruction* of the city” or *-ance* in “the *acceptance* of the offer”, and are thus intermediate between unmarked nouns and unmarked verbs. It is not surprising, therefore, that the ME action noun in *-ing* should pick up verb-like properties.

Though this account is substantially correct, it is not generally the case that lexical nominalizations across languages develop into members of the verb system —as happened with the English gerund, or, in prehistoric times, with the Indo-European infinitive, originally also a noun of action (cf. Disterheft 1980). In other words, we need to identify, for any individual change, the mechanism or mechanisms that get the process started. In the case of ME *-ing* nominals (cf. Fanego forthcoming a), it is clear that they exhibited a range of properties that made them unique candidates for a shift towards the verbal end of the nomino-verbal scale, namely: *a*) being nominalizations, they preserved the argument structure of the corresponding base verb; *b*) they had (almost) unlimited productivity; *c*) they were fully transparent morphosemantically in that they lacked any kind of morphophonemic alternation with respect to their bases and their meaning was quite consistently an “act/process of doing x”, two necessary conditions for them to develop from a derivational into an inflectional type; and *d*) they were exposed to other subsidiary factors, such as influence from the French gerund (see also footnote 9 above). Features (a)-(d) no doubt provided the necessary preconditions for the change to take place, yet the fact that the transition from noun to verb started with Type I gerunds (*writing of it*), rather than with those of Type II (*the/his writing of it*), and was for a long time exclusively restricted to them suggests that the specific trigger for verbalization was the availability in Middle English of sequences such as (21)-(24) above, in which the dependents of the *-ing* form were categorially ambiguous, and there was no determiner marking the gerund phrase as overtly nominal. In my view, the existence of such surface patterns, which had “the potential for more than one structural analysis” (Harris/Campbell 1995:72), made it possible to reinterpret a nominal structure as a (more) verbal one. This is not to say, however, that *-ing* nominals switched abruptly from N to V, with concomitant reanalysis of NP as VP, for in that case we would not expect to find categorial hybrids such as *by recevyng worthily of the precious body of Jhesu Crist* (c1390 Chaucer *CT Parson’s Tale* 385) or *touching teaching of childer their letters* “as regards teaching children their letters” (1569 Hart *Orthographie* 71). As proved by these and similar examples, the gerund acquired

its verbal traits gradually, a course of development which is compatible with Harris/Campbell's (1995) model of syntactic change.

¹¹ The following abbreviations will be used henceforth to indicate the source of each example in the quotations from Early and Late Modern English: HC (= Helsinki Corpus, 1500-1710; a sample of 392,110 words); COPC (= Century of Prose Corpus, 1680-1780; 240,000 words); ECF (= Chadwyck-Healey's Eighteenth Century Fiction, 1705-1797; a sample of 360,000 words); NCF (Chadwyck-Healey's Nineteenth Century Fiction, 1849-1872; a sample of 200,000 words). For further details on these corpora and text samples, see Fanego (1998, forthcoming b).

¹² Note that in examples (35)-(36) the appositive markers *viz* and *to wit* also help to mark the boundaries of the *-ing* clause. Compare also ECF 1722 Defoe *Moll Flanders* 15: "he began with that unhappy Snare to all Women, *Viz. taking Notice upon all Occasions how pretty I was.*"

¹³ The restrictions on the occurrence of *-ing* clauses applied also after purposive *for* (e.g. 1725 Haywood *Fantomina* 280: "in search of a House for the compleating her Project"), *to* (e.g. 1749 Cleland *Memoirs of a Woman of Pleasure* 16: "infinitely superior to the being touch'd with any glare of dress, or ornaments"), *but* "except", and comparative *as* and *than*. For various reasons discussed in detail in Fanego (forthcoming b), for the purposes of governing verbal gerunds these five items behaved as conjunctions rather than as prepositions, hence hybrid gerunds with *the* (24 ex. in all) were common after them. In addition, my data for the eighteenth and nineteenth centuries contains 12 more examples of *the*-hybrids occurring after prepositions other than those just mentioned; it seems likely that a corpus larger than the one I have used would reveal further conditioning factors to account for some of these instances.

¹⁴ I.e., without being formally dependent on a matrix predicate, as in NCF 1855 Mrs. Gaskell *North and South* 89: "But the papers. What taste! And *the overloading such a house with colour and such heavy cornices!*"

¹⁵ Table 1 is based on a subsample of only 440,000 words; in the Late Modern English corpus as a whole (800,000 words) the number of unintroduced subject gerunds is 10.

¹⁶ For one of the exceptions found in my data, cf. *perfect playing upon velvet* in quotation (39) above. In Present-day English, *-ing* nominals lacking modification are still frequently not marked for definiteness, as in "Fred has given up *smoking*", "*anchoring* is almost impossible in deep water".

¹⁷ On *extension* as a mechanism of syntactic change see Harris/Campbell (1995:97ff) and Croft (2000:148ff); on the extension of verbalization across different classes of gerunds see Fanego (forthcoming a). A process of extension or analogical generalization in the spread of the verbal gerund needs to be assumed in view of the existence of complex mixed structures such as (25) above, where the gerunds are premodified by adjectives but govern direct objects (for details cf. Fanego 1996b, 1998, forthcoming a-b). This kind of structure is found chiefly in the seventeenth century, when the nominal system of gerundial complementation was breaking down and was giving way to the new verbal system; by the eighteenth century most hybrid gerunds were completely verbal except for the presence of an initial determiner, whether *the* or a possessive. It is roughly at this stage, I believe, when *the* could be "exapted" and used in a new function, as will be discussed below.

¹⁸ As noted by Huddleston/Pullum (2002:1256) with regard to Present-day English, the potential ambiguity of *-ing* clauses in certain environments is sometimes exploited in advertising slogans such as *Our business is working for you*: “as a specifying construction this identifies what our business is; with *be* a marker of progressive aspect, it says what our business is doing”.

¹⁹ This is in fact Martin Haspelmath’s (2002) current definition of grammaticalization: “A grammaticalization is a diachronic change by which the parts of a schematic construction come to have stronger internal dependencies.”

²⁰ Another piece of evidence against a grammaticalization analysis of subordinating *the* could perhaps be found in the fact that the shift from article to complementizer does not seem to be attested cross-linguistically as a grammaticalization process; see Heine/Kuteva (2002). In this respect articles differ from demonstratives, which, as happened with English *that* or German *das* “that” (ibid.:106-107), can develop into complementizers in constructions where they are used with deictic force to refer cataphorically to a subsequent paratactic clause (e.g. *she said that: there is no money*).

²¹ Exaptive processes have also been discussed by Greenberg (1991), who refers to them as *regrammaticalization*; see also Croft (2000:126ff).

²² See also Lyons (1999:339-340) on the life cycle of definite articles: “Definiteness itself can expand its range of application, taking in generics, specifics etc., and a point can come at which its exponent is reanalysed as grammatically and semantically empty (perhaps leading to its being pressed into service with some other function).”

²³ Norde discusses exaptation within the larger context of *degrammaticalization*, which she defines (2002:47) as a shift from right to left on Hopper/Traugott’s cline of grammaticality (cf. (52) above), for instance the shift from affix to clitic, or the shift from an inflectional to a derivational affix (as in the case of Swedish *-on*). Exaptive changes, however, need not involve degrammaticalization, but simply, as already noted, “reinterpretation in a new function” (Greenberg 1991:301); see also Hopper/Traugott (2003[1993]:174).

²⁴ The connection between exaptive processes and system restructuring is also clear from Lass’s (1990, 1997) discussion.

²⁵ Hogan/Hogan (p.191) point out that the users of impertinent *by* are “good students” who wish to show their familiarity with information they consider to be of special relevance. This suggests that the construction may arise through hypercorrection, which, according to Disterheft (1990), is particularly often found in writing.

²⁶ Since Hogan/Hogan (p.183) speak of “final exams”, I am assuming that the pages in question are handwritten, rather than typed.

²⁷ Mair (1990:84-92) is another study containing information on gerunds in contemporary usage. His figure of 126 *-ing* subjects (80 in pre-verbal position) occurring in a large corpus of British English (895,000 words) is, however, unreliable, since he has included in the count (cf. p.245) unmodified *-ing* nominals such as “*anchoring* is impossible in deep water” or “*banking* is free”; see in this respect footnote 16 above. Even including these nominals, gerunds in pre-verbal position (about nine in every 100,000 words) remain infrequent in Mair’s data. With respect to the low figure for *-ing* subjects

recorded by Huddleston, it is unlikely that it could be attributed to the nature of his corpus. The fourteen bare *-ing* clauses occurring as pre-verbal subjects (see (29)-(30) above for two of them) in my Early and Late Modern English material (1500-1872; 1,117,621 words) are mostly found in relatively formal and technical registers, so it would be unexpected for *-ing* subjects to have subsequently become a feature of colloquial speech avoided in formal styles.

²⁸ The actual number is 23, but two of their examples are “grammatical” sentences which clearly do not belong in the type under discussion. Cf. (a) for one of them and also my discussion of the semantics of *by*-sentences in 5.3 below:

(a) By judging a person by his or her color, one is immediately put in a close-minded category.

²⁹ Note that the declarative version of (38) would be a sentence with two clausal complements (i.e. “the cutting down trees, and burning of bricks means that my Lord is going to build a new apartment”), as in the examples quoted further down.

³⁰ The term *pseudo-specifying* is adapted from Hernanz’s (1994) “pseudoeuativo”, which she applies to Spanish infinitival structures such as *querer es poder* “you can do it if you really try”. Most analyses of pseudo-specifying sentences appear to assume that the two complement clauses involved must be of the same structural type (i.e. two infinitives or two *that*-clauses), but at least in the case of English this is not correct, possibly because *-ing* clauses have encroached upon both *that*-clauses and *to*-infinitives and hence are often in variation with them.

³¹ I restricted my search to the string *by *ing*. This retrieves all *by Xing* clauses without an overt subject (e.g. *by banning Welsh*), but not those like *by the company addressing the problem*. The great majority of *-ing* clauses are, however, of the former type.

³² This desemanticization of *by* can be observed in other contexts too. Witness the following quotation, where the function of the second *by* (underlined) is to help clarify a long and complex construction, much as the anacoluthic *of* occurring in (32) and (36) above:

FROWN [Science] J38 12: Underlying India’s dismal productivity growth performance are structural and intrasectoral distortions. To some degree, these could be reduced *by appropriate rural policies (notably, providing insurance incentives to encourage peasant farmers to switch to higher yield but riskier crop mixes: see Singh, 1979) and industry-oriented packages (in particular, by removing location and input pricing distortions created by current policies: see Sekhar, 1983)*. In the presence of such distortions, traditional policies that are oriented toward capital formation will have only a modest impact, mainly because returns are low.

³³ Prepositional phrases can be detached from the main clause when they function as topic introducers (e.g. *as for Sam, I haven’t seen him in two weeks*), but this is clearly a different construction (cf. Huddleston/Pullum 2002:1409ff). Though left-dislocation is most typically a property of NPs, complement clauses have also occurred as left-dislocated constituents at all stages of English; cf. Fanego (1992:59-60, 77-78) and Mair (1990:56), who quotes, from spoken British English, “*to start off at half cock you know it’s stupid.*”