Subject Omission in English Diaries

Supervisor:
Prof. dr. Liliane Haegeman
Department: Linguistics
   English

Thesis submitted in partial fulfillment of the requirements for the degree of Master in de Taal- en Letterkunde: Engels by Varduhi Nanyan

2013
Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. dr. Liliane Haegeman, for her generous support and patience. Throughout my studies she has inspired me with useful and interesting discussions and helpful comments and has led me in the right direction of research. Without her continual guidance and assistance this thesis would not have been possible.

Many thanks to my family and friends for their continual love and support.
TABLE OF CONTENTS

Preface ................................................................................................................. 7

Chapter 1. Null subjects in English ................................................................. 11
1.1 The subject requirement in English ......................................................... 13
1.2 The pro-drop parameter ........................................................................... 15
1.3 Null subjects in English ........................................................................... 18
   1.3.1 The properties of diary null subjects .............................................. 18
      1.3.1.1 The interpretation of diary null subjects .................................. 18
      1.3.1.2 Null subjects as syntactically and semantically active empty categories? ....20
      1.3.1.3 The syntactic properties of null subjects ................................... 21
1.4 The nature of diary null subjects ................................................................. 23
   1.4.1 pro ................................................................................................. 23
   1.4.2 PRO ............................................................................................... 24
   1.4.3 A-trace ............................................................................................ 25
   1.4.4 A´-trace .......................................................................................... 25
      1.4.4.1 The phenomenon of topic drop .............................................. 25
      1.4.4.2 Diary subject ellipsis as topic drop? ....................................... 27
1.5 The pragmatic conditions on subject omission ........................................... 29
   1.5.1 Recoverability: syntax or pragmatics? ............................................ 30
   1.5.2 Economy ......................................................................................... 31

2.1 Introduction ................................................................................................. 35
2.2 The nature of the DNS revisited ............................................................... 36
   2.2.1 The typology of empty categories ................................................. 36
   2.2.2 The Empty Category Principle ...................................................... 38
   2.2.3 Root clauses ..................................................................................... 39
   2.2.4 Truncation ....................................................................................... 39
2.3 Adjunct preposing ..................................................................................... 42
   2.3.1 The problem and proposals............................................................ 42
   2.3.2 The Split CP hypothesis ................................................................. 43
   2.3.3 DNS in a split CP ............................................................................ 48
2.4 The Interpretation of subjects ................................................................. 56
2.5 DNS and subject omission in coordinate clauses .................................... 57

Chapter 3. Ihsane’s (1998) analysis: Virginia Woolf’s Diary (1940) ............... 60
3.1 Introduction ............................................................................................. 60
3.2 The data and methodology ...................................................................... 61
3.3 The realization of the subject ................................................................. 63
3.4 Grammatical properties of null subjects ................................................. 66
   3.4.1 Referential and expletive subjects .................................................... 66
   3.4.2 The person and number of null subjects ......................................... 68
   3.4.3 The verb type and verb tense ......................................................... 70
   3.4.4 Material to the left of the subject ................................................... 72
3.5 Conclusion ............................................................................................... 74

Chapter 4. The analysis of the Truman Diary (1947) .................................... 75
4.1 Introduction ............................................................................................. 75
   4.1.1 The corpus ....................................................................................... 76
4.2 The data .................................................................................................... 77
   4.2.1 The classification criteria ................................................................. 80
4.3 Overall figures ........................................................................................ 86
   4.3.1 The first results: comparison with Ihsane’s (1998) findings ............... 91
4.4 The classification and grammatical properties of null subjects ............... 92
   4.4.1 Referential and expletive subjects .................................................... 92
      4.4.1.1 Comparing the results with Ihsane: referential and expletive subjects 95
   4.4.2 The interpretation of null subjects ................................................... 96
      4.4.2.1 The categories of person and number in root clauses ............... 98
      4.4.2.2 The categories of person and number: comparison with Ihsane’s findings ................................................................. 102
      4.4.2.3 The categories of person and number in clauses coordinated with roots 102
      4.4.2.4 The categories of person and number in coordinate embedded clauses ... 105
      4.4.2.5 Concluding remarks ................................................................. 107
   4.4.3 The type of verbs ............................................................................. 108
      4.4.3.1 The verb type: root clauses .................................................... 109
         4.4.3.1.1 Comparison with Ihsane’s findings .................................... 111
4.4.3.2 The verb type: coordinate root clauses .............................. 111
4.4.3.3 The verb types: coordinate embedded clauses ..................... 113
4.4.3.4 Concluding remarks ......................................................... 114
4.5 Preposing ........................................................................... 114
  4.5.1 Preposed constituents: roots ............................................. 115
    4.5.1.1 Comparing the results with Ihsane ................................. 117
  4.5.2 Preposed constituents: coordinate root clauses ..................... 118
  4.5.3 Preposed constituents: coordinate embedded clauses ............... 120
  4.5.4 Concluding remarks .......................................................... 122
4.6 Coreferentiality of subjects in Standard English and in diaries ............ 122
4.7 Discourse factors of DNS ......................................................... 126
  4.7.1 The position of the subject in the diary entry ......................... 126
4.8 Concluding remarks ............................................................... 131

Chapter 5. Conclusion .................................................................. 134

Appendix .................................................................................... 136
References .................................................................................. 143
Languages differ parametrically in terms of the “realisation of pronominal subjects in finite clauses” (Haegaman 2002b: 135, see also Haegeman 1997, Rizzi 1997). While subject ellipsis in English finite clauses gives rise to ungrammaticality (Huddleston and Pullum 2002: 238, Haegeman and Guéron 1999: 127), the pronominal null subjects are licit in languages like Italian and Spanish. So, (1a) is grammatical while (1b) is not, the Italian examples in (1c) and (1d) are both grammatical.

(1)  
a. I study English.
b. *(I) study English.
c. Io studio inglese.
   I study English.
   I study English.
d. Studio inglese.
   study-PRES-1SG English.

(1) Study English.

(1d) is licit because in Italian verbs have inflectional endings for all person and number categories and the implicit subject can be recovered (Rizzi 1994, Haegeman 1997, 2002b). The inflectional ending –o denoting the first person singular allows for the pronominal subject I to be implicit. In Standard English the subject of a finite clause cannot be deleted even if it can be recovered from the discourse context (Haegeman 2011). Nevertheless, there are specific registers (e.g. diaries, personal letters, e-mails, chat, notes) in English which do allow for a subject to be non-overt (see also Haegeman and Guéron 1999: 614-16). The phenomenon of subject ellipsis in one of these registers, namely diaries, will be studied in this thesis. (2) is an example from a diary.¹

(2) Ø Spent the day at work. (Truman’s Diary, 1947, 1 Jan.) ²

During the recent years the phenomenon of the diary subject drop has been studied by a number of researchers. Specifically, Haegeman (1990, 1997, 2002b, 2007, 2011) discusses

¹ I use the symbol Ø to denote null subjects.
² The diary is available at http://www.trumanlibrary.org/diary/transcript.htm.

This thesis studies null subject patterns in finite clauses in English non-fiction diaries with the aim to identify the syntactic features of the pronominal subject drop in diary registers through examination of empirical data. I build the thesis on Haegeman’s (1990, 1997, 2002b, 2007, 2011) analysis and proposals on diary null subjects and test the compatibility of the empirical data with them. Accordingly, I examine whether the diary drop is a root phenomenon, i.e. subject ellipsis is not available in embedded clauses, whether there are any null subjects in wh-questions, yes/no questions, and whether there is an adjunct/argument asymmetry with respect to the material preceding null subjects.

The thesis also observes whether the subject omission in diary coordinate clauses differs from the same phenomenon in core grammar. And if it does, what are the patterns? Coordinate clauses are of interest for my research for two reasons: (i) subject omission in the second conjunct is licit in Standard English; (ii) null subjects in the second conjunct of coordinate clauses and DNS have similar distribution constraints (Haegeman 2011: 25-26). Haegeman (2007, 2011) briefly discusses the syntactic patterns of subject omission in coordinate clauses and in diaries; the discussion will be presented in Chapter 2.

The similarities between diary null subjects and coordinate ellipsis in Standard English was first highlighted by Wilder (1996) (see also Haegeman 2011). I investigate whether this observation is compatible with my data. If subject ellipsis in diary style is the same phenomenon as the subject omission in coordination, then this suggests that diary subject ellipsis is a grammatical omission: the diary style seems to explore and generalize an option that is already available in core grammar, namely, the subject ellipsis in coordination. I will test this hypothesis on my data.

Note that I am not concerned with providing a theoretical account for the derivation or structure of the subject omission in coordinate clauses. My primary objective is to investigate

3 Haegeman and Ihsane (1999, 2001) show that there is a difference between fictional and non-fiction diaries in terms of the syntactic constraints of null subjects. Namely, fictional diaries allow null subjects in embedded clauses while no such examples have been found in non-fiction diaries. Note, however, that my interest lies in examining non-fiction diaries.
the null subject patterns attested in diary coordinate clauses and to identify their similarities/differences with the patterns which occur in Standard English.

For my analysis I have chosen a diary which has an American author, namely, *Harry S. Truman 1947 Diary*. The abovementioned diaries studied in some detail by Ihsane (1998) and Becquet (2000) were both written by British writers, there are no comparable studies of American diaries yet. Although at first sight the distinction between the US and UK diaries might seem irrelevant, I examine whether diary registers in the two variants of English differ. Specifically, I explore whether the syntactic constraints typical of null subjects attested in British English non-fiction diaries extend to American English non-fiction diaries. It is also observed whether British English non-fiction diaries differ from American English non-fiction diaries with respect to the rate of subject omission.

Given my aim to study non-fiction diaries, for comparative analysis, among the two British diaries mentioned above I have chosen for *The Diary of Virginia Woolf* (1940) analysed in Ihsane (1998). In order to be able to compare the attested patterns first an overview of Ihsane’s (1998) analysis is presented. Then I examine the distribution of subject patterns in my own corpus data. Finally, a comparative analysis of the results of both corpuses is conducted.

To make the comparative analysis more efficient I build my research on the methodology applied by Ihsane (1998). However, in order to refine her findings, my work departs from hers in a number of ways: while she mainly looks at null subjects, I examine the patterns of both overt and null subjects in my corpus in order to be able to draw more accurate conclusions on the distribution of diary null subjects (henceforth DNS). In addition, I also study coordinate clauses. Although Ihsane (1998) considers this clause type in her general overview of the distribution of overt and null subjects, she doesn’t study it with respect to subject ellipsis.

---

4 Although there is a universally accepted variant of English, widely known as Standard English, there are still differences between its ‘regional’ varieties. According to David Crystal, “Standard English is essentially written, printed English, seen in the textbooks, newspapers, and periodicals of the world – and also, these days on the World Wide Web. It is largely identical in its global manifestation; we must allow only for the small amount of variation in vocabulary, grammar and spelling which make up the differences between Am, Br, Aus and other ‘regional’ standards” (2004: 39). Thus, if these lexical, grammatical and phonetic differences exist in the standard variety of the language, it would be logical to think that they might be attested in its specific registers as well. From here arises my interest to examine American English diaries.

5 Teddiman & Newman (2007) examine British and American blog diaries and observes that the rate of subject omission in British diaries is higher than in the American ones. I examine whether this is the case in non-fiction diary registers.

6 Becquet (2000) studies a fictional diary, namely, *Bridget Jones’s Diary*, as mentioned above.

---

7 See the reference list for all cited authors.
Chapter 1. Null subjects in English

1.1 The subject requirement in English

Given the subject requirement hypothesis (Haegeman and Guéron 1999: 127) in English all finite clauses must have a subject, and the canonical position of the subject, SpecIP, must be filled. Thus, subject is an obligatory element in a finite clause in English (3a), and subjectless sentences are considered ungrammatical (3b) (Huddleston and Pullum 2002: 238, see also Quirk et al. 1985: 724).

(3) a. Jane has arrived.
    b. *Has arrived.

However, non-overt subjects are allowed in some finite clause types: imperative (4a) and coordinate clauses (4b) (Huddleston & Pullum 2002: 238, Quirk et al. 1985: 724).

(4) a. Come here!
    b. Joana ate her dinner and (Joana/she) went for a walk.

In both cases despite being non-overt the subject can be recovered. In (4a) the empty category can be identified as the second person singular or plural depending on the situation since imperative clauses are used in contexts when there is a potential addressee. In case of (4b) the omitted subject has an antecedent in the first conjunct and can be easily recovered, in this case, as Joana or she. Note that in core grammar the subject of the second conjunct in a coordinate clause cannot be omitted if it is not co-referential with the subject of the first conjunct (Quirk et al. 1972: 555).

I will not focus on null subject patterns in imperative clauses in this thesis because irrespective of the register imperatives have an implicit subject and, hence, they cannot be of interest for my research.

Conversely, coordinate clauses are interesting for the given analysis as they seem to be a ‘link’ between core grammar and diary registers in terms of the subject ellipsis with this

---

8 Note that null subjects occur in non-finite clauses, too, but I do not discuss them in this thesis; I focus only on the realization of subjects in finite clauses.

9 See also Biber et al. (1999: 219-221) for realization and distribution of imperative clauses.
phenomenon being licit in both. As mentioned above, Haegeman (2011) observes that null subjects in diaries and coordinate null subjects in core grammar demonstrate similar syntactic patterns. I investigate whether this observation is borne out in Truman’s Diary.

Furthermore, it is investigated whether the subject omission in the second conjunct in core grammar and in diary registers is the same phenomenon. As it has been mentioned above, in Standard English the second conjunct can have a non-overt subject only if the co-referentiality of subjects applies. I examine whether in my data the null subject of the second conjunct is always identical to that of the first clause. Observe that Becquet (2000)\textsuperscript{10} shows that in \textit{Bridget Jones’s Diary} the subject of the second conjunct can be null even if it is not co-referential with the subject of the first conjunct. But this pattern has been found in a diary style which allows for embedded null subjects (i.e. in a fictional diary). I examine whether the same pattern applies to the diary style in which null subjects are unavailable in embedded environments\textsuperscript{11}.

As already noted above, in this thesis I do not elaborate a theoretical account for the derivation or structure of coordination ellipsis. My primary aim is to examine null subject patterns attested in diary coordinate clauses and to identify their similarities/differences with the patterns which are licit in core grammar. I refer the reader to Chapter 4 for discussion.

Thus, as discussed above, Standard English by and large doesn’t license non-overt subject patterns in finite clauses\textsuperscript{12}. On the other hand, in languages like Italian and Spanish, null subjects are fully acceptable in finite clauses and are often the preferred option. I will investigate this cross-linguistic variation with respect to the realization of the pronominal subject in finite clauses in the next section. Section 1.3 focuses on the null subjects in specific registers of English. Sections 1.3.1 deals with the interpretation of diary non-overt subjects. In the following two sections I present the semantic and syntactic properties of DNS. Section 1.4 discusses the nature of diary null subjects. Section 1.5 focuses on the pragmatic conditions on subject ellipsis.

\textsuperscript{10} See also Haegeman (2002b: 141-142)
\textsuperscript{11} The non-fiction diaries studied so far have shown that DNS are a root phenomenon.
\textsuperscript{12} As shown above, imperative and coordinate clauses allow non-overt subjects, though.
1.2 The pro-drop parameter

According to the Principles and Parameters approach (Chomsky and Lasnik 1993), all languages have a set of common principles, which build the model of Universal Grammar. Yet they vary from each other in terms of a set of parameters and these differences give rise to cross-linguistic variations. One of these parameters referred to as a pro-drop parameter determines the realization of a pronominal subject in a finite clause, i.e. whether the pronominal subject may be omitted or should be overtly realized in a tensed clause (Haegeman 1997: 233). In pro-drop languages, like Italian, the subject can be non-overt (5a) while in a non-pro-drop language, like English (5b)\(^\text{13}\), the subject cannot be omitted in a finite clause (Haegeman 1997: 233, Rizzi 1997: 270). Consequently, (5a) is grammatical while (5b) is not.

(5) a. Italian  (Io) Compro un giornale.
         (I) buy        a   newspaper

b. English  *(I) buy a newspaper.

What licenses this difference among languages? According to Rizzi (1997: 270), in the late seventies it was attempted to define a number of properties which might account for this cross-linguistic variation. I will outline these properties here without discussing them and refer the reader to Rizzi (1994, 1997) for analysis.

1. In pro-drop languages both referential (as shown in (5a) and non-referential subjects\(^\text{14}\) (6a, c) can be non-overt in a tensed clause while non-pro-drop languages don’t allow this (see (5b) for referential subjects). (6b, d) provide examples of non-referential subjects. Rizzi (1997: 271) states that Italian not only licenses non-overt expletive subjects but also “disallow[s] an overt non-referential pronoun”, as shown in (6a). (6a) and (6b) illustrate examples of a quasi-argument subject of a weather verb in Italian and English respectively. (6c) and (6d) are examples of expletives\(^\text{15}\).

\(^{13}\text{All examples in this section are taken from Haegeman (1997) and are presented as given in the source.}\)
\(^{14}\text{i.e. expletives and quasi-argument subjects of the clauses with weather verbs}\)
\(^{15}\text{In my analysis I use the term expletives both for the weather-verb subjects, like in It is snowing, and the dummy subjects, like in It is strange that he is late.}\)
(6) a. (*Ciò) piove.
   rains (3sg.)
   ‘It is raining.’
b. *(It) is raining.
c. (*Chiò) è chiaro che Lara non partirà.
   is clear that Lara not will leave.
   ‘It is clear that Lara will not leave.’
d. *(It) is clear that Lara will not leave.

2. In pro-drop languages the subject can occur in a post-verbal position (7a). This pattern is not compatible with non-pro-drop languages (7b).

(7) a. Ha telefonato il decano.
   has telephoned the dean.
   ‘The dean has phoned.’
b. *Has telephoned the dean.

3. In languages like Italian (8a) “an embedded subject can be moved to the matrix domain across the overt complementizer che, without giving rise to a that-trace violation” (Haegeman 1997: 235). This is impossible in English: "the subject trace ti in the embedded domain is not governed by an appropriate head governor” (Haegeman 1997:235): the complementizer that is “inert for government” (8b).

(8) a. Chi credi che abbia telefonato ti?
   Who believe you that have (subj.) telephoned?
   ‘Who do you think has called?’
b. *Who, do you think that ti has telephoned?

4. Finally, in Italian “the morphological paradigm of subject-verb agreement is rich and transparent if compared to English” (Rizzi 1997: 272) and each person-number combination has a distinct inflectional ending (9a). In English the inflectional paradigm is poorer (9b). It has a marked form only for the third person singular in present simple, in all the other cases the verb has identical forms for all person-number combinations (Rizzi 1997: 272).
Note that not all pro-drop languages do necessarily share all the abovementioned properties. Specifically, as Cole (2010) reports (see also Huang 1984), not all pro-drop languages, which allow a thematic null pronominal subject, do necessarily have a rich inflectional paradigm of verbs for the categories of person and number. To explain why such a discrepancy occurs Jaeggli and Safir (1989) offer the Principle of Uniformity (see also Harvie 1989), according to which a language must have “either only underived inflectional forms or only derived inflectional forms” (Jaeggli and Safir, 1989: 30). English stands in-between the two; it has a morphological ending only for the third person singular, which means that English does not belong to either of the groups and cannot be considered a pro-drop language. Chinese and Japanese, on the other hand, have no morphological endings at all and, hence, qualify as a pro-drop language. Cole (2010) suggests that languages of the latter type allow null subjects due to their contextual strength and concludes that accessibility is a determining factor for the occurrence of null subjects. I refer the reader to Cole (2010) for relevant analysis.

1.3 Null subjects in English

As shown above, the pro-drop parameter is set negatively in English, and null pronominal subjects are considered ungrammatical\(^\text{16}\) in core grammar. Yet subject ellipsis is attested in English abbreviated registers, like diaries (Haegeman 1997, 2011, Haegeman and Ihsane 1999, see also Weir 2008) personal letters, informal messages, e-mails, chat, short notes and in spoken English (Napoli 1982, Weir 2008). According to Haegeman (1997: 238, see also

\(^{16}\) As illustrated in Sections 1.1, non-overt subjects are licit in some clause types in Standard English, though.
Haegeman & Guéron 1999: 615), these registers allow for both referential and expletive subjects to be non-overt. (10) – (14) give examples without discussion.

(10) *diary*: ø Spent the day at work. (Truman’s Diary, 1947, 1 Jan.)

(11) *personal e-mail*: ø Still have horrible cold. (from Scott 2010:12)

(12) *personal SMS/text message*: ø Must have missed one another. (from Scott 2010:12)

(13) *informal note*: ø Wish you had come. (from Ihsane 1998:10)

(14) *spoken English*:


  b. ø Isn’t much we can do about it. (Thrasher 1977:44, reported in Haegeman and Guéron 1999:616)

Note that null subject patterns also occur in the early production (Rizzi 1994, 2006, Hyams 1986, Bromberg and Wexler 1995). (15) provides an example of a child-language null subject. (15a) is from Ihsane (1998:11), (15b) is from Haegeman (2002b:135).

(15) a. ø Want more.

  b. ø Was a green one. (Eve 1; 10)

I don’t discuss child production null subjects in this thesis and refer to the abovementioned authors for analysis. As noted above, Haegeman adopts Rizzi’s (1994, 1997, 2006) account for subject omission in early production for her analysis of DNS and suggests that subject ellipsis in adult written registers “is analogous to the early subject omission” (Haegeman 2002b: 136). Given that diary null subjects and subject omission in child production have similar manifestations she observes that both “follow from similar, or even identical, mechanisms” (Haegeman 2011:3). Following Thrasher’s (1977) analysis of null subjects in spoken English, Haegeman (2002b) shows that although at first sight one might think that early production is closer to spoken language given that both are oral models, it is not so.

---

17 As already mentioned above, my analysis is based on Haegeman’s (1990, 1997, 2002b, 2007, 2011) proposals which will be presented in Chapter 2.
Haegeman (2002b: 137) provides several arguments; I will present one of them here and refer the reader to Haegeman (2002b: 136-138) for sufficient discussion. Specifically, in spoken English, the article can be omitted if it occurs in a sentence-initial position while “a sentence-internal omission leads to ungrammaticality” (Haegeman 2002b: 137). (16) gives an example from Thrasher (1977: 35), reported in Haegeman (2002b: 137).

(16) a. Damn dogs are taking over the city.
    b. *Damn dogs are taking over city.

Haegeman (2002b) observes that in abbreviated written registers such as diaries, articles can also be deleted sentence-internally. (17) is from Haegeman (2002b: 137).

(17) Had to stop, wet to – skin… (V. Woolf Diary, vol 5: 89)

The same holds for the child language where article omission is not restricted to the sentence-initial position (18) (from Haegeman 2002b: 137).

(18) Paula play with ball. (Paula 1; 6)

Haegeman (2002b) concludes that “the adult written register is a better candidate for comparing with the early production than the oral register” (2002b: 138) and suggests that the ellipsis patterns which are relatively freely available in child language become “restricted to the adult written register in the course of language development” (2002b: 138).

In his account for informal style subject omission, Weir (2008) argues that non-overt subject patterns in written registers and in spoken language are different phenomena. According to him, “subject pronoun drop in spoken English is best analysed, not in fact as “subject pronoun drop” as such, but one instantiation of a metrical phenomenon “it is permitted to delete initial weak syllables” (2008: 24). The same conclusion was drawn by Napoli (1982). I will not discuss the differences of subject ellipsis in spoken and written English here and refer to Thrasher (1974), Napoli (1982), Weir (2008) for sufficient discussion.

In what follows I will confine the analysis to the null subject patterns which occur in diaries. To understand the nature of the diary non-overt subjects let us discuss the properties of DNS.
1.3.1 The properties of diary null subjects

1.3.1.1 The interpretation of diary null subjects

Given that null subjects can “be spelt out by an overt pronoun” (Haegeman 2007: 95) diary null subjects are commonly understood as pronouns: “in diary writing we find examples with first and third person interpretations, both for singular and for plural” (Haegeman 2007: 95, see also Weir 2008). Given this observation it seems plausible to treat DNS on a par with a pronoun. Observe also that lexical NPs are strongly informative and cannot easily be dropped because there is no discourse referent from which the deleted information can be retrieved. Pronouns, however, can be omitted as they are less informative, i.e. they contain ‘given’ information, and can be easily recovered from a discourse antecedent. Consider (19).

(19) John likes reading. He has read too many books already.

The lexical DP John establishes an entity in the discourse and the subsequent reference to that entity would be with a pronoun, he and it is the pronoun that can be omitted. Following this logic and taking into account Haegeman’s observation above, I treat diary null subjects as omitted pronouns.

Examination of empirical data shows that in diaries I-omissions are more frequently attested as compared with the other person/number categories. In her analysis of Virginia Woolf’s Diary (1940) Ihsane (1998) reports that 53 out of 111 attested null subjects can be interpreted as I, there are 17 cases of non-overt we, 36 instances of he/she/it and 5 cases of they. It follows that the omission of the first person singular makes nearly the half of the attested cases of implicit subjects. In view of the material studied such a distribution of non-overt subjects in terms of the grammatical categories of person and number seems logical: the diary is about the life events of the writer and, therefore, I is the default topic in the written text. Hence, I-omissions “can arise ‘out of the blue” while the other person/number categories need an antecedent to be non-overt (Haegeman, 2011: 6).

There are no instances of you-omissions in Ihsane’s corpus. However, empirical data evidence that cases of the second-person subject omissions are attested, too. (20) is taken

(20) Note that the second-person subject omissions are rare in diaries. This can be conditioned by the fact that in diaries there is no addressee. This has been observed by Haegeman and Ihsane (1999), too. No second-person null subjects have been attested in Truman’s Diary, 1947. As noted above, in her analysis of Virginia Woolf’s
from Haegeman and Ihsane (1999: 132), also reported in Scott (2010: 13). (21) gives examples of non-overt first person singular (21a) first person plural (21b), third person singular (21c) and third person plural (21d) subjects.

(20) Trouble is, Rebecca’s stings are aimed so subtly at one’s Achilles’ heels, like Gulf War missiles going ‘Fzzz Whoosssh’ through Baghdad hotel corridors, that ø never see them coming. (Fielding 1996: 146)

(21) a. ø Had breakfast at 9 A.M. (Truman’s Diary, 1947, 5 Jan.)
   b. ø Land at 10:00 on the dot. My pilot never misses a schedule. (Truman’s Diary, 1947, 4 March)
   c. ø Said she was for me. (idem, 3 Jan.)
   d. ø Sat there for 3/4 hours. (V. Woolf 1940: 334, from Haegeman 2013: 8)

Note that in (20) the omitted subject is a generic you and can be substituted for one. This means that you in this example does not imply an addressee.

A final point to be made with regard to the interpretation of the DNS is that, as Haegeman (2007: 95) observes, null subjects and their “pronominal counterparts” differ. This difference is illustrated in (22). Specifically, in (22a) “the overt pronoun she can be coreferential with the DP Mary in the sentence initial adjunct”, while in (22b) “the null subject is grammatical as such but it cannot be interpreted as coreferential with the DP Mary” (Haegeman, 2007: 95). I will return to this later in Section 2.4.

(22) a. In John’s picture of Mary, she smiles.
   b. In John’s picture of Mary, [ec n/*i] smiles.

1.3.1.2 Null subjects as syntactically and semantically active empty categories?

This section gives evidence that though phonetically non-overt, diary null subjects are syntactically and semantically active\textsuperscript{19}.

---

\textsuperscript{19}Here I limit my discussion to diary null subjects, however, Scott (2010) shows that these properties are in line with the non-overt subjects of the other registers as well.

---
Given that the null subject can be interpreted and carries “the grammatical properties of a subject which is not spelt out” (Haegeman, 2011: 6), one can assume that DNS are also actively realized at the semantic level as otherwise this would bring to the violation of theta-criterion (see also Scott 2010). This criterion is repeated in (23) as given in Haegeman and Guéron (1999: 138).

(23) a. Each argument must be associated with one and only one theta-role.
   b. Each theta-role must be associated with one and only one argument.

According to (23), there should be ‘one-to-one mapping’ between the arguments and the semantic roles assigned (see also Haegeman 2006: 193). Note that theta-roles are assigned by the verb to its arguments and that the number of the thematic roles assigned in a sentence depends on the properties of the verb (transitive/intransitive). Thus, we cannot assume that if a sentence has a transitive verb with a non-overt subject but an overt object, then the latter can be assigned two roles: both its role and that of the subject. There should be two potential arguments to be assigned the roles. Consider (24):

(24) ø Cooked dinner.

In this sentence the verb cook assigns two roles: the agent and the theme. It would be logical to expect, then, that the sentence should have two arguments. The role of the theme is assigned to dinner, but there is no potential candidate for the role of agent. According to (23), dinner cannot be assigned both roles. To solve this problem and to avoid the theta-criterion violation we can postulate an empty category in the subject position which is assigned the agent role.

Thus, I postulate that diary null subjects are empty categories in SpecIP, which though phonetically non-overt, carry out their thematic role as assigned by the verb of the clause. It can be assumed, therefore, that the logical form (LF) of a clause containing a null subject should “be identical to that of a parallel sentence containing an overt pronoun” (Scott 2010: 15). It follows that (25) should be equivalent to (26). Accordingly, in (25a) the null subject should be seen as an empty category which is assigned a thematic role of agent. (26a) gives the same sentence with a pronounced pronoun in the subject position.

---

(25) a. ø Spent a quiet pleasant day at Stanley Woodward's place. (Truman Diary, 1947, 5 July)
    b. LF: X spent a quiet pleasant day at Stanley Woodward's place.

(26) a. I spent a quiet pleasant day at Stanley Woodward's place.
    b. LF: X spent a quiet pleasant day at Stanley Woodward's place.

Now let us consider the example in (27) reported in Haegeman (2011: 6).

(27) ø Have busied myself to-day in unpacking my trunk and arranging my things for a visit to Vancouver. (1836)

The omitted subject binds the reflexive pronoun myself. This means that the empty category in the subject position is syntactically represented and is syntactically active and carries “the grammatical properties of a subject which is not spelt out” (Haegeman, 2011: 6).

Thus, it can be concluded that **DNS are syntactically and semantically active empty categories.**

Given the syntactic sensitivity of diary null subjects my next move is to present the syntactic features identified in relation to the diary subject drop.

### 1.3.1.3 The syntactic properties of null subjects

#### a) root phenomenon

Haegeman (1997, 2002b, 2011) and Ihsane (1998) observe that subject ellipsis in diaries is attested only in declarative root clauses and that it is not compatible with interrogative (yes/no questions and wh-questions) and embedded clauses. Thus, (28) is allowed in diaries, but (29) and (30) are not.

(28) Ø Like swimming.
(29) *She says that Ø likes swimming.
(30) a. *Do Ø like swimming?
    b. *Why do Ø like swimming?
b) argument/adjunct asymmetry

The observation of the phenomenon of the diary subject drop conducted so far (Haegeman 1997, 2007, 2011, Ihsane 1998) has shown that there is asymmetry with respect to the material preceding null subjects: DNS can have a preposed adjunct (31) while they cannot have a topicalized complement (32).

(31) At 3:30 today Ø had a very interesting conversation with Gen[eral] Eisenhower. (Truman Diary, 1947, 25 July)

(32) *This jacket Ø don’t like.

Note that (31) provides evidence that null subjects in diaries do not necessarily occur sentence-initially. At the same time, (32) reveals that DNS are sensitive to the material which can precede them: diary null subjects can be preceded by a preposed adjunct but cannot have a topicalized argument. In Chapter 2 a theoretical account for this asymmetry will be provided. In Chapter 4 I will investigate whether these properties also constrain the subject ellipsis in the second conjunct of coordinate clauses.

All these syntactic constraints suggest that the diary subject omission is a grammatical (syntactic) phenomenon. Of course, pragmatic factors also play a role in the omission of subjects and for a full description of the distribution of null subjects the pragmatic factors should also be taken into account. Before discussing the pragmatic factors, I will investigate the nature of this syntactically active empty category comparing the properties of DNS with those of the other empty categories.

1.4 The nature of diary null subjects

In this section I consider the four potential empty categories distinguished within the Government and Binding Theory (Chomsky 1982, Haegeman 1994, see also Haegeman and Ihsane 1999) to find out which of these empty categories can account for the analysis of DNS:

- pro
- PRO
- A-trace
• A´-trace

I discuss each of these empty categories in turn.

1.4.1 pro

At first sight, it is tempting to consider pro as a potential candidate for diary null subjects as like DNS, pro is also an implied, syntactically active empty category (see also Scott, 2010). Haegeman (2011: 8) suggests that treating DNS as pro could be viewed as a register-specific “parametric resetting”. However, there is an apparent negative setting of the pro-drop parameter in diary registers in that the latter lack the syntactic properties of pro-drop languages21. Specifically, in diaries we have no post-verbal subjects as shown in (7) repeated here as (33), no cases of that-trace violation as seen in (8), repeated here as (34). Besides, the inflectional endings of a verb in diaries do not differ from those in Standard English (7b). Hence, DNS cannot be assimilated to pro.

(33) a. Ha telefonato il decano.
      has telephoned the dean.
      ‘The dean has phoned.’

b. *Has telephoned the dean.

(34) a. Chii credi che abbia telefannato ti?
      Who believe you that have (subj.) telephoned?
      ‘Who do you think has called?’

b. *Who, do you think that t has called?

Moreover, English diary drop is not attested with subject-auxiliary inversion (35b) and with wh-movement (36b), as demonstrated in Section 1.3.1.3, whereas pro-drop languages like Italian license a null subject in both cases, as shown in (35a) and (36a) respectively (Haegeman 1994, 2011, Haegeman and Ilhane 1999, Weir 2008). (35) is from Haegeman (2000: 140), also reported in Scott (2010: 26), (36) is from Haegeman (2011: 9).

---

21 See Section 1.2.
(35) a. Tornerà [pro] presto?
   Return-FUT-3SG soon?
   ‘Will he/she return soon?’

b. *Will Ø return soon?

(36) a. Quando [pro] tornerà?
   When return-FUT-3SG?
   ‘When will he/she return?’

b. *When will Ø return?

Finally, pro-drop languages allow for a non-overt subject to be preceded by an argument (37a), while this pattern is incompatible with DNS\(^{22}\) (37b) (Haegeman 1997, 2011). The example is from Haegeman (2011: 10).

(37) a. Questo libro, [pro] non lo voglio.
   This book non it want-1SG?
   ‘This book, I don’t want.’

b. *This book Ø don’t like.

1.4.2 PRO

PRO is the null subject which occurs in non-finite clauses as illustrated in (38) (from Haegeman and Ihsane 1999: 119, also reported in Weir 2008: 32)

(38) The cat expects [PRO to get regular meals].

PRO does not occur in finite clauses and does not “alternate with an overt subject” (Weir 2008: 32, see also Haegeman and Ihsane 1999, Scott 2010). This means that PRO cannot be a potential candidate for DNS as the latter occur in finite clauses and can “alternate with overt subjects” (Weir 2008: 32).

\(^{22}\) See also (32) in Section 1.3.1.3.
1.4.3 A-trace

An A-trace is left after a movement to an argument position (A-position), e.g., “in cases of passivization [(39a)] and rising” (39b) (Scott 2010: 23).

(39) a. Tom was killed.
    b. She wants [t. to travel around the world].

A-traces do not “alternate with overt subjects” and, hence, cannot qualify for the analysis of DNS (see also Scott 2010, Weir 2008).

1.4.4 A´-trace

An A´-trace is related to a movement to a non-argument position, i.e. an A´-movement, and is left in case of wh-movement (40) and topicalization (41).

(40) Who did you speak to there?
(41) That jacket I don’t like.

To be able to answer the question whether an A´-trace can account for the diary subject omission let us discuss the properties of this type of movement.

1.4.4.1 The phenomenon of topic drop

Based on Raposo (1986), Haegeman (1997, 2011) observes that in some languages like European Portuguese the topicalized constituent may remain implicit. (42) is from Haegeman (1997: 240). The topicalized object leaves an A´-trace but it itself is non-overt. The null topic operator in SpecCP binds the object trace while the implicit operator is recovered from a salient antecedent.

(42) OP A Joana viu t. na televisao ontem à noite.
    the Joana saw ec on television last night.
‘Joana saw (it) on television last night.’

Haegeman (2011: 11) reports that “[a]rgument ellipsis may affect an argument in the main clause [(43a)] or in the embedded clause [(43b)].” The examples are from Raposo (1986: 381), also reported in Haegeman (2011: 11).

(43)  
(a) O Manel guardou _ no cofre da sala de jantar.
Manuel kept _ in the safe of the dining room.
(b) Eu disse ao António [que gardase _ no cofre da sala de jantar].
I told Antonio that he asked Manuel to keep _ in the safe of the dining room.

According to Haegeman (2011), Raposo (1986) elaborates an account for the movement of the null topic. Specifically, “the topic is moved to the CP of the matrix clause where it is allowed to be non-overt” (Haegeman 2011: 11). There are three island-domains which do not allow for the object extraction: the object cannot be extracted from “within the casual complement of a noun” (44a), “from inside an adjunct” (44b) and “from within an embedded interrogative” (44c) (Haegeman 2011: 11). The examples are from Raposo (1986: 381-383), reported in Haegeman (2011: 11-12).

(44)  
(a) *Eu informei a policia da [possibilidade [de o Manel ter guardado _ no cofre da sala de jantar]]. (Raposo 1986: 381: (16c))
I informed the police of the possibility that Manuel had kept _ in the safe of the dining room.
(b) *O pirata partiu para as Caraibas [dep ois de ter guardado _ cuidadosamente no cofre]. (Raposo 1986: 382: (19))
The pirate left for the Caribbean after he had kept _ carefully in the safe.
(c) *Eu sei [[em que cofre] o Manuel guardou _] (WH-island) (Raposo 1986: 383: (20))
I know in which Manuel kept _.

Haegeman observes that this analysis accounts for the topic drop patterns which occur in German and Dutch. (Haegeman 2011: 12). The data in both languages show that “a topical argument may be non-overt if moved to the left edge of the clause” (Haegeman 2011: 12). As both are V2 languages, a fronted topic will occur immediately before the finite verb and in
case of a topic drop the verb will occupy an initial position in the sentence (Haegeman 2011: 12). (45) gives examples from German, taken from Haegeman (2011: 12). In (45a) the sentence has a fronted topic *das* (‘that’), in (45b) the fronted topic is implicit. (45c) is not grammatical because only the topics in the leftmost position can be implicit (Haegeman 2011: 12).

   (45) a.   Das habe ich schon gesehen.
             that have I already seen
             ‘I have already seen that/it.’
   b.   ec Habe ich schon gesehen.
   c.   *Ich habe ec schon gesehen.

Further Haegeman (2011) observes that the topic drop analysis does not account for expletive subjects. They cannot be non-overt even if they occupy an initial position in the sentence. (46) is from Haegeman (2011: 12).

   (46) *(Es) wurde viel getanzt.             (German)
         (It) was a lot danced.

1.4.4.2 Diary subject ellipsis as topic drop?

Based on Huang’s (1984) proposal that the “null subject is an A´-trace bound by a non-overt topic operator in CP domain” (Haegeman, 1997: 240), Haegeman (1990a, 1990b) proposes that the topic drop analysis can account for DNS (Haegeman 1997). Haegeman (2011: 13) points out that topic drop “analysis would account for a number of restrictions on” DNS. These are presented below as in Haegeman (2011: 13).

   a)   The implicit subject can be interpreted either as the diary writer or as a salient referent in the discourse context (see also Section 1.3.1.1).
   b)   DNS are not allowed in adjunct and relative clauses: as these are island-domains, “the subjects would have to be illicitly extracted” (Haegeman 2011: 13).
   c)   Diary subject ellipsis is not licit “in complement clauses introduced by an overt C-element”: the subject cannot be extracted across an overtly realized C-element as it
would lead to the violation of the *that* trace filter (Haegeman 2011: 13) (47).

(47) This book I don’t think that *(that) this book is interesting.

d) DNS are not compatible with argument topicalization (cf 32); for the derivation of (32), both the subject and the object must be fronted with the subject being implicit. “However, in English multiple topic fronting is independently known to be degraded” (Haegeman 2011: 13). (48) is from Haegeman (2011: 13).

(48) *John, this book, I will give.

e) Given that “English fronted topics are not easily compatible with yes/no questions [(49a)] or with fronted wh-items” (49b), diary null subjects are not compatible with yes-no questions (30a) and wh-questions (30b).

(49) a. *Her story about whales, did you really believe?  
   b. *The story about whales, when did she tell?

Though initially attractive, this analysis raises a number of problems. I will outline them as given in Haegeman (1997, 2011).

a) While subjects can be commonly dropped, object ellipsis does not occur systematically in diaries (50).

(50) The Governor visited *(me) and we discussed the recent political affairs.

b) Expletive subjects can be non-overt in diary registers (51a) though they cannot operate as topics (cf. 46). (51b) is from Haegeman (2011: 14).

(51) a. Ø Makes a person ashamed to be gloomy even if world affairs are mixed up. (Truman Diary, 1947, 25 Dec.)  
   b. Cloudy & humid all day Ø sprinkled tonite. (Carolyn’s diary, 1964, 29 August)
c) Embedded null subjects are incompatible with non-fiction diaries (52) while “the Portuguese null topic operators may bind an embedded trace” (Haegeman, 1997: 240), (cf. 43b). (53) is from Haegeman (1997: 240).

(52) *He said that Ø will come tomorrow.

(53) OPi O José sabe que a Maria viu ti. (Raposo, 1986)
that Jose knows that the Mary saw ec.
‘Jose knows that Mary saw (it).’

It follows that none of the four empty categories discussed above are compatible with the properties of diary null subjects and, hence, none of them can account for the analysis of DNS. For the moment I will not try to give any solutions to the problem. In Chapter 2 Haegeman’s (1997) proposal on the nature of DNS will be presented which is fully compatible with the non-fiction diaries investigated so far. I will test whether her hypothesis is borne out in my corpus. (54) summarises Haegeman’s proposal.

(54) Diary null subjects are antecedentless non-variable empty categories which are licensed by CP truncation.

1.5 The pragmatic conditions on subject omission

While the preceding discussion proves that diary subject omission is a syntactic phenomenon, the relevance of pragmatics in interpreting the non-overt subject patterns cannot be neglected. In some instances syntax plays an essential role for the identification of the implicit subjects (the third person inflectional ending, reflexive pronoun, etc.). However, I postulate that questions like why overt and null subjects alternate with each other throughout a diary, and why diary subjects are omitted, can be interpreted by discourse-related factors while syntax would be a poor candidate here.

Haegeman (2011) and Weir (2008) also stress the importance of pragmatics in diary ellipsis but propose that only pragmatics cannot give an answer to many questions and here is when the syntax should be addressed. “There is clearly a pragmatic, or discourse-related, component to subject pronoun drop in written English, insofar as only subjects that can be reconstructed from the discourse (and not, for example, from verbal morphology as in the
case of pro-drop …) can be dropped. But relying solely on pragmatic factors cannot explain the distribution of the phenomenon which we see” (Weir, 2008: 29).

Before drawing any final conclusions, let us first consider some pragmatic factors which favour/disfavour DNS. Section 1.5.1 focuses on the recoverability of null subjects. Section 1.5.2 considers the economy deletion hypothesis and Grice’s maxims as potential explanations for the diary drop phenomenon.

1.5.1 Recoverability: syntax or pragmatics?

Recoverability is one of the most essential factors favouring ellipsis, and prima facie it is difficult to say whether it has syntactic or pragmatic foundations. On closer observation it becomes obvious that both are at stake. From the syntactic point of view, an element can be omitted if it is recoverable (Quirk et al 1972: 536). This means that the omitted constituent should have a salient referent in the linguistic context which will make the recoverability possible. But is this always the case in diaries? Does the recoverability of DNS depend only on the linguistic context?

Examination of diary data reveals that the omitted subjects can be recovered either from the linguistic or the extralinguistic context. Specifically, in case of I-omissions the subject can be recovered from the extralinguistic context: given our general knowledge concerning the nature of diary writing, namely that the text refers to the writer and his or her life events, we assume that I is the default pattern for omission, as already noted above. The other person/number categories can be recovered from the linguistic context due to the availability of a salient referent. It follows that the non-overt subject is commonly interpreted as first person singular unless there is a salient referent in the context which points in the other direction. Consider (55).

(55) a. Went down F St[.] and back G. Like to look in merchants’ windows. Had breakfast at 9 A.M. At 12:45 had the G. D. message in shape. Read & reread. Spent the afternoon in study on the same message and the Economic one too. (Truman Diary, 1947, 5 Jan.)

b. The Ambassador speaks terrible English. Said Argentine [sic] wanted to get along with us, etc. (idem, 31 March)
In (55a) pragmatics plays a significant role for the recoverability of the subjects: all 7 clauses lack a subject and there are no antecedents to make the retrieval of the deleted elements possible. But based on our general knowledge that the diary writing is about the narrator (i.e. I), it seems plausible to recover the omitted subjects as I. Thus, we can tentatively assume that somehow in our mental representation the first person subject is ‘always available’ as an antecedent.

In (55b), on the other hand, the null subject of said in Said Argentine [sic] wanted to get along with us, etc. can be easily recovered as he because there is a salient referent in the context - the lexical NP the Ambassador which is the subject of the preceding clause.

Observe that the only inflectional ending denoting the third person singular –(e)s can serve as a cue for retrieving the omitted subject but it does not always ensure the sufficient interpretation of the null subject (56).

(56) Reads a lot.

Although it is clear that the subject is a third-person-singular entity, it is still ambiguous whether the non-overt subject should be identified as he or she. This will be possible only if there is an antecedent in the discourse context.

Thus, the recoverability of diary null subjects depends on both syntax and pragmatics and the role of neither of them should be underestimated. While I-omissions ‘arise out of the blue’ (Haegeman, 2011: 6) and can be retrieved based on the general knowledge, the recoverability of other person/number categories depends on a salient referent in the discourse context. A question arises: why are diary subjects omitted? I consider this in the following section.

1.5.2 Economy

An option to explain the phenomenon of DNS would be to say that diary subjects are omitted for reasons of economy. With respect to economy two principles seem relevant:

a) Chomsky’s (1981) Avoid Pronoun Principle;
a) Avoid Pronoun Principle - Given that diary null subjects are interpreted as pronouns\(^{23}\), it can be postulated that they arise as a result of the *Avoid Pronoun Principle* (57) (Chomsky, 1981). The principle can be considered “a subcase of the conversational principle of not saying more than is required, or might be related to a principle of deletion-up-to recoverability, but there is some reason to believe that it functions as a principle of grammar.” (Chomsky 1981: 65)

\[
\begin{align*}
(57) & \quad \text{Avoid pronoun.} \\
& \quad \text{Avoid overt pronoun, whenever possible.}
\end{align*}
\]

Though initially attractive, this proposal does not seem to apply to diaries. Avoiding a pronoun does not imply only omitting subject-pronouns but possessives and object-pronouns as well. Empirical data prove it otherwise. So far we have no cases of object or possessive pronoun deletion.

Besides, this principle cannot account for DNS: in diaries pronominal subjects can be realized both overtly and non-overtly while the *Avoid Pronoun Principle* is mandatory, not optional. Consider (58).

\[
\begin{align*}
(58) & \quad \text{a. He remembers him closing the door.} \\
& \quad \text{b. He remembers closing the door.}
\end{align*}
\]

In (58a) *him* is overt as it is not co-referential with the subject of the sentence, *he*: the verbs *remembers* and *closing* have different subjects. On the other hand, in (58b) the subject of the verb *remembers* is co-referential with the subject of the verb *closing*, and the second pronoun has to be non-overt (PRO).

Based on the discussion above I conclude that the *Avoid Pronoun Principle* should be ruled out as a potential explanation for the diary subject omission.

b) Grice’s Maxim of Quantity – Another plausible explanation to the phenomenon of DNS would be Grice’s Maxim of Quantity (59) (Grice, 2004: 45).

---

\(^{23}\) See sections 1.3.1.1
(59) 1. Make your contribution as informative as is required (for the current purposes of the exchange)
   2. Do not make your contribution more informative than is required.

*Grice’s Maxim of Quantity* might have two implications for the realization of subjects in diaries: (i) diary subjects are omitted to avoid redundancy (60); (ii) diary subjects are overtly realized to avoid ambiguity (61).

(60) Arose at 7:30, shaved, dressed and had breakfast at 8:15 with Adm[iral] Foskett the naval aide and Capt[ain] Freeman, Commander of the Williamsburg. Arrived at the White House about 9 A.M. Had a beautiful snow the night before. Trees in the White House yard were beautiful. R[ea]r Adm[iral] Foskett came to the House with me and then went home. Spent the day working on messages. Called all the members of the Cabinet[,] wished them a happy New Year. Called Henry Stimson, Miss Perkins, Gen[eral] Eisenhower and Gen[eral] Flemming too. (Truman Diary, 1947, 2 Jan.)

(61) Spent the day working on State of the Union, Economic and Budget messages. Having a terrible time with the Economic one. The Commission have no White House experience. I've turned them over to Steelman[,] Harriman, Snyder and Schwellenbach, and I hope for the best. The awful 79th Congress put me on the spot. Now I've a job putting the 80th on the same spot to make us even. (idem, 4 Jan.)

If all the clauses in (60) had an overt subject, we would have to pronounce the pronoun *I* 9 times in one paragraph. Given that *I* is recoverable as it refers to the author and can be easily understood even if it is implicit, as discussed above, it seems plausible to conclude that the author does not use the first-person-singular subject every time to avoid redundancy.

In (61) *I* is omitted once\(^{24}\) and is overt 3 times. If we examine the entry, we see that in one case *I* is overtly realized in a coordinate clause where the subjects of the two conjuncts are co-referential. It follows that *I* could have been omitted without giving rise to ambiguity but it is overt. In the remaining 2 cases *I* is overt as there is a change of topic, i.e. the subject of the preceding clause is other than *I* and the author uses the pronoun to avoid ambiguity.

\(^{24}\) Recall that I do not consider the cases with subject+auxiliary omission.
Observe, however, that in both cases the subjects occur with a contracted verb which means that they cannot be deleted irrespective of the change of the topic. The contracted verb is a clitic and needs a host, so the subject cannot be deleted\textsuperscript{25}. This means that the overt realization of these two subjects is determined by a syntactic factor.

Observe also that, as discussed above, DNS are unavailable in embedded environments as well as in questions, which is another indication that a subject is overtly realized not only for pragmatic reasons but also because DNS are subject to certain syntactic constraints. The analysis shows that although ambiguity might sometimes explain why a subject is overt, it cannot account for all the overtly realized instances; in certain cases the syntactic factors are determining.

To conclude I can say that while syntax plays a major role in identifying the patterns of diary null subjects, the importance of the discourse-related factors cannot be underestimated. Rather, syntax and pragmatics complement each other and both should be accounted for to make a full picture of the diary drop phenomenon.

\textsuperscript{25} See Section 4.4.1 for discussion.

2.1 Introduction

It has been discussed in the preceding chapter that DNS are subject to a set of syntactic constraints which have been examined by Haegeman in a series of papers (1990a, 1990b, 1997, 2000, 2007, 2011). Her proposals, which are adopted as basis for data analysis in this thesis, will be reported in this chapter.


Like the early null subject, the (adult) null subject in abbreviated registers is an antecedentless empty category in the A-specifier of the root. Null subjects depend on the truncation of CP, which turns the specifier of IP into the highest specifier of the clause. (Haegeman 1997: 233)

As noted above, I do not discuss the account for child language null subjects in this thesis and refer the reader to Rizzi (1994, 1997, 2006) for detailed analysis (see Hyams (1986) for an alternative account). Nevertheless, I present some of the core proposals of Rizzi’s accounts (1994, 1997), namely, those that are essential for the analysis of DNS. Though in this chapter I focus on the theoretical proposals suggested in Haegeman (1997), I would like to point out that in later work Haegeman (2007, 2011) develops a second account for DNS “based on Rizzi’s (2006) analysis of null subjects in early child production, using the cartographic model of formal syntax (Cinque and Rizzi 2010) and insights from Phase theory (Chomsky 2001)” (Haegeman, 2011: 3). Given that both analyses focus on the same phenomenon and both account for the syntactic constraints identified in diaries, only one of them will be presented in this thesis: I will discuss the analysis presented in Haegeman (1997) and refer the reader to Haegeman (2007, 2011) for the phase-based analysis.

Recall that, as discussed in Chapter 1, DNS are a root phenomenon, i.e. they are incompatible with subject-auxiliary inversion, wh-preposing and embedding. Diary null subjects are attested with preposed adjuncts, but not with preposed arguments. With this in mind my next step will be to show how Haegeman’s proposals can account for these syntactic constraints attested in the empirical data.
Chapter 2 is organized as follows: Section 2.2 revisits the discussion on the nature of DNS and outlines Rizzi’s (1994) proposals on which Haegeman’s (1997) analysis is built. Section 2.3 deals with Haegeman’s (1997) account for adjunct/argument asymmetry with respect to diary null subjects. Based on the analysis elaborated in the previous sections, Section 2.4 argues that null subjects and their pronominal counterparts differ. Section 2.5 shows that subject omission in the second conjunct of coordinate clauses is governed by the same restrictions as DNS.

2.2 The nature of the DNS revisited

In the previous chapter it was concluded that a diary null subject is an empty category. When trying to identify the nature of this empty category it was observed that neither the pro-drop analysis nor the topic drop analysis are sufficient for DNS. Haegeman (1997) proposes that, differently from these two patterns in which a null category is bound by an A´-element in the left periphery, diary null subjects are antecedentless empty categories. Let us discuss this step by step.

2.2.1 The typology of empty categories

According to the Government and Binding theory (Chomsky 1981, 1982, see also Haegeman 1994, Scott 2010), the typology of empty categories is determined by two properties: [± anaphoric] and [± pronominal]. Based on Lasnik and Stowell’s (1991) proposal, Rizzi (1994) elaborates a more refined typology of A´-traces26 (cf. Haegeman 1997: 241). Specifically, he suggests that for distinguishing A´-bound traces, which typically have the features [-anaphoric, -pronominal], another feature should be introduced: [± variable]27. A variable trace implies “quantification ranging over a possibly non-singleton set, in the latter the null element never ranges over a non-singleton set, rather it has its reference fixed to that of the antecedent” (Rizzi 1994: 158, see also Haegeman 1997: 242). A variable trace is left in case of wh-movement, a non-variable trace is related to the movement of null operators28.

26 See Section 1.4.4 for discussion of A´-traces.
27 See also Haegeman (1997: 242).
28 See Lasnik and Stowell (1991) for relevant discussion.
(64) illustrate the difference between the two. The examples are from Haegeman (1997: 242), also reported in Ihsane (1998: 34).

(62) gives examples of A´-movement. In (62a), the trace in the complement position of the preposition to is represented as ti which is left as a result of the wh-movement of the interrogative element who; it is a variable. In (62b) the assumption is that there is a trace, ti, in the complement position of to which is bound by the null operator OP and it is a non-variable.

(62) a. Whoi did you get John to talk to ti?
   b. Bill is easy [OPi [PRO to get John to talk to ti]]

(63a) and (63b) are ungrammatical with the interpretation in which the complement of to is co-referential with him, because in both examples the A´-movement – the “wh-movement” in (63a) and the “null operator movement” in (63b) - leads to a “strong crossover: the movement must not cross a coindexed pronoun” (Haegeman, 1997: 242). In both examples the moved elements, who in (63a) and OP in (63b), cross the coindexed pronoun him. In these configurations the pronoun him c-commands and binds the trace in A´-position, which leads to the violation of the Principle C of the Binding theory (Chomsky 1981: 188).

(63) a. *Whoi did you get himi to talk to ti?
   b. *Johni is easy [OPi [PRO to get himi to talk to ti]]

(64) shows that while the movement of an overt wh-phrase and that of a null operator share certain properties, as shown in (63), they are not to be fully assimilated. In (64a) the fronted wh-phrase gives rise to a weak crossover violation: it illicitly crosses the coindexed pronoun his which is part of a constituent (his mother). As shown by the grammaticality of (64b), the movement of the null operator apparently does not give rise to a weak crossover violation (Haegeman 1997: 242). This difference between the two traces left by the movement is conditioned by the feature [+ variable]: the trace left by wh-movement is a variable while the one left by a null operator is not.

(64) a. *Whoi did you get his motheri to talk to ti?
   b. Johni is easy [OPi [PRO to get his motheri to talk to ti]]
2.2.2 The Empty Category Principle

Since Chomsky (1981), it is standardly assumed that non-overt elements are regulated by the Empty Category Principle (ECP) (see also Rizzi 1994, Haegeman 1997: 242), presented in (65).

(65) a. Licensing: an empty category must be head-governed.
   b. Identification: an empty category must be identified by a chain-connected antecedent.

In his account for null subjects in the early production Rizzi (1994) modifies the identification clause (65b) as in (66):

(66) A non-pronominal empty category must be chain-connected to an antecedent if it can be.

(66) stipulates that a “non-pronominal empty category” doesn’t have to “be chain-connected to an antecedent” if it cannot. Basing on (66), if there is no maximal projection to c-command an empty category, then there is no available position for a potential antecedent; hence, the empty category “cannot be chain-connected to an antecedent” and will “remain antecedentless” (Haegeman 1997: 242).

Haegeman (1997: 243) observes that “antecedentless empty categories cannot be [+variable]: variables need an operator to bind them. Non-variable empty categories, on the other hand, may in principle remain antecedentless as long as they occur in a position for which there is no potential antecedent.”

As we shall see, the reformulation of the identification condition of empty categories presented in (66) plays a significant role in terms of the distribution of DNS; it can explain why diary null subjects do not occur in embedded clauses and wh-questions.
2.2.3 Root clauses

With respect to the adult grammar, Rizzi (1994, 1995) makes two proposals given in (67):

\[(67) \begin{align*}
(a) & \text{ Root } \rightarrow \text{ CP} \\
(b) & \text{ Avoid structure.}
\end{align*}\]

According to (67a), CP must be projected in every root clause. This can be accounted for if we assume that C denotes the illocutionary force of the sentence and that “all root clauses must have an illocutionary force” (Haegeman, 1997: 243). Based on the economy principle hypothesis (Chomsky 1991), (67b) implies that only the required structure will be projected. It follows from (67b) that no projections can be inserted to salvage a structure which would otherwise give rise to ungrammaticality. For instance, one cannot insert a projection only for the reason of avoiding the violation of ECP.

2.2.4 Truncation

In his account for the non-overt subjects in the child production data Rizzi (1994, see also Haegeman 1997) introduces the concept of structural truncation: while, in general, root clauses are CP (according to 67a), there are exceptions to this principle and, hence, root clauses do not always have to be projected at CP level. In certain circumstances root clauses may be truncated, and the projection can be terminated at a level which is lower than CP. Rizzi’s assumption is that in the default case, in the adult language root clauses cannot be truncated and in root clauses the CP level must be projected (Haegeman 1997: 244) but that in child grammars root clauses can be truncated at IP level (68).

\[(68) \text{ Want more. Rizzi (1994: 1, from Hyams, 1986)}\]

Assuming a truncation account, Rizzi (1994) then analyses early null subjects as antecedentless non-variable empty categories in SpecIP (69). The ec in SpecIP does not give

---

29 See also Haegeman (1997: 243).
30 See also Haegeman (1997: 243).
rise to the violation of the ECP identification principle: it occurs in the highest position of the sentence and because there is no potential antecedent, it can remain antecedentless. This leads to the prediction that the child null subject cannot occur in a clause in which CP is projected for independent reasons, as this will mean that there is an available position (SpecCP) to host a potential antecedent and, thus, the identification through an antecedent becomes mandatory. This can explain why early subjects cannot be non-overt in embedded clauses and \textit{wh}-questions as in both cases the CP layer must be projected. Recall that DNS are incompatible with embedding and \textit{wh}-questions, too.

\begin{equation}
\text{(69)}
\end{equation}

Since, as observed in Section 1.3, null subjects in the early production and in diary registers are subject to similar distributional constraints, Haegeman (1997) extends Rizzi’s (1994) proposals to the analysis of diaries and postulates that a diary null subject is “an antecedentless non-variable empty category in the specifier of IP, with the functional projection CP being truncated” (1997: 245). Following Rizzi (1996 class lectures) Haegeman (1997: 245) suggests that CP truncation in the child language occurs due to the economy principle. She proposes “that in the child grammar [(67b)] is stronger than [(67a)], while in the adult grammar [(67a)] wins out over [(67b)]” (Haegeman 1997: 246). Given that in the abbreviated registers economy (or ‘brevity’) plays a role, Haegeman (1997: 246) suggests that like the child grammar, in diaries, too, (67b) prevails over (67a) “by economy considerations”.

As mentioned for early null subjects, the hypothesis pursued by Haegeman (1997) implies that DNS will be illicit in embedded clauses and \textit{wh}-questions because in both cases the CP layer must be projected and, hence, the empty category in the subject position of the
embedded clause is subject to the identification requirement and cannot remain antecedentless. As shown in Chapter 1, diary style null subjects are not attested in embedded clauses or in interrogative clauses.

From the account outlined above, it would seem to follow that null subjects in diaries must occupy the highest position in the clause and that no material can precede DNS: if the subject is not the “highest specifier in the structure”, then there is a maximal projection available to c-command the empty category, hence, there is a position to host a potential antecedent. This means that the empty category will be “chain-connected to an antecedent” and cannot remain antecedentless (Haegeman 1997: 242). At first sight, though, this expectation is only partially borne out. Although diary subjects cannot be non-overt in the presence of a preposed argument, which is in line with the hypothesis elaborated above, DNS are compatible with adjunct-preposing, as already discussed in Chapter 1 (Section 1.3.1.3). I repeat the examples in (31) and (32) here as (70) and (71).

(70) At 3:30 today Ø had a very interesting conversation with Gen[eral] Eisenhower.
(Truman Diary, 1947, 25 July)

(71) *This jacket Ø don’t like.

As expected, (71) is illicit as it violates the ECP. (70), on the other hand, is licit and the representation given in (70) is problematic: if null subjects are not the leftmost constituents in the sentence, this should lead to the ECP violation because preposed material implies that there are higher projections in the clause and, hence, one of these can host a potential antecedent for the non-overt subject. This gives rise to a problem: if the ECP can account for the ungrammaticality of (71), the fact that examples such as (70) are attested and, I assume, are acceptable, goes unexplained. Haegeman (1997) provides an account which explains how adjuncts can precede DNS without giving rise to the ECP violation. The next section deals with her proposals.
2.3. Adjunct preposing

2.3.1 The problem and proposals

As seen in the previous section, the compatibility of adjunct preposing with DNS (72) gives rise to a problem for the hypothesis that diary null subjects are antecedentless empty categories.

(72) At 3:30 today Ø had a very interesting conversation with Gen[eral] Eisenhower. (Truman Diary, 1947, 30 Sept.)

Haegeman (1997) considers several alternative proposals for the adjunct/argument asymmetry. First, it could be related to the intrinsic properties of the preposed material in the sense that one might postulate that a non-argument constituent cannot be an antecedent for an argumental empty category. If the adjunct cannot be a potential antecedent for a non-overt subject, then the sentences with a preposed adjunct do not violate the ECP. Haegeman (1997) observes, however, that if this line of reasoning were correct, then *wh*-adjuncts should also be compatible with diary null subjects. However, no examples of fronted *wh*-elements and null subjects have been found so far (73).

(73) %*When will visit us?

A second hypothesis might be that adjuncts are compatible with DNS because they are adjoined to IP rather than occupying the specifier of a projection, as proposed by Lasnik and Saito (1984), among others. Haegeman (1997) challenges Lasnik and Saito’s (1984) proposal that sentence-initial adjuncts are in an adjoined position. Based on Rizzi (1997), she suggests that fronted constituents occurring with DNS are not “merely adjoined to projections but that they are associated with their own functional head” (Haegeman 1997: 251). The examples in (74) provide empirical evidence for this assumption (taken from Haegeman (1997: 251).

(74) a. *This is a law which I think [CP that [IP t will be approved next week]]

b. This is a law which I think [CP 0 [IP t will be approved next week]]

c. This is a law which I think [CP that next week [IP t will be approved]]
(74a) is ungrammatical because there is a that-trace violation: the subject cannot be extracted from a lower projection across the complementizer that. The latter “is inert for head-government” and the subject trace remains ungov erned (Haegeman, 1997: 251). (74b) is grammatical: following Rizzi (1990), the assumption is that in this example the complementizer has been substituted for by a non-overt counterpart, represented as 0, and that this null C element is able to carry AGR features. The complementizer is no longer “inert” and the trace of the subject can be governed. In (74c), in spite of the presence of the complementizer that which precedes a subject trace, the sentence is licit: this is apparently achieved by inserting the adjunct next week after the overt complementizer. If the adjunct can improve the sentence, this suggests that the presence of the preposed adjunct involves the presence of an additional functional head which can head-govern the subject trace.

Considerations such as those above, as well as additional comparative data have led Rizzi (1997) to postulate what has come to be known as the ‘split CP’: “a more articulated structure for the CP domain, in which CP is decomposed into a number of functional projections” (Haegeman 1997: 251). The next section briefly discusses the proposal.

### 2.3.2 The Split CP hypothesis

According to Rizzi (1997), the head C, in fact, consists of two functional heads: the first one, Force°, “encodes the illocutionary force” (the type of clause); the second one, Fin°, “encodes finiteness” (the type of selected IP), (Haegeman 1997: 251). In addition to ForceP and FinP there are two more projections - Topic (TopP) and Focus (FocP). TopP is activated if there is a topicalized element in a sentence (75a), FocP is projected if the sentence contains a focalized element (75b). Rizzi (1997) suggests that TopP can sit both before and after FocP in a projection. Both TopP and FocP are optional and are projected only if required (Rizzi 1997).

(75) a. This book I have read.
    b. Never again will I read this book.

Force° and Fin° split when “the Topic-Focus field is activated” (Rizzi, 1997: 313). (76) illustrates the CP structure. The stars denote the recursive projections; the parentheses mean
that the projections are optional (Rizzi 1997). I refer the reader to Rizzi (1997) for relevant analysis of the ‘split CP’.

(76) \[ \text{ForceP} > (\text{TopP})^* > (\text{FocP}) > (\text{TopP})^* > \text{FinP} > \text{IP} \]

In terms of the articulated CP structure (76), (74c) is represented as in (77). The adjunct \textit{next week} sits in the specifier of TopP while Top\textsuperscript{o} remains available. The adjunct “provides the head” (TopP) that can govern the subject trace (t\textsubscript{i}) (Haegeman 1997: 252). That is not the functional head any longer and the adjunct can head-govern the subject trace properly. “The finiteness head (Fin\textsuperscript{o}) is separately projected; the subject moves via [Spec, FinP] where it triggers agreement” (Haegeman 1997: 252). Associated with Agr features, Fin\textsuperscript{o} ensures that the subject trace (t\textsubscript{i}) is head-governed properly. Fin\textsuperscript{o} moves to Top\textsuperscript{o} by head movement and provides a position from which it can govern the intermediate subject trace (t\textsubscript{i}) in [Spec, FinP]: the subject has left an intermediate trace in [Spec,FinP] while moving through this position, as described above (Haegeman 1997: 252).
The interaction between negative inversion and preposed adjuncts in English provides further support for the proposal that fronted adjuncts are associated with a specific functional head (Haegeman, 1997). (78) and (79) are from Haegeman (1997: 252-53).

(78) a. I swear that on no account will I write a paper during the holidays.
   b. I swear that during the holidays I will write a paper.

(78a) is an example of negative inversion. Haegeman (1997: 252) proposes that the auxiliary will is in FocP and the negative constituent is in [Spec, FocP]. In (78b) we have an adjunct topicalization where the PP during the holidays sits in [Spec, TopP].

Consider (79). (79a), in which the topicalized adjunct precedes the negative inversion structure, is grammatical. (79b), in which the topicalized adjunct follows the negative inversion structure, is ungrammatical.

(79) a. I swear that during the holidays on no account will I write a paper.
   b. *I swear that on no account will during the holidays I write a paper.

The difference in grammaticality between the two clauses gives further evidence that adjunct fronting is associated with a functional head. In (79a) the auxiliary will can move to Focº through Finº without crossing the head Topº because TopP is projected higher. In (79b), on the other hand, will has to cross the head Topº in order to move to Focº. This will lead to the violation of the Head Movement Constraint, according to which “a head X cannot skip an intervening head Y” (Haegeman 1994a). Accordingly, Topº doesn’t allow will to move to Focº. (80a) and (80b) provide the representations for (79a) and (79b) respectively.
that during the holidays on no account will I
Thus, the fronted adjunct is associated with a functional projection and provides additional evidence for the Split CP.

Haegeman (1997: 254) shows that assuming the articulated CP “has implications for truncation theory”\(^{31}\) and proposes that when FocP is truncated, the lower projections (TopP) can remain available for hosting a topicalized constituent. Using this modification of the truncation account, she makes a proposal which can account for the adjunct/argument asymmetry: implicit subjects by-pass the adverbial adjunct and leave a co-indexed trace in [Spec, IP]. These two hypotheses will be further elaborated in the following section.

\(^{31}\) See Section 2.2.4.
2.3.3 DNS in a split CP

This section focuses on Haegeman’s proposals to account for the adjunct/argument asymmetry introduced in Section 2.2.4. Recall that, as empirical data evidence, a preposed adjunct can co-occur with a diary null subject but a preposed argument cannot.

Given the split CP hypothesis, according to which CP consists of a number of functional projections, Haegeman (1997: 254) postulates that truncation can “apply at any level of the CP domain.” If the truncation can occur at different levels of CP, Haegeman suggests that “ForceP may be truncated leaving the lower projections available” (Haegeman 1997: 254). This proposal, however, as such does not solve the problem: Haegeman (1997) observes that even if ForceP is not projected, TopP remains a potential position to host an antecedent for the subject in [Spec, TP]. In (81) the empty category in [Spec, IP] is c-commanded by the adjunct at 3:30 today which sits in [Spec, TopP]. This representation again gives rise to ECP violation as there is a potential binder for the antecedentless empty category. The structure representation of (81) is given in (82).

(81) * At 3:30 today Ø had a very interesting conversation with Gen[eral] Eisenhower.
    (Truman Diary, 1947, 25 July)

(82)                                            * TopP
       Spec         Top´
       |
       At 3:30 today
       |
        Topº  FinP
        |
      IP

had a very interesting conversation …
And yet, (81) is acceptable. Thus, (82) cannot be the correct representation. Haegeman (1997) suggests that the non-overt subject in (81), in fact, is not in [Spec, IP]. Rather it moves out and by-passes the fronted adjunct in the specifier of TopP, “leaving a coindexed trace” in its canonical position (Haegeman 1997: 255). As a result, (81) will have a representation as in (83). The moved null subject is now the leftmost element and has no potential antecedent that can bind it.

(83) Ø at 3:30 today t i had a very interesting conversation with Gen[eral] Eisenhower. (Truman Diary, 1947, 25 July)

Given the argument-adjunct asymmetry\(^{32}\) with respect to the DNS, the by-passing strategy must be restricted to adjuncts and should not be extended to arguments: for instance, the analysis should not apply to representations such as (84a) and lead to the representation in (84b). We have to be able to assume that the preposed argument in (84a) remains a potential antecedent for the non-overt subject and, hence, the null subject cannot remain antecedentless. This means that adjunct/argument asymmetry has to be captured: the null subject in (83) should be able to cross a topicalized adjunct but it cannot by-pass a preposed argument (Haegeman 1997: 256).

(84) a. *This jacket don’t like.
   
   b. *Ø this jacket t i don’t like.

In order to account for the adjunct/argument asymmetry, Haegeman (1997) first observes that the asymmetry occurs more generally in sentences with subject extraction. She examines several French examples given in Rizzi (1997) illustrating the adjunct/argument asymmetry. In (85a) the subject of the embedded clause *qui* is extracted from the embedded domain to the matrix clause (Haegeman 1997: 256). It can cross an adjunct. In case of argument preposing “the by-passing is blocked” (Haegeman 1997: 256), as shown in (85b).

(85) a. Qui crois-tu qui l’année prochaine pourra nous aider?
   who think-you that year next will-be-able us help?
   ‘Who do you think that next year will be able to help us?’

\(^{32}\) See also Sections 1.3.1.3, 2.2.4.
b. *Qui crois-tu qui ton livre pourra l’acheter?  
who think-you that your book will-be-able it to buy?  
‘Who do you think that will be able to buy your book?’

(86) illustrates the adjunct/argument asymmetry on the basis of data with long wh-extraction.  
For (86a), the hypothesis is that like that in English, the French complementizer que “is inert for head government” and the subject trace remains ungoverned and violates the ECP licensing clause (Haegeman, 1997: 256). In (86b) and in (86c) que is replaced by qui and the assumption is that “qui is the spell-out of the agreeing complementizer which can head-govern the subject trace” (Haegeman, 1997: 256). We observe that, with qui replacing que in in (86b), the fronted adjunct l’année prochaine can be crossed by the wh-movement of the subject, apparently without violating the ECP. In (86c), however, as the ungrammaticality of the sentence shows, the subject cannot be extracted across the preposed argument. Below I provide Haegeman’s (1997) argumentation for this difference.

(86) a. *Voici l’homme que je crois que t pourra nous aider l’année prochaine.  
this is the man that I think that will-be-able us help year next.  
‘This is the man that I think that will be able to help us next year.’

b. Voici l’homme que je crois qui l’année prochaine pourra nous aider.  
this is the man that I think that year next will-be-able us help.  
‘This is the man who I think that next year will be able to help us.’

c. * Voici l’homme que je crois qui ton livre pourra l’acheter.  
this is the man who I think that your book will-be-able it to buy.  
‘This is the man who I think that your book will be able to buy.’

To explain this asymmetry in (86) Haegeman (1997) adopts Rizzi’s (1997) proposal for the derivation of subject extraction across adjuncts which is presented in (87). Rizzi (1997) suggests that:

[T]he head of a substantive (contentful) functional projection, ‘FP’, may freely be associated with AGR features, and that this association will license an agreement projection which immediately dominates FP. The features of the head of the dominating agreement projections will match the AGR features of Fº.

33 See Example (72a).
Applied to the CP domain, Top° may host AGR features and TopP may be dominated by an associated AgrP. (Haegeman, 1997: 256-57)

So, if, as (87) shows, TopP is dominated by AgrP, then the features of the head Agr° and those on Top° match.

Then, Rizzi proposes that rather than targeting a specifier position, fronted adjuncts adjoin to TopP. He then proposes that in case of adjunction to TopP the agreement features on the head Top° are not activated –since there is no specifier - and as a result, they remain available.

[Spec, AgrP] can serve as a landing site for (A)-movement. Haegeman (1997: 257) assumes that the subject may move to [Spec, FinP], cross the preposed adjunct and move to [Spec, AgrP]. “Carrying agreement features, the subject will activate the agreement head Agr” (Haegeman 1997: 257), the features of which match those on Top. The subject trace in FinP can be head-governed by TopAGR. The trace in FinP activates Fin° which governs the subject trace in [Spec, IP] (Rizzi 1997, Haegeman 1997).

According to Rizzi (1997), fronted arguments, on the other hand, must be hosted in [Spec, TopP]. As a result, the argument in [Spec, TopP] itself triggers the agreement features on Top°. Hence, the AGR features on Top° will be exhausted (Haegeman 1997: 257). The agreement features of the argument and those on Top° match. Therefore, the subject will not
be able to move via [Spec, AgrP] as this move would give rise to a clash of agreement features. The features of Agr° and Top° differ: Agr° is associated with the features of the subject (index i) while Top° carries the features of the argument (index j) in its specifier. Since the features of TopAGR do not match those of the subject, the functional head TopAGR cannot head-govern the trace of the subject\textsuperscript{34} and the trace violates the ECP (Rizzi 1997, see also Haegeman 1997).

(88a) and (88b) give the representations of (86b) and (86c) respectively.

The trace \( t_i \) and the intermediate traces in [Spec, FinP], [Spec, AgrP] and [Spec, ForceP] show the movement path of the subject. In (88a) the subject passes through [Spec, AgrP] and activates the AGR features on the head, Agr°. The notation TopAGR\textsubscript{i} serves to indicate that the features of Agr° and Top° match. As a result, the intermediate subject trace in [Spec, FinP] can be head-governed by TopAGR\textsubscript{i}. The intermediate subject trace in [Spec, FinP] itself also triggers agreement on Fin° and, hence, Fin° head-governs the subject trace in [Spec, IP]\textsuperscript{35}. There is no ECP violation, hence, the sentence is grammatical.

\textsuperscript{34} See Rizzi (1997) for relevant discussion.
\textsuperscript{35} See also Ihsane 1998: 48.
In accordance with the discussion above, in (88b) the fronted argument moves to [Spec, TopP]. Again, the assumption is that Topº is associated with agreement features and TopP is dominated by AgrP. Given that the features of the topicalized complement in [Spec, FinP], and those of Topº match, the fronting of the argument exhausts the agreement features on Topº. Consequently, the subject cannot move via the [Spec, AgrP]: if it did and if it triggered agreement on Agr, then this would give rise to a clash of agreement features, as explained above. In (85b), the features of TopAGR do not match those of the subject and it follows that the subject trace cannot be head-governed by TopAGR. This leads to the violation of the licensing clause of the ECP, hence, the sentence is ungrammatical.
As it becomes obvious from the diagrams and the discussion, the subject can cross a topicalised adjunct but cannot by-pass a preposed argument (Haegeman 1997: 258).

To apply the proposal to the derivation of the DNS pattern, Haegeman (1997) postulates that one can “insert an antecedentless non-variable empty category” in the A’-specifier position (AgrP) (Haegeman 1997: 259). As a result, we get sentences like (83) and (84b). (89a) and (89b) are the diagrams for the grammatical (83), with adjunct preposing, and the ungrammatical (84b), with argument preposing. In both ForceP has been truncated and [Spec, AgrP] hosts the null subject with AgrP being the root, i.e. the maximal projection in the representation. Haegeman (1997: 260) observes that if ForceP were projected, [Spec, ForceP] would be a potential position for an antecedent. In (89a) the null subject successfully by-passes the preposed adjunct and moves to [Spec, AgrP]. Thus, with AgrP being the highest projection, the subject sits in the specifier of the root and remains antecedentless. (Haegeman 1997: 260). In (89b) the null subject cannot cross the argument in [Spec, TopP]
due to the clash of features, as discussed above. It follows, that the preposed argument is not compatible with DNS.

Haegeman (1997) concludes that null subjects occur thanks to the truncation of ForceP. A diary null subject is “an antecedentless [non-variable] empty category in the A´-specifier of the root” where the non-overt subject can cross a preposed adjunct but cannot by-pass a preposed argument (Haegeman 1997: 233, 260). It also accounts for the fact that null subjects are illicit in embedded clauses: in such cases the matrix clause can provide a potential
antecedent for the subject in the embedded clause, and the null element will be c-commanded by a higher projection leading to the ECP (66) violation. Besides, DNS are not compatible with *wh*-questions\(^\text{36}\) (90): as discussed in Haegeman (1997: 260), “*wh*-phrases move to the CP domain and enter into an agreement relation with a head carrying the *wh*-feature. This agreement exhausts the possibilities of agreement on the relevant head just as in case of argument preposing.” The relevant functional head in this case is not Top\(^\circ\) but Foc\(^\circ\). Thus, the features on Agr\(^\circ\) clash with those on Foc\(^\circ\) giving rise to ungrammaticality. The null subject cannot cross the *wh*-phrase which leads to the violation of ECP, hence, the sentence cannot be grammatical\(^\text{37}\) (Haegeman 1997: 260-61).

(90) *When will \(\emptyset\) see you again?

2.4 The Interpretation of subjects

Based on the analysis above, Haegeman (1997: 261) observes that sentences that contain a third person null subject do not pattern exactly like those containing a corresponding personal pronoun. This contrast follows from the analysis of DNS in Section 2.3.3. Recall that, after crossing the fronted adjunct, the null subject moves to [Spec, AgrP] where it agrees with the functional head. Haegeman proposes that by virtue of the agreement, [Spec, AgrP] is a c-commanding A-position and can A-bind lower A-positions (Haegeman 1997). If a fronted adjunct contains a DP, the nominal will be A-bound by the null subject in [Spec, AgrP]. Haegeman (1997) gives examples for French and English. The English example is illustrated in (91).

(91) a. In John’s picture of Mary, she smiles.
   b. [\(\emptyset\) [In John’s picture of Mary, [*i/j smiles]]]

Haegeman observes that in (91a) the NP Mary is intended to be co-referential with the subject of the matrix clause, she, and the example is grammatical. Conversely, in (91b) the same co-reference gives rise to ungrammaticality. The null subject in [Spec, AgrP] A-binds the DP Mary and violates the Principle C of the Binding theory.

\(^{36}\) See Section 1.3.1.3.

\(^{37}\) For detailed discussion I refer the reader to Haegeman (1995).
The discussion shows “that the null subject is not strictly equivalent to the pronominal counterpart” (Haegeman, 1997: 261).

The following section focuses on the syntactic patterns of null subjects attested in diary registers and coordinate clauses in core grammar.

2.5 DNS and subject omission in coordinate clauses

This section deals with the similarity of distribution of null subjects in the second conjunct of coordinate clauses and DNS. Recall that, as noted in Section 1.1, null subjects in diaries and coordinate null subjects in core grammar demonstrate similar syntactic patterns. According to Haegeman (2007, 2011), “the striking parallelism” between the two phenomena was first highlighted by Wilder (1994, 1996). In diaries the subject can be omitted if it has a salient referent in the discourse. In coordinate clauses the subject of the second conjunct can be non-overt if it is co-referential with that of the first conjunct (Quirk et al., 1972: 555). However, this restriction is not always applicable to diaries. I will return to this point below. Examples in (92) are from Haegeman (2007: 117).

(92) a. I went home and ø wrote a few letters.
   b. ø Wrote a few letters
   c. It was half past seven and ø felt like midnight.
   d. ø Felt like midnight.

Haegeman (2007, 2011) observes that both coordinate null subjects and DNS are subject to the same constraints (Haegeman 2007: 117, 2011: 25-26). The examples provided are from Haegeman (2011: 25-26). Thus, like DNS the subject of the second conjunct cannot be non-overt:

1) in case of subject-auxiliary inversion (93a),
2) when there is a preposed argument (93b),
3) in embedded clauses (93c),
4) with a focalized constituent (93d).

(93) a. *Did you go home and did ø find anything?

(c) The first house we visited is too old. *The second house is very grand and John is sure Ø will increase in value over time.

d. *After Boston we will pass through Providence and then FINALLY Ø can begin the last leg to NYC.

Coordinate null subjects are compatible with a preposed adjunct, like DNS (93e).

(93) e. They are dedicated golf fans and for years Ø have travelled to the Open.

Returning to the co-referentiality of subjects in both conjuncts of coordinate clauses, I would like to refer to an important point made in Haegeman (2002b). Specifically, based on Becquet (2000), she observes that in some examples of the diary registers the occurrence of coordinate null subjects, which are not co-referential with the subject of the preceding clause, seems licit. (94) provides examples from Haegeman (2002b: 141).

(94) a. It is already 6.30 and Ø have to go out to Cullens for Grand Marnier soufflé ingredients. (Fielding, 83)

b. Ø Badly need water but Ø seems better to keep eyes closed. (Fielding, 68)

In (94a) the null subject of the second conjunct can be interpreted as I and it is not identical to the subject of the previous clause. In (94b) the subjects of both conjuncts are non-overt, nevertheless, it is evident that they are not co-referential. While the non-overt subject in the first conjunct can be understood as I, that of the second conjunct is the expletive it.

These examples illustrate two patterns of coordinate null subjects which are not attested in the core grammar of English. First, there is a violation of the principle of co-referentiality, second, the subject of the first conjunct is always overt in core grammar but it can be implicit in diaries as (94b) shows. An important point to be made about this discussion is that the data are from a fictional diary. In the same diary style embedded null subjects are also found, and Ihsane (1998), Haegeman and Ihsane (1999, 2001) and Haegeman (2002b) postulate that the two patterns are related, i.e. that they both illustrate a different ‘diary dialect’. Given that my analysis is based on a non-fiction style one can assume that these patterns should be non-relevant for this research. However, it will be interesting to see whether these patterns are
also attested in non-fiction diaries. I will pursue this objective in Chapter 4 which deals with the analysis of Truman’s Diary (1947). Besides, I will observe whether the theoretical proposals discussed in this chapter account for my data.

As already mentioned, in order to make a more complete idea about the distribution of DNS, I will compare my results with those of Ihsane (1998). For making the comparison more efficient I adopt her methodology in this thesis. Ihsane’s methodology and results for Virginia Woolf’s Diary (1940) are presented in the following chapter.
Chapter 3. Ihsane’s (1998) analysis: Virginia Woolf’s Diary (1940)

3.1 Introduction

In the previous chapter, Haegeman’s (1997, 2007, 2011) analysis and proposals have been presented, from which my paper takes its lead. We saw that in the data which she examined DNS are a root phenomenon, i.e. they are not attested in embedded clauses. Diary subject omission is also unavailable in wh- and yes/no questions. In addition, as Haegeman shows, with preposed material subject ellipsis is selectively possible. While the subject can be non-overt in case of adjunct preposing, subject ellipsis is not licit when there is a preposed argument. Examples (28-32) are repeated here as (95-97).

(95) Ø Like swimming.
(96) *She says that Ø likes swimming.
(97) a. *Do Ø like swimming?
    b. *Why do Ø like swimming?

As the analysis will show, the data in my corpus are totally in line with Haegeman’s proposals discussed above.

In terms of the methodology and classification criteria for the analysis, my research is modelled on Ihsane’s (1998) analysis of Virginia Woolf’s Diary (1940). Ihsane provides a detailed analysis of diary null subjects focusing on both the realization and the interpretation of the attested DNS patterns. She identifies overt and null subjects according to the clause type they are attested in and discusses the properties of the diary null subjects with respect to the verb type, verb tense, preposed constituents they occur with. In this chapter I present Ihsane’s analysis and results. Given that Ihsane’s work has not been published, sometimes I provide as much material as is appropriate for the reader’s understanding, including her tables and numbers.

Though based on Ihsane’s classification criteria and methodology, my own analysis will diverge a bit from hers in that it will also include root and coordinate clauses while Ihsane’s study excludes coordinate clauses from the analysis after giving the general figures.

But see Haegeman and Ihsane (1999, 2001) for a divergent type of usage.
For more details I refer the reader to Chapter 4 which provides the analysis of my data and presents a comparative review of Ihsane’s and my results.

The chapter is organized as follows: Section 3.2 presents the data and Ihsane’s methodology. Section 3.3 deals with her results in relation to the realization of the subject and Section 3.4 discusses the grammatical properties of null subjects she has identified. Section 3.5 summarizes Ihsane’s findings.

3.2 The data and methodology

Ihsane’s analysis is based on a 30-page corpus drawn from The Diary of Virginia Woolf (1940: 310 – 339). The purpose of the research is to identify the syntactic constraints of DNS and the ratio of occurrence of null and overt subjects. As the first part of her analysis, Ihsane classifies the finite clauses of the corpus into three groups: root clauses (98a), embedded clauses (98b), coordinate clauses (98c). All the examples presented in this chapter are Ihsane’s:

   b. I know that Virginia lives in Paris.
   c. and/but/or Virginia lives in Paris.

Ihsane further divides root clauses into: interrogatives (99a), exclamatives (99b) and “roots which are neither questions nor exclamatives” (99c) (Ihsane 1998: 12). The latter are referred to as ‘simple roots’:

(99) a. Who will be killed tonight? (The Diary of Virginia Woolf 1940: 330)
   b. How instinctive the mother’s reaction is! (idem: 333)
   c. He’s agile. (idem: 333)

The interrogative clauses are classified into two types: yes/no questions (100a) and wh-questions (100b):

(100) a. Is it thunder? (The Diary of Virginia Woolf 1940: 311)
b. Who lived there? (idem: 316)

Ihsane excludes imperatives from the study.

As far as the methodology is concerned, Ihsane presents a “set of rules” she has taken into account during the data analysis. First, given the punctuation is inconsistent in Woolf’s Diary, Ihsane considers colons, dashes and semi-colons preceding a finite verb as full stops (101):

(101) a. Morgan lost his spectacles – found them in a book. (The Diary of Virginia Woolf 1940: 315)

b. The air saws; the wasps drone; (idem: 313)

Ihsane classifies sentences starting with how as exclamatives despite the absence of the exclamation mark (102):

(102) a. How I should like to write poetry all day long – (The Diary of Virginia Woolf 1940: 312)

b. How free, how peaceful we are. (idem: 328)

Thirdly, Ihsane reports a case when she counts a sentence with a question mark as a simple root, not as an interrogative (103), explaining that “there is nothing in the context of the sentence which allows us to conclude that Woolf is asking a question” (Ihsane 1998: 15):

(103) I must add that today is as hot as August: walked on downs; heard gunfire at Dover? – shelling Calais; summer clothes; L. cleaning beds. Too hot for fire. Mist rising, must black out. (The Diary of Virginia Woolf 1940: 332)

Furthermore, Ihsane clarifies that the clauses following the verbs say, tell, admit, feel, suppose, think… have been considered as embedded clauses (104a) even when the complementizer that is not present. However, when what is said, thought, felt… precedes the matrix clause with e.g. he said, she thought, I felt… both are counted as simple roots (104b):

(104) a. I said he was a prostitute. (The Diary of Virginia Woolf 1940: 314)

b. We’ve not had our raid yet, we say. (idem: 313)
A clause introduced by *and, but or* is categorized as a coordinate even if it follows a full stop (105a) or marks the beginning of a new paragraph (105b):

(105) a. It’s done. And I’m writing PH. (The Diary of Virginia Woolf 1940: 311)
    b. And then Morgan slightly damped me. (idem: 310)

Given Woolf’s inconsistency in punctuation, it becomes clear why Ihsane chooses to consider such cases as coordinate clauses. The sentences, in which the coordinate conjunction is followed by a subordinate conjunction, are classified as embedded clauses (106):

(106) No more reviews; & if I had solitude – no man driving stakes digging fresh gun emplacements & no neighbours, doubtless I cd. expend & soar – into PH. Into Coleridge; (The Diary of Virginia Woolf 1940: 310)

Finally, Ihsane has left out two types of infinitival verbs from her analysis. They are shown in (107), the first one being the instances like (107a) and the second one questions with *Why + infinitive…?* and *Why not + infinitive…?* (107b,c):

(107) a. so *pray* God the Church bells don’t ring tonight. (The Diary of Virginia Woolf 1940: 321)
    b. *Why come* to the top when I suffer so in that light? (idem: 339)
    c. *Why not write* the real life for the Memoir Club? (idem: 314)

After presenting the classification criteria and the methodology adopted, Ihsane gives the overall numbers, which are reported in the next section.

### 3.3 The realization of the subject

First, as reported above, Ihsane classifies all the finite clauses in her corpus into three main types: roots, embedded and coordinate clauses. Table 1 shows that in her 30-page extract, Ihsane identifies 1421 clauses with finite verbs. 989 (69.60%) of these are root clauses, 287
(20.20%) are embedded clauses and 145 (10.20%) are coordinate clauses. I reproduce Ihsane’s tables below\textsuperscript{39}.

**Table 1: Finite verbs**

<table>
<thead>
<tr>
<th>Type</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coord.</td>
<td>145</td>
<td>10.20%</td>
</tr>
<tr>
<td>Embedded</td>
<td>287</td>
<td>20.20%</td>
</tr>
<tr>
<td>Roots</td>
<td>989</td>
<td>69.60%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1421</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2 illustrates the proportions of root clauses in terms of three categories: simple roots, exclamatives and interrogatives. Among the total of 989 roots there are 938 (94.84%) simple roots, 6 (0.61%) exclamatives and 45 (4.55%) questions.

**Table 2. Root clauses**

<table>
<thead>
<tr>
<th>Type</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple roots</td>
<td>938</td>
<td>94.84%</td>
</tr>
<tr>
<td>Exclamatives</td>
<td>6</td>
<td>0.61%</td>
</tr>
<tr>
<td>Questions</td>
<td>45</td>
<td>4.55%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>989</td>
<td>100%</td>
</tr>
</tbody>
</table>

Questions are classified into yes/no and wh- questions. The figures for these are presented in Table 3. Among them 31 (68.89%) yes/no questions and 14 (31.11%) wh-questions have been found.

\textsuperscript{39} Ihsane is inconsistent in presenting the total of “100%” which is mostly given as “100.00%” and once as “100%”. In all the tables I report it as 100%.
Table 3. Questions

<table>
<thead>
<tr>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes/no questions</td>
<td>31</td>
</tr>
<tr>
<td>Wh- questions</td>
<td>14</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>45</strong></td>
</tr>
</tbody>
</table>

Finally, Ihsane presents the proportions of null and overt subjects according to the clause types they are attested in. Her results are reported in Table 4. Among the 1421 finite clauses 1241 (87.33%) have an overt subject and in the remaining 180 (12.67%) the subject is omitted. Out of the 938 simple roots, 827 (88.17%) have an overt subject and 111 (11.83%) display a null subject pattern. Among 145 coordinate clauses Ihsane identifies 76 (52.41%) cases with an overt subject and 69 (47.59%) cases with an omitted subject. There are no instances of subject ellipsis in yes/no and wh- questions, exclamatives and embedded clauses. Ihsane concludes that these clause types are “irrelevant to determine the distribution” of DNS (Ihsane 1998: 20).

Table 4. Overt and null subjects

<table>
<thead>
<tr>
<th>Overt subjects</th>
<th>% Total</th>
<th>Null subjects</th>
<th>% Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple roots</td>
<td>827</td>
<td>111</td>
<td>11.83%</td>
<td>938</td>
</tr>
<tr>
<td>Yes/no questions</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Wh- questions</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Exclamatives</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Embedded</td>
<td>287</td>
<td>0</td>
<td>0</td>
<td>287</td>
</tr>
<tr>
<td>Coord.</td>
<td>76</td>
<td>69</td>
<td>47.59%</td>
<td>145</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1241</strong></td>
<td><strong>180</strong></td>
<td><strong>12.67%</strong></td>
<td><strong>1421</strong></td>
</tr>
</tbody>
</table>

So, as the analysis shows, null subjects occur in two clause types: simple roots and coordinate clauses. In her discussion of null subjects Ihsane only considers root clauses.

Ihsane claims that null subjects are not “uniformly distributed in the text” (Ihsane 1998:20). However, this finding is difficult to assess as she reports numbers for only two entries (Friday 13 September and Sunday 29 September), which does not give a good picture.
of the variation in the distribution of null subjects in the corpus. Ihsane notes that the 39-line entry for 13 September has 15 cases of subject omission while in the first 18-line paragraph of the entry for September 29 there are 19 null subjects. Thus, 30.63% of the 111 attested null subjects are found in these two paragraphs. Ihsane doesn’t pursue this point any further.

To summarize, the crucial finding of Ihsane’s work is that DNS are a root phenomenon, confirming Haegeman’s (1997) conclusion. The phenomenon is attested in 111 (11.83%) simple root clauses. No cases of subject omission have been found in yes/no and wh- questions, exclamatives and embedded clauses.

Observe that though the figures above give us some indication as to the overall ratio of the finite sentences with null subjects in relation to all finite sentences with overt subjects, the figures are very broad and do not allow us to assess precisely the distribution of null subjects. The major problem – and one that will be relevant also in the next sections – is that Ihsane groups all sentences with overt subjects together, i.e. both sentences with pronominal subjects such as I, he, she, they as well as the sentences with lexical subjects such as Mary, Virginia, the doctor, etc. Now it is clear that subject omission is not readily available with a lexical NP as a subject, in particular if that NP introduces the referent in the discourse. For instance it would not be readily expected that the subject Morgan would be dropped in (101a) above. As shown in that example, the introduction of the referent by means of the DP Morgan establishes an entity in the discourse. Subsequent reference to that entity would be with a pronoun, and it is the pronoun that can be omitted. So in order to fully assess the realization of subjects it seems that we should not only have been provided with the total number of finite sentences with overt subject, but it would also have been important to be able to see the proportion of pronominal subjects. This information is not given by Ihsane.

3.4 Grammatical properties of null subjects

3.4.1 Referential and expletive subjects

Ihsane starts the analysis of the null subjects attested in the corpus with classifying them into two types: referential (108a) and expletive (108b).

    b. Seems there will be an election soon. (idem)
There is only 1 (0.90%) expletive (109) among the 111 null subjects:

(109) Seems impossible. (The Diary of Virginia Woolf 1940: 332)

Table 5 illustrates the proportions of the referential and expletive null subjects in the corpus.

**Table 5. Referential and expletive null subjects**

<table>
<thead>
<tr>
<th></th>
<th>Referential</th>
<th>Expletive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS</td>
<td>110</td>
<td>1</td>
<td>111</td>
</tr>
<tr>
<td>%</td>
<td>99.10%</td>
<td>0.90%</td>
<td>100%</td>
</tr>
</tbody>
</table>

On the basis of this table, Ihsane concludes that null expletive subjects are rare, though, on the basis of the example that she has identified and also using another diary example (110) from Haegeman (1997) and some examples of null subjects in spoken English (111) given in Thrasher (1977), the author points out that expletive subjects may be null.

(111) a. Isn’t much we can do about it. (Thrasher 1977: 44)
        b. Won’t be too difficult to reconstruct his argument. (Thrasher 1977: 44)

However, Ihsane’s conclusion is not fully supported by the data. As was the case with respect to the overall ratio of overt subjects to null subjects, the figures are not precise enough. In particular, the only information that is given concerns the ratio of the expletive subjects among null subjects. Ihsane fails to discuss the frequency of overt expletive subjects. We can only really assess whether the unique example of an expletive null subject can be labelled rare if we can compare it with the overt subjects and if we can compare the ratio of null/overt expletive subjects with that of null/overt third person pronouns. I will pursue this point in my own analysis to see what the conclusions are and whether the null expletives are indeed rare in diaries.
3.4.2 The person and number of null subjects

Ihsane then pursues the studies of the 110 non-overt referential subjects. Among the null subjects she identifies the grammatical categories of person and number of the implicit subjects. Two main criteria are taken into consideration while identifying the null subjects: the inflectional ending –(e)s occurring in the present tense to denote the third person singular in English and second, the context. Ihsane identifies 17 third person singular null subjects, 15 out of which are attested with a present-tense verb, the 2 others are followed by an auxiliary in present tense: *may* and *can*. In other words, Ihsane mainly relies on the context to identify the 95 null subjects.

In trying to recover the non-overt subjects from the context, Ihsane observes the position of the antecedent with regard to the implicit subject. For this purpose, four context categories are defined:

Category 1 includes the first two strings of words preceding the null subject, Category 2 the first two strings of words following the omitted subject, Category 3 represents the paragraph in which the implicit subject occurs and Category 4 consists of the more general context or meaning beyond the first paragraph (Ihsane 1998: 22-23).

Though this system of classification is sufficient to give a general view of the position of the antecedent of the omitted subjects in the corpus, the classification of the “strings of words” does not seem to be consistent and reliable enough. In one case these are represented by a root clause, another time by a coordinate clause, in a third case by a material separated by punctuation marks, like *Bath* or *Out to Lodge* (Ihsane 1998: 23). Ihsane calls the system “approximate”, too. What I would suggest is that each finite clause be considered as a string of words for the simple reason that the analysis focuses on null subjects attested in finite clauses. To consider *Out to Lodge* as a relevant string of words, the cases of *subject + auxiliary* omission should also be taken into consideration because this material can be reconstructed as meaning *I am/was out to Lodge* or *They are/were out to Lodge*, etc., thus, also introducing potential antecedents. However, we do not know whether the analysis includes the cases of *subject + auxiliary* omission: no details about this are reported in the paper.

---

40 Observe that Ihsane includes the example with the attested expletive null subject in her table below (Table 6).
To continue with the study, it should be noted that in cases when the implicit subject precedes another null subject, Ihsane counts the first empty category as an antecedent for the second empty category. After the detailed analysis, Ihsane concludes that non-overt subjects are always recoverable and the subjects are left out only when they are redundant. The details about the distribution of null subjects according to the context categories are presented in Table 6, which provides information on the realization of the implicit referential subjects, identifying the grammatical categories of person and number of the latter. The results reveal that 53 (47.75%) out of the 111 subjects refer to the narrator. There are no omitted subjects which can be identified as you. There are 36 (32.43%) instances of the third person singular null subjects. The first person plural occurs 17 times (15.32%). The category which is least attested is the third person plural with 5 cases in total (4.50%).

As far as the context categories are concerned, in most cases the antecedent can be found in “the first two strings of words preceding the null subject” (Category 1). There are 83 identified cases. Then, we get 17 instances for “the paragraph in which the implicit subject occurs” (Category 3). The “first two strings of words following the omitted subject” (Category 2) and the general context (Category 4) are identified in 6 and 5 cases respectively.

Table 6. Referential null subjects

<table>
<thead>
<tr>
<th>Category</th>
<th>Category 1</th>
<th>Category 2</th>
<th>Category 3</th>
<th>Category 4</th>
<th>Total</th>
<th>% out of 111</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg⁴¹</td>
<td>41</td>
<td>32</td>
<td>13</td>
<td>3</td>
<td>53</td>
<td>47.75%</td>
</tr>
<tr>
<td>2sg</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3sg M</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>22.52%</td>
</tr>
<tr>
<td>3sg F</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>6.31%</td>
</tr>
<tr>
<td>3sg N</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>3.60%</td>
</tr>
<tr>
<td>1pl</td>
<td>12</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>17</td>
<td>15.32%</td>
</tr>
<tr>
<td>2pl</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3pl</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4.50%</td>
</tr>
<tr>
<td></td>
<td>83</td>
<td>6</td>
<td>17</td>
<td>5</td>
<td>111</td>
<td>100%</td>
</tr>
</tbody>
</table>

⁴¹ Note that sg means singular, pl plural, M masculine, F feminine and N neutral.
To sum up, the figures reveal that in the corpus Ihsane examined, nearly the half of the null subjects refer to I. Third person plural null subjects are attested least frequently. As to the context categories, in more that two thirds of the cases the antecedent can be found in the first two strings of words preceding the omitted subject. Because the antecedent is closely situated in most of the cases, Ihsane concludes that the null subjects can, thus, be easily recovered and “the overt filling of the overt subject is therefore not essential for the understanding of the sentence”. (Ihsane 1998: 26)

Once again, though, I need to point out that the figures given here raise additional questions. Since we have no information concerning the nature of the overt subjects, we cannot assess, for instance, how often a first person pronoun is spelt out and how often it is non-overt. Similarly, Ihsane shows that there are no null subjects to be interpreted as second person pronouns, but we have no information as to the availability of second person pronouns in the corpus. If these are rare or non-existent, too, then it is not surprising that their null counterparts do not arise. The same remark extends to third person subjects: we can only assess the relative rareness of, say, third person plural null subjects, once we know how many third person plural pronouns are present in the corpus. If these are rare, too, then it might well be the case that though there are few cases of third person plural null subjects, these are not rare in relation to the overt pronouns of the same category.

In my own research I will, therefore, refine the figures, and compare the distribution of null subjects not just with the overall totals of overt subjects, but I will consider partial comparisons between subjects of the same type.

The following section deals with the type and tense of the verbs null subjects are attested with.

3.4.3 The verb type and verb tense

In Ihsane’s work, verbs are classified into 3 types: lexical verbs, auxiliaries, copula be. In addition, the tense of the verb (present or past) the implicit subjects occur with is studied. Ihsane also examines whether the null subjects are attested with positive or negative verbs.

Lexical verbs are verbs like go, sleep, write, eat, know, etc. The class of auxiliaries consists of the following items: have, may, must, can, shall, will, be, should, need, dare, do-support. The distinction between the copula be and auxiliary be is determined by the type of complement they take: when followed by a VP, be is classified as an auxiliary (112a), when
followed by a NP, PP, AP, it is counted as a copula (112b) (Ihsane 1998: 26). The examples are mine:

(112) a. He is arriving on Sunday.
    b. He is a nice person.

The proportions of the tense for each verb type are presented in Table 7. Accordingly, among the 111 null subjects 47 (42.34%) are attested with a verb in present tense. Most of these, 37, are lexical verbs, there are 9 auxiliaries and there is 1 occurrence of copula be. As to the past tense, 59 out of the total of 64 are lexical verbs, 4 are auxiliaries and there is 1 copula be. As we can see, for both tenses, the majority of null subjects occur with lexical verbs, about two thirds of which are in past tense. In absolute terms, fewer null subjects occur with an auxiliary, two thirds of these auxiliaries are in present tense. Copula be occurs least frequently with null subjects and is equally distributed for each tense. Thus, Ihsane concludes that verb tense doesn’t play a significant role in the distribution of null subjects.

Table 7. The verb tense and verb type

<table>
<thead>
<tr>
<th></th>
<th>Present</th>
<th>% Total</th>
<th>Past</th>
<th>% Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical V</td>
<td>37</td>
<td>38.54%</td>
<td>59</td>
<td>61.46%</td>
<td>96</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>9</td>
<td>69.23%</td>
<td>4</td>
<td>30.77%</td>
<td>13</td>
</tr>
<tr>
<td>Copula be</td>
<td>1</td>
<td>50%</td>
<td>1</td>
<td>50%</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>42.34%</td>
<td>64</td>
<td>57.66%</td>
<td>111</td>
</tr>
</tbody>
</table>

I point out that the same reservations that arose before are to be stated here. While we do have absolute figures for the numbers of null subjects with lexical verbs and with auxiliaries, these figures cannot allow us to conclude the relative frequency of subject omission according to the verb type because we do not have any information as to the overall totals of lexical verbs and auxiliaries with overt (pronominal) subjects. This is important with respect to auxiliaries since we know that the sequence subject+auxiliary itself may be deleted (113) which reduces the number of the sentences with an auxiliary. We would need to know the ratio of lexical verbs/auxiliaries with overt pronominal subjects in order to be able to assess the relative weight of the figures in Table 7.
(113) a. A cold day.
   b. Feeling exhausted.

Next, Ihsane presents the numbers for positive and negative verbs, again according to the verb types. As Table 8 reveals, the large majority (102 out of 111) of sentences with a null subject are positive. All the lexical verbs and both copulas are attested in positive sentences. As to the auxiliaries, only 4 (nearly 31%) out of 13 sentences are positive. Given the small numbers, Ihsane draws no firm conclusions. The point made is that “lexical verbs and copula be seem to be more often in positive sentences containing a null subject” (Ihsane 1998: 28).

Table 8. Positive/Negative

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>% Total</th>
<th>Negative</th>
<th>% Total</th>
<th>Total</th>
<th>% 111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical V</td>
<td>96</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>96</td>
<td>86.49%</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>4</td>
<td>30.77%</td>
<td>9</td>
<td>69%</td>
<td>13</td>
<td>11.71%</td>
</tr>
<tr>
<td>Copula be</td>
<td>2</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>2</td>
<td>1.80%</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td></td>
<td>9</td>
<td></td>
<td>111</td>
<td>100%</td>
</tr>
</tbody>
</table>

Again, like in the previous section, all the 111 null subjects, including the only expletive subject, have been counted.

As before, in the absence of information on the relative ratio of positive/negative sentences with overt (pronominal) subjects, we cannot establish any firm conclusions as to the distribution of null subjects in relation to clausal polarity.

3.4.4 Material to the left of the subject

This section examines whether the null subjects are the first elements in the clause or if there are any constituents to their left.

Pursuing the question about what kind of constituent to the left of the canonical subject position might be compatible with subject omission, Ihsane looks at four patterns:
adjuncts (114a), complements (114b), *wh*-elements (114c), no preposed material (114d). The examples are given as in Ihsane (1998: 29).

(114) a. Yesterday Virginia went to London.
   b. Her books, she will never throw away.
   c. How can Virginia read so much?
   d. Virginia went to London.

Among the instances of subject omission, the majority (100) have not material to the left of the subject position. There are 11 instances with a preposed adjunct (8 AdvP and 3 PPs) and there are no instances of preposed complements.

Table 9 summarizes the findings and also recapitulates the results for interrogative contexts. Thus, among the 111 clauses only 11 (9.91%) occur with a preposed constituent. All of them are adjuncts. The data are in line with Haegeman’s analysis on argument/adjunct asymmetry presented in the previous chapter.

Table 9. Preposing

<table>
<thead>
<tr>
<th>Null subject…</th>
<th>Preposing</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>With no preposed material</td>
<td>100</td>
<td>90.09%</td>
</tr>
<tr>
<td>With a preposed adjunct</td>
<td>11</td>
<td>9.91%</td>
</tr>
<tr>
<td>With a preposed complement</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>With a preposed <em>wh</em>- element</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>111</td>
<td>100%</td>
</tr>
</tbody>
</table>

Given the insufficient data Ihsane refrains from drawing any firm conclusions.

As before I would like to point out that the figures remain incomplete, and that even leaving aside the problem of the low numbers, we need to take into account that there is no information on the occurrence of overt subjects in the contexts with preposing. While we know that there are 6 instances of exclamatives with overt subjects and 45 instances of interrogatives (see Table 4), we do not know how many of the relevant overt subjects are pronominal.
As for preposing: we have no information concerning the relative frequency of adjunct preposing and complement preposing with overt pronominal subjects, so there is no way we can assess the significance of the figures above, regardless of the fact that they are low.

3.5 Conclusion

To conclude, on the basis of her material Ihsane’s main conclusions are that:

- subject omission in diaries is a root phenomenon (no null subjects in yes/no or wh-questions found)
- no null subjects with a preposed complement attested \textbf{BUT:}
- attested preposed adjuncts with null subjects $\rightarrow$ adjunct/complement asymmetry.

However, as I have pointed out repeatedly throughout the chapter, she only discusses the properties of the sentences with null subjects and apart from the cases of interrogative and exclamative clauses, she fails to compare the distribution of null subjects with that of overt pronominal subjects. Where she does look at the ratio of overt subjects vs. null subjects she does not take into account that only pronominal subjects can alternate with null subjects.

In the following chapter my analysis of Truman’s Diary (1947) is presented. In most of the cases throughout my research I follow Ihsane’s classification criteria and methodology as given in this chapter. However, my own work will depart from Ihsane’s in a number of ways. First, I will examine not only the realization of null subjects but I will also examine the overt subjects identified in the corpus to give a more clear idea on the distribution of null subjects and the syntactic constraints which may favor or disfavor subject omission in diaries.

In addition, my analysis will also include coordinate clauses. In particular, given that the core grammar of English allows subject ellipsis in second conjuncts, I will try to assess whether the omission of a subject in the second conjunct of coordinate clauses in diaries is the same phenomenon as in core grammar.
Chapter 4. The analysis of the Truman Diary (1947)

4.1. Introduction

After presenting Haegeman’s (1997, 2007, 2011) theoretical proposals, on which this thesis is based, and Ihsane’s (1998) analysis of Virginia Woolf’s Diary (1940), on which my research is modeled, the next logical step would be to examine my corpus data based on Harry S. Truman 1947 Diary. As already mentioned above, the pursued aim of my analysis is to identify the features of null subjects through examination of the empirical data and observe whether any so-far-unidentified patterns of null subjects occur in diaries besides the syntactic constraints of subject ellipsis already defined. If yes, what are they and how are they realized? Although DNS have both syntactic and pragmatic foundations, as concluded in Chapter 1, I mainly focus on the syntactic factors.

The data are analyzed according to the clause types the null subjects are attested in. To be able to compare the distribution of null subjects with that of overt pronominal subjects, for each clause type and grammatical pattern the proportions of both overt and null subjects are presented. Then I compare the results of my corpus with those discussed in Ihsane (1998), where a comparable study is made of a corpus based on the diary of Virginia Woolf. The attested patterns drawn from the corpus are supplemented, where relevant, with additional examples in order to illustrate the occurring and non-occurring patterns.

The identified null and overt subjects will be classified into two groups: referential subjects and expletive subjects. I will examine whether one type has a privilege over the other with respect to the frequency of omission. My discussion further focuses only on referential subjects. I investigate the grammatical properties of null and overt subjects according to the clause type they occur in.

After discussing the distribution of the patterns for each clause type I focus on coordinate clauses and compare the constraints of the coordinate null subjects attested in the corpus with the patterns which are allowed in core grammar. As already mentioned above, this clause type is of interest for my research for two reasons: (i) subject omission in the second conjunct is licit in English core grammar, (ii) null subjects in the second conjunct of coordinate clauses and DNS have similar distribution constraints, as discussed in Section 2.5.

---

42 See previous chapters for the identified syntactic constraints.
43 See Chapter 3 for discussion of Ihsane’s (1998) results.
44 For classification of clause types see Section 4.2.
The pursued objective is to identify the similarities and differences between the coordinate null subjects occurring in Standard English and in diaries through examination of empirical data.

The following section provides information on the corpus and motivates my choice for this particular diary.

4.1.1 The corpus

My research is based on an eleven-page (5,469-word) corpus consisting of all the entries for *Harry S. Truman 1947 Diary*, which can be found at [http://www.trumanlibrary.org/diary/transcript.htm](http://www.trumanlibrary.org/diary/transcript.htm). The choice of this particular diary was made taking into account several criteria: first, given my aim to compare the syntactic constraints identified in the US and UK\(^{45}\) diaries to see whether the two variants of English differ in terms of the distribution of DNS\(^{46}\), I have chosen a diary written by an American author.

The second criterion I have taken into consideration when choosing Truman’s Diary, is the assumption of ‘the expected audience’. Specifically, I suppose that unlike most diaries which are written to oneself with no expected or targeted reader, Truman might have expected his diary to be read later\(^{47}\) and, thus, might have tried to be less personal and less ambiguous which would result in a lower rate of subject omission. The same assumption might be extended to Virginia Woolf’s Diary (1940) given that she was famous, too, and might have expected that her diary would be read later. These shared criteria will contribute to the efficiency of the comparative analysis and will allow me to draw more precise conclusions on the distribution patterns and the rate of subject deletion\(^{48}\).

The chapter is organized as follows: Section 4.2 presents the general data and the classification of the clause types. Section 4.2.1 elaborates on the methodology and the criteria

---

\(^{45}\) Ihsane (1998) provides a detailed analysis of a UK diary.

\(^{46}\) See also Preface.

\(^{47}\) Truman uses the second person pronoun two times and one of the possible interpretations on the use of *you* can be that in this way Truman addresses the reader. See Section 4.4.2.1 for interpretation of *you*. Besides, the attested 7 examples of imperative clauses might suggest that there is an implied addressee. See Section 4.2 for imperative clauses.

\(^{48}\) Another criterion to take into account for efficient comparison of the data would be choosing writers of the same gender. The suggestion is based on the difference of language use by men and women (Mesthrie, 2000: 216-247). I am not sure whether these differences can anyhow affect the DNS distribution. This could be an avenue for future work.
taken into account for the classification of the data. Section 4.3 demonstrates the general figures of null and overt subjects according to different clause types. Section 4.4 provides a more fine-grained analysis of the distribution of overt and null subjects in the corpus concentrating on the syntactic properties of overt and implicit subjects. Section 4.5 examines what kind of constituent to the left of the canonical subject position might be compatible with DNS. Section 4.6 discusses the similarities and differences between the coordinate null subjects in Standard English and in diaries. Section 4.7 sheds light on the discourse properties of the attested overt and null subjects. Section 4.8 concludes the chapter.

4.2 The data

To study the constraints of DNS as well as to assess the occurrence and the frequency of null and overt subjects attested in the corpus, first I identify the type of clauses the implicit and overt subjects appear in. For this purpose, like Ihscane, I have classified all the finite clauses in my corpus into three main groups: root clauses (or ‘roots’ for short), as illustrated in (115a), coordinate clauses as in (115b), embedded clauses as in (115c).

(115) a. They study linguistics at university.
    b. and/but/or they study linguistics at university.
    c. They told me that they study linguistics at university.

I have separated out coordinate clauses coordinated with a root clause (i) from those coordinated with an embedded clause (ii). The third category I have identified is the single coordinate clause (iii) which lacks the first conjunct and begins with a coordinator, like but, and, or.

(109b) (i) He will check the data and she will report back to me.
    (ii) I think that he will check the data and (that) she will report back to me.
    (iii) And/But/Or she will report back to you.

As it has been mentioned throughout the thesis, the reason why I keep coordinate clauses separate is that the subject ellipsis in coordinate clauses is fully grammatical in Standard
English. Besides, DNS and the second conjunct null subjects in core grammar have similar distribution patterns as already discussed in Section 2.5. The subject of the second conjunct of a coordinate clause can be omitted if it is co-referential with the subject of the first conjunct. This is illustrated in (116): in (116a) the root clause *I have returned from Spain* is coordinated with a second root clause *will visit you tomorrow will visit you tomorrow*, in which the subject *I* can be omitted as it is co-referential with the subject of the preceding conjunct. The same type of ellipsis in coordination is illustrated for embedded clauses in (116b):

(116) a. I have returned from Spain and (I) will visit you tomorrow.
   
b. I told him that I had returned from Spain and (I) would visit him the following day.

Observe, however, that in (116c) *I* cannot be omitted because it would lead to a *that*-trace filter violation (Chomsky and Lasnik 1977): the overt complementizer *that* “is inert for head-government” and the subject trace remains un governe d (Haegeman, 1997: 251). Conversely, (116b) allows subject ellipsis because the complementizer *that* is non-overt and the subject trace can be head-governed (Quirk et al., 1972: 556). It will be interesting to observe whether there are any instances like (116c) in the corpus.

(116) c. I told him that I had returned from Spain and that *(I) would visit him the following day.

Root clauses fall into three categories: interrogative clauses/questions (117a), exclamative clauses (117b), declarative clauses (clauses functioning neither as questions nor as exclamations) (117c). The latter clause type is also referred to as a ‘simple root’ (Ilhsane, 1998: 12).

(117) a. What are they waiting for?
   
b. How sensitive she is!
   
c. They are waiting for Mary.

49 For subject omission in coordinate clauses in Standard English see Quirk et al. (1972: 555-56, 574-76)

50 See also Section 2.3.1.

51 See also Section 2.3.1.
Furthermore, I divide questions into: yes/no questions (118a) and wh-questions (118b).

(118) a. Did I hurt her with my words?
   b. When did she write the letter?

I will not include imperatives\(^{52}\) in my discussion as in general and quite independently of register, imperatives in English (as in many other languages) typically have a non-overt subject. In my corpus I have found 7 imperative clauses which are reported in (119). For clarity, the clauses are given in the context they appear in and are marked in bold. I will not discuss them as they do not demonstrate any specific features which are not attested in Standard English. What seems relevant with respect to these clauses is that (119a, b, e) might explain my assumption with respect to the ‘expected audience’ expressed above. Given that imperatives typically address an interlocutor who is characteristically absent in the diary one can assume that Truman addresses a reader who might read his diary later. (119c, d) suggest an alternative explanation: Truman might have addressed himself or a specific person who would have access to the schedule or the guest list he implies. Both examples are parenthetical sentences\(^{53}\) and they provide additional information/further clarification rather than denote a command/request. Given that my aim is to examine the syntactic features of DNS I will not go into this here.

(119) a. Maybe there was something on both sides in this situation. It is a pity a great man has to have progeny! Look at Churchill's. Remember Lincoln's and Grant's. Even in collateral branches Washington's wasn't so good-and Teddy Roosevelt's are terrible. (Truman Diary, 1947, 3 Jan.)
   b. Some of the crackpots will in all probability yell their heads off-but let 'em yell! (idem, 3 Jan.)
   c. Had quite a day. (Look at schedule for this day, the day before and the day before that.) The Crown Prince of Arabia with his retinue and the Minister from his country to ours came in with the Secretary of State (Mr. Byrnes) and discussed Mid East Affairs at some length. (idem, 16 Jan.)
   d. They afterwards came to lunch. It was a gala affair. See guest list. (idem, 16 Jan.)

\(^{52}\) See also Section 1.1.
\(^{53}\) In the diary entry (119c) appears in parentheses but (119d) does not.
e. The Jews, I find are very, very selfish. They care not how many Estonians, Latvians, Finns, Poles, Yugoslavs or Greeks get murdered or mistreated as D[isplaced] P[ersons] as long as the Jews get special treatment. Yet when they have power, physical, financial or political neither Hitler nor Stalin has anything on them for cruelty or mistreatment to the under dog. **Put an underdog on top** and it makes no difference whether his name is Russian, Jewish, Negro, Management, Labor, Mormon, Baptist he goes haywire. I've found very, very few who remember their past condition when prosperity comes. **Look at the Congressional attitude on D[isplaced] P[ersons]**-and they all come from D[isplaced] P[ersons]. (Truman Diary, 1947, 21 July)

Thus, the classification of the clauses with attested null subjects will be based on the following clause types:

1. Root clauses
   a. simple roots
   b. interrogatives
      i. yes/no questions
      ii. wh-questions
   c. exclamatives
2. Embedded clauses
3. Coordinate clauses
   i) coordinated with a root clause
   ii) coordinated with an embedded clause
   iii) single coordinate clause

4.2.1 The classification criteria

Before presenting the figures which show the distribution of subjects in finite clauses in my material, it would be appropriate to mention several classification criteria which have been taken into account.
1. The first point which requires clarification is punctuation. Like it is the case with V. Woolf’s Diary, as mentioned by Ihsane (1998: 14), Truman shows inconsistency with respect to punctuation marks, too. I will discuss the cases which are essential for my data.

   a) *dash* – This punctuation mark is widely used throughout the diary. In all the cases when a finite clause follows a dash, I have treated this punctuation mark on a par with a full stop. Among all these cases I would like to single out the 6 clauses in which a dash is followed by a coordinate clause with an overt coordinator. I have classified all these cases as a single coordinate clause because they are preceded by a dash (treated as a full stop, as noted above). If there were no overt coordinators present, I would have classified them as roots. All of them have an overt subject. (120) gives some examples, (120 a, b, c) illustrate examples of the single coordinate clauses which follow a dash.

   (120) a. Had a most pleasant evening-and so did everyone, apparently. (Truman Diary, 1947, 1 Jan.)
   b. Well only my Secretarial Staff knew of it-and they had known since April 19, 1946! (idem, 7 Jan.)
   c. I'm sure he'll regret it-and I know I do. (idem, 1947, 8 Jan.)
   d. I've read thousands of messages from all over the world in the White House study and I can shed tears as I please-no one's looking. (idem, 28 July)
   e. Had all the cadets lined up and the Foreign Minister and the Commandant of the Cadets wept-so did news men and photographers. (idem, 4 March)

   b) *exclamation mark* – In the corpus I have found 3 exclamative clauses, there are no interrogative clauses. Among them only one of the exclamative clauses occurs with an exclamation mark, the other two are followed by a full stop. Following Huddleston and Pullum (2002: 918-19), the presence of exclamative phrases *how* and *what* (121a, b) have been taken into account for classifying these clauses as exclamatives. (121c) is ambiguous because it can be classified both as an interrogative and an exclamative; for clarity I have given the clause in the context it appears in and have highlighted it in bold. Given the grammatical structure of the clause prima facie it seems logical to classify it as an

---

54 Note that there is one more instance when an imperative introduced by *but* follows the dash (i):
   (i) Some of the crackpots will in all probability yell their heads off-but let 'em yell! (Truman Diary, 1947, 3 Jan.)
As already mentioned, I do not examine imperative clauses and haven’t included this clause into my analysis.
55 For the other two clauses I haven’t given the context as their interpretation is clear.
interrogative: a subject-auxiliary inversion is typical of interrogative clauses while in exclamatives the subject precedes the predicator (Huddleston and Pullum 2002: 920). “However, subject-auxiliary inversion is available as an option in exclamatives, though it is relatively infrequent and characteristic of fairly literary style” (Huddleston and Pullum, 2002: 920). Taking into account the context and the observation above that subject-auxiliary inversion is compatible with exclamatives, I have classified the clause as an exclamative.

(121) a. How we'll miss Mrs. Patterson! (Truman Diary, 1947, 23 July)
    b. What a job he can do there. (idem, 25 July)
    c. Doc tell's [sic] me I have Cardiac Asthma! Ain[']t that hell. Well it makes no diff[erence,] will go on as before. I've sworn him to secrecy! So What! (idem, 7 March)

In cases when an exclamation mark is used after a declarative (122) I have treated it as equal to a full stop since these clauses cannot be considered exclamatives only because of the presence of the exclamation mark. Rather, according to Beijer (2002), they are “expressive/emotional utterances”, and should be differentiated from exclamative clauses. I will not discuss this distinction here and refer the reader to Beijer (2002) for relevant analysis.

(122) a. I wonder! (Truman Diary, 1947, 3 Jan.)
    b. Haven't had a more pleasant week end since moving into the great white jail, known as the White House! (idem, 5 July)

2. Like Ihsane (1998: 15), when a clause follows a verb such as say, tell, hope, think… I classify it as an embedded clause (123a). Conversely, when the clause containing the abovementioned verbs is preceded by what is said/thought…, both are classified as roots (123b).

(123) a. I said (that) I hadn’t heard from them for a while.
    b. I haven’t heard from them for a while, I said.

3. I have not included the null subjects in contexts such as those in (124).

(124) Tired in the afternoon.
(124) may be interpreted as a finite clause: *I was tired in the afternoon* and would represent a case of subject and auxiliary omission. However, in diary style writing examples can be found which present simply the combination of a subject and a predicate, without an overt auxiliary (125):

(125) Leonard sick.

This means that (124) might also be seen as an instance of subject omission on the basis of a subject-predicate combination, in which case the omission of the subject might be viewed as related to the omission of the copula. Though they are of interest, I will not examine such cases here.\(^\text{56}\)

4. Examples such as (126), in which the coordinate clause lacks a subject and an auxiliary, have also been excluded from the analysis because such examples do not necessarily illustrate subject ellipsis. On the one hand, (126) might be a case of the ellipsis of a subject AND an auxiliary like shown in (127):

(126) He will talk to his mother and phone me back later.

(127) He will talk to his mother and he will phone me back later.

But according to an alternative analysis, (126) may be seen as VP coordination (128):

(128) He will [talk to his mother] and [phone me back later].

In this context the status of examples such as (129) below is unclear:

(129) He talked to his mother and phoned me back later.

---

This example may illustrate a case of subject ellipsis in a coordinate clause (130a), but it may also illustrate a VP-coordination (130b). As Haegeman suggests, “sentences are centered around I” and I “links a VP and a subject” (2006: 164). Note that I stands for Inflection where tensed auxiliaries and inflectional endings of a verb are inserted. (129) contains two finite verbs, both occur with a tense morpheme -ed: talked and phoned. If we consider that in (129) we have a VP-coordination, then it appears that we have one I position where two inflected endings should be inserted. We cannot have two I positions for one IP projection. But since in English the finite verb does not leave the VP and raise to I but instead “I lowers onto V” (Haegeman, 2006: 169), it seems plausible to assume that the strings talked to his mother and phoned me back later might instantiate just a VP. However, we immediately run into a problem when we consider (131).

(130) a. He talked to his mother and Ø phoned me back later
    b. He [VP talked to his mother] and [VP phoned me back later].

(131) He has talked to his mother and will phone to me later.

In (131) we have two strings: has talked to his mother and will phone to me later. The presence of the auxiliaries implies that neither of them can instantiate a VP. This is because in case of (129) the inflectional ending –ed is a bound morpheme and lowers onto V. In (131) the auxiliaries has and will are free morphemes and are inserted under I while the verb sits under V (Haegeman, 2006: 174). It appears that there should be two I positions to host the two auxiliaries, hence, (131) cannot be a VP-coordination but should be treated as an IP-coordination.

With respect to cases like (129) Quirk et al. (1985) observes that such cases can either be considered as instances when “the conjoined predicates share the same subject” or “as elliptical alternatives to coordinated clauses” (Quirk et al., 1985: 948) which occur due to the “reduce where possible” principle. Similarly, according to Huddleston and Pullum (2002: 1349), (129) can be interpreted in two alternative ways: as a VP-coordination (132a) or as a clause-coordination where the subject of the second clause has undergone ellipsis (132b).

(132) a. He [VP talked to his mother] and [VP phoned me back later].
    b. [IP He talked to his mother] and [IP (he) phoned me back later].
Given the discussion above and given that DNS are treated on a par with a pronoun\textsuperscript{57}, in my corpus I consider cases like (129) as IP-coordination where the pronominal subject of the second/third/fourth conjunct has been omitted.

5. There is one clause in the corpus which prima facie can be considered both as a root and an embedded clause. (133) gives the sentence in the context it appears in.

\begin{quote}
(133) ø Arose at 5:45 A.M., ø read the papers and at 7:10 ø walked to the station to meet the family. ø Took 35 minutes. It was a good walk. Sure ø is fine to have them back. This great white jail is a hell of a place in which to be alone. While I work from early morning until late at night, it is a ghostly place. (Truman Diary, 1947, 6 Jan.)
\end{quote}

The string \textit{is fine to have them back} can be considered as an embedded clause and \textit{sure} can be regarded as the matrix clause where the subject and the auxiliary, namely \textit{I am}, have been omitted. As discussed above, subject + auxiliary omissions are also attested in diaries, so, this postulation seems to be logical. If this is true, then it would mean that we have a null subject in an embedded clause. However, on closer examination, it becomes clear that the clause should be classified as a root because \textit{sure} should not be assimilated to \textit{I am sure} but rather it should be regarded as synonymous to \textit{of course, certainly, definitely}. If \textit{sure} were equal to \textit{I am sure}, we would paraphrase it like (134).

\begin{quote}
(134) I don’t doubt is fine to have them back.
\end{quote}

The reader will hardly agree with such an interpretation of the clause because here Truman does not want to refute his doubts on the fact that \textit{it is fine to have them back}, rather he emphasizes that it \textit{is fine to have them back}. Following this logic, I have classified this clause as a root which has a null subject with what might be considered a preposed adverbial adjunct\textsuperscript{58}.

The next section presents the general figures of the distribution of overt and null subjects according to the clause types as defined above.

\textsuperscript{57} See Sections 1.3.1.1 and 3.3.  
\textsuperscript{58} See Section 4.5.1 for the discussion of preposed constituents.
4.3 Overall figures

This section presents the results of the sample corpus analysis. I consider the realization of the subject of the finite verbs in the three clause categories (roots, coordinate and embedded clauses).

Table 10 illustrates the results for the distribution of finite verbs in the corpus. In the entire diary entry of 1947, 692 finite verbs have been identified. 7 of them occur in imperative clauses which I do not discuss, as noted above. More than half of the 685 finite verbs - 404 (58.98%) - occur in roots, the rest of the verbs are distributed in embedded and coordinate clauses: 151 (22.04%) and 130 (18.98%) respectively. The graph below presents the proportions of each type of clause in percentages.\(^{59}\)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>% 685</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roots</td>
<td>404</td>
<td>58.98%</td>
</tr>
<tr>
<td>Embedded clauses</td>
<td>151</td>
<td>22.04%</td>
</tr>
<tr>
<td>Coordinate clauses</td>
<td>130</td>
<td>18.98%</td>
</tr>
</tbody>
</table>

\(^{59}\) I will present two graphs in this section to illustrate the most important numbers.
Graph 1. The proportions of finite verbs in the corpus presented according to clause types

In Table 11, which illustrates the classification of non-coordinated root clauses, we see that the simple root clauses form the majority among roots - 401 (99.26%), and that the corpus contains 3 (0.74%) exclamative clauses. There are no interrogative clauses in the diary.

Table 11. Root clauses

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>% 404</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple roots</td>
<td>401</td>
<td>99.26%</td>
</tr>
<tr>
<td>Exclamatives</td>
<td>3</td>
<td>0.74%</td>
</tr>
<tr>
<td>Interrogatives</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 12 presents the figures of coordinate clauses attested in the corpus. As mentioned before, I have subdivided the coordinate clauses into 3 types: clauses coordinated with a root clause, clauses coordinated with an embedded clause and single coordinate clauses. According to the data illustrated in the table, among the total of 130 coordinate clauses, there are 100 (76.92%) clauses coordinated with roots, 15 (11.54%) clauses coordinated with embedded clauses and 15 (11.54%) single coordinate clauses. Given the small number of occurrences, any conclusions in terms of the patterns in coordinate embedded clauses and single coordinate
clauses will be tentative. Among the 15 coordinate embedded clauses only 5 have an overt subordinator. In all these cases the subject is overt\textsuperscript{60}. I will present the numbers for the distribution of overt and null subjects for this and all other clause types below.

Table 12. Coordinate clauses

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>% 130</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinated with root clause</td>
<td>100</td>
<td>76.92%</td>
</tr>
<tr>
<td>Coordinated with embedded clause</td>
<td>15</td>
<td>11.54%</td>
</tr>
<tr>
<td>Single coordinate clause</td>
<td>15</td>
<td>11.54%</td>
</tr>
</tbody>
</table>

Table 13 illustrates the overall information on the distribution of overt and non-overt subjects in the diary corpus. The table provides information on the total number of the overt and null subjects attested in the corpus, as well as the proportions of both overt and non-overt subjects occurring in each clause type.

As we can see in Table 13, the first findings based on the material show that in the sample corpus null subjects are found in simple root clauses and in coordinate clauses. They are absent from the embedded and exclamative clauses.

Among the total of 685 finite clauses 465 (67.88%) have an overt subject and accordingly, the remaining 220 (32.12%) have a non-overt subject, i.e. roughly one out of three sentences has a null subject. Given that I have a relatively small corpus, such percentages clearly do not have any statistical value but at the same time they are indicative of the relative frequency of null subjects.

After breaking down the clauses according to clause types, we see that 263 (65.1%) out of 404 root clauses have an overt subject and, accordingly, 141 (34.9%) of them have an implicit subject. Among 401 simple roots 260 (64.84%) have an overt subject and 141 (35.16%) of them lack a subject.

With respect to the coordinate clauses we observe that only 51 (39.23%) out of the total of 130 are attested with an overt subject while there are 79 instances (60.77%), i.e. more than half of the coordinate clauses, with an implicit subject. A more precise breakdown shows that among 100 coordinate roots there are only 30 clauses (30%) which have an overt subject whereas there are 70 instances (70%) with a null subject. 7 (46.67%) out of 15 coordinate

\textsuperscript{60} For subject omission in coordinate embedded clauses see also Section 4.2.
embedded clauses have an overt subject and there are 8 clauses (53.33%) with a non-overt subject. All 8 clauses with an implicit subject have a non-overt subordinator. Although null subjects are not found in the embedded clauses in the material, they have been attested in coordinate embedded clauses. This cannot be surprising as this type of deletion is licit in core grammar, too. I will not discuss this here but will return to it later in Section 4.5.3. As mentioned above, 5 out of seven clauses with an overt subject are introduced with a subordinator. Among 15 single coordinate clauses 14 (93.33%) have an overtly realised subject, 1 (6.67%) is attested with a null subject.

The corpus provides no evidence for the availability of embedded null subjects. All the 151 embedded clauses have an overt subject. As already mentioned in Chapter 2, Haegeman and Ihsane (1999) give examples of embedded null subjects attested in the fictional diary style, however, the data in Truman’s Diary are in line with the fairly generally accepted rule of the non-fiction diary writing: the corpus data provide further evidence that diary subject ellipsis is a root phenomenon.

All the exclamative root clauses attested in the corpus have an overt subject.

Observe that the figures here are just a first rough outline; I will return to a more fine-grained and nuanced analysis of these data below.

Table 13. Overt and null subjects

<table>
<thead>
<tr>
<th></th>
<th>Overt subjects</th>
<th>% Total</th>
<th>Null subjects</th>
<th>% Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple roots</td>
<td>260</td>
<td>64.84%</td>
<td>141</td>
<td>35.16%</td>
<td>401</td>
</tr>
<tr>
<td>Yes/no questions</td>
<td>0</td>
<td>0%</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Wh- questions</td>
<td>0</td>
<td>0%</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Exclamatives</td>
<td>3</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>3</td>
</tr>
<tr>
<td><strong>Roots (total)</strong></td>
<td><strong>263</strong></td>
<td><strong>65.1%</strong></td>
<td><strong>141</strong></td>
<td><strong>35.9%</strong></td>
<td><strong>404</strong></td>
</tr>
<tr>
<td>Embedded</td>
<td>151</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>151</td>
</tr>
<tr>
<td>Coordinated with roots</td>
<td>30</td>
<td>30%</td>
<td>70</td>
<td>70%</td>
<td>100</td>
</tr>
<tr>
<td>Coordinated with embedded</td>
<td>7</td>
<td>46.67%</td>
<td>8</td>
<td>53.33%</td>
<td>15</td>
</tr>
<tr>
<td>Single coordinate</td>
<td>14</td>
<td>93.33%</td>
<td>1</td>
<td>6.67%</td>
<td>15</td>
</tr>
<tr>
<td><strong>Coordinate (total)</strong></td>
<td><strong>51</strong></td>
<td><strong>39.23%</strong></td>
<td><strong>79</strong></td>
<td><strong>60.77%</strong></td>
<td><strong>130</strong></td>
</tr>
</tbody>
</table>
The data in Table 13 are illustrated in Graph 2 below. As there are no cases of embedded clauses and exclamative root clauses with null subjects in the corpus, and, consequently, they present no interest for the given research, I have not included them in the graph.

Graph 2 shows the proportions of overt and non-overt subjects in the types of clauses they are attested.

**Graph 2. The proportions of overt and null subjects in the corpus presented according to clause types**

(135) gives some examples of the null subjects found in the corpus without discussion.

(135) a. Ø Spent the day at work. (Truman Diary, 1947, 1 Jan.)

b. Ø Arose at 5:45 A.M.[,] Ø read the papers and at 7:10 Ø walked to the station to meet the family. (idem., 6 Jan.)

c. Ø Said she didn't like Byrnes and Ø was sure he was not reporting Elliott correctly. (idem., 3 Jan.)

d. Ø Had a most delightful week end. (idem., 3 July)
Arrived in Grandview about 3:30 CST[,] ø went to the house and ø met sister
& brother. (idem., 26 July)

4.3.1 The first results: comparison with Ihsane’s (1998) findings

When we compare the proportions of the distribution of null subjects in the root clauses and in
the coordinate clauses, it becomes evident that in my corpus nearly one third of the roots have
a null subject, whereas among the coordinate clauses nearly two thirds lack a subject. I
tentatively predict that the higher rate of subject omission in coordinate clauses can be
conditioned by the grammaticality of this phenomenon in the core grammar of English.

According to Ihsane’s (1998: 19) findings, the difference of the rate of subject
omission between the two clause types is much bigger. Roughly, only 1 out of 10 root clauses
have a non-overt subject while in case of coordinate clauses the distribution is 76 overt against
69 non-overt subjects. This proportional difference among the clause types again can be
explained by the fact that coordinate null subjects are allowed in Standard English and this
phenomenon is not register-specific while the same cannot be said with regard to the root
clauses.

After drawing parallels between my and Ihsane’s (1998) first results, it appears that in
my corpus root null subjects are more frequent than in hers. Specifically, according to her
findings, only 111 (11.83%) of the 938 root clauses lack a subject while in my corpus the rate
of subject omission is much higher 141/401 (35.16%). With respect to coordinate clauses
Ihsane (1998) reports that 69 (47.59%) out of 145 have a null subject; in my corpus the
coordinate null subjects are more frequently attested: 79 (60.77%) out of the 130 coordinate
clauses lack a subject. It is difficult to say what this higher frequency of subject omission in
my corpus might be conditioned by. Based on these general figures I can assume that in
American diaries the rate of subject omission is higher than in British diaries. Note, however,
that the comparison of only two diaries cannot be sufficient for drawing conclusions.

61 Although Ihsane’s (1998) analysis has been presented in Chapter 3, sometimes I will repeat her figures or
some of the data here.
62 Recall that Teddiman and Newman (2007) examine American and British blog diaries and concludes that the
rate of subject omission in British diaries is higher than in the American ones. The comparison of the results of
Ihsane’s and my corpuses points in the other direction. However, one should examine more diaries to be able to
draw more accurate conclusions.
In both corpuses there are no null subjects in exclamative and embedded clauses. The non-availability of null subjects in these environments follows from the analysis presented in Chapter 2 and the findings of both corpuses show that subject omission in diaries is a root phenomenon.

The following sections elaborate on the classification of null subjects and examine the grammatical properties of referential null subjects. Recall that in her analysis Ihsane (1998) presents only the proportions for the null subjects. I will provide the numbers both for the overt and the non-overt subjects in order to give a more clear idea about their distribution patterns.

4.4 The classification and grammatical properties of null subjects

This section provides a more fine-grained analysis of null and overt subjects attested in the sample corpus: first, I distinguish between referential subjects and expletive subjects and give the proportions for the null and overt subjects of both types according to the clause types they occur in. Furthermore, the grammatical features of the referential implicit and overt subjects, that is, their grammatical person and number are examined according to the clause types. Given that expletives belong to the category of the third person a comparative overview of the distribution of the third person subjects and expletive subjects will be presented.

Then, I observe the null and overt subject patterns according to the type of the verb (lexical verb, auxiliary, copula be) they are attested with. Recall that Ihsane (1998) also studies the proportions of null subjects according to their occurrence in negative/positive clauses and the tense of the verb. I do not investigate these categories because I don’t see in what ways a negative or a positive verb or a verb tense can be related to the diary drop phenomenon. Finally, I examine whether null subjects in the sample diary are sensitive to the material which can precede them.

In this section I do not discuss the exclamative and embedded clauses because there are no attested null subjects in these types of clauses. Given that there is only 1 case of subject omission in the single coordinate clauses in my corpus and it will be impossible to draw any conclusions based on the single example, I exclude this clause type from my further analysis, too.

63 There are no implicit subjects in interrogative clauses in Ihsane’s corpus either. Recall that in Truman’s diary there have not been found any interrogatives at all.
1.1.1 Referential and expletive subjects

To study the distribution of subject patterns further I classify them into referential (136a, b) and expletive subjects (*it, there*) (136c, d). The examples are taken from Tuman’s Diary and illustrate cases of non-overt subjects. In the given section the two types of subjects will be analyzed according to the type of clause they occur in: root clauses, clauses coordinated with roots and clauses coordinated with embedded clauses.

(136) a. Ø Spend a pleasant day. (Truman Diary, 1947, 3 March)
    b. Ø Landed in Washington at 4:16. (idem, 29 July)
    c. Ø Looks as if we've lost a grand, honest man & wife of the same caliber and have gained a good man and a baby talking, henna haired lady. (idem, 23 July)
    d. Ø Makes a person ashamed to be gloomy even if world affairs are mixed up. (idem, 25 Dec.)

Table 14 illustrates the occurrence of referential and expletive subjects in the diary corpus. The table shows the proportions of referential null and overt subjects as well as the expletive null and overt subjects which occur in the corpus. In addition, it provides the relative proportions of referential and expletive null subjects according to the clause types they occur in.

As the results show, among the 516 attested subjects there are 13 (2.52%) expletive subjects. 5 (38.5%) out of the 13 expletives are implicit; all these non-overt instances occur in root clauses. Among the 8 overt expletive subjects 7 are attested in root clauses, there is 1 case in a coordinate root clause. There are no cases of null expletive subjects attested in coordinate roots and coordinate embedded clauses. Observe that in the core grammar omission of an expletive in the second conjunct is licit provided the first conjunct contains an expletive, i.e. the subjects of both conjuncts should be co-referential (137):

(137) a. It will rain all night and Ø may be quite windy at times.
    b. He said that it will rain all night and Ø may be quite windy at times.

Recall that in this and the upcoming sections the embedded, exclamative as well as single coordinate clauses have been left out from the analysis.
The absence of such examples in my corpus is probably due to the relatively low numbers of expletive subjects (taking into account the general numbers). Recall that I have excluded from the discussion all the instances in which subjects are deleted with an auxiliary. It is possible that such examples would also contain expletive subjects (138):

(138) a. Perfect day. (Truman Diary, 1947, 4 March)
   b. A very pleasant meeting. (idem. 15 Dec.)

Table 14 also makes it apparent that out of 516 subjects 503 (97.48%) are referential. Among them there are 214 cases (42.54%) when the subject is non-overt. Both overt and null referential subjects occur in all three clause types. 253 (65.04%) out of 389 referential subjects attested in root clauses are overt while 136 (34.96%) of them are implicit.

Among the total of 99 coordinate roots with an attested referential subject there are 29 cases (29.29%) when the subject is overtly realized. The referential subject is non-overt in 70 cases (70.71%). 7 (46.67%) out of the 15 referential subjects attested in coordinate embedded clauses are overt while 8 (53.33%) are omitted. As mentioned above, embedded clauses and exclamative clauses consistently do not display subject omission in the diary examined.

If we compare the proportions of referential and expletive null subjects which occur in root clauses\(^\text{65}\), we can see that the picture is roughly similar with the expletive subjects displaying a slightly higher rate of deletion as compared with the referential ones: 136 (34.96%) of the total of 389 referential subjects are non-overt, while in case of expletive subjects 5 (41.67%) out of 12 are implicit.

In order to draw more accurate conclusions with respect to the distribution of the expletive and referential subjects it will be necessary to see how many of them occur with a contracted verb. This is important because a contracted verb (‘m, ‘ve, ‘s…) makes the deletion of the subject impossible: the subject must be overt because the contracted verb is a clitic and needs a host\(^\text{66}\). “In most cases, that host is a pronoun (e.g. I’m, you’d, she’ll, that’s)” (Biber et al. 1999: 1128). According to Biber et al., “full nouns, wh-words, and there” can also “serve as a host” (1999: 1128). Further Biber et al. (1999: 1128) indicates that the contraction is not possible if there is no preceding host. It follows that a subject cannot be deleted if it is followed by a contracted form of a verb (139).

\(^{65}\) It will be impossible to draw any conclusions with respect to the coordinate clauses as in the corpus in this clause type there is only one attested instance of an expletive subject which is overtly realized (see Table 14).

\(^{66}\) See also Biber et al. 1999: 1128.
(139) a. He’s a nice person.
b. We’ll be back next Sunday.
c. *φ’d join you but I have to work on my paper.

There is only one instance when in roots an expletive occurs with a contracted verb whereas in case of the referential subjects 22 instances have been identified. If we refine the data removing these cases from the total of overt expletive and referential subjects, we get that in roots 5 (45.45%) out of 11 expletive subjects are implicit. The referential subjects are distributed as follows: 136 (37.06%) out of 367 subjects are non-overt. The first results show that in roots expletive subjects are omitted more frequently than the referential ones.

However, as the expletive belongs to the category of the third person, in Section 4.4.2.1 I will do a further breakdown and will look at the proportions of the expletive null subjects and the referential null subjects which can be interpreted as the third person singular.

Table 14. Referential and expletive null and overt subjects according to clause types

<table>
<thead>
<tr>
<th></th>
<th>Overt subjects</th>
<th>%Total</th>
<th>Null subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential</td>
<td>289</td>
<td>57.56%</td>
<td>214</td>
<td>42.54%</td>
<td>503 (97.48%)</td>
</tr>
<tr>
<td>Roots</td>
<td>253</td>
<td>65.04%</td>
<td>136</td>
<td>34.96%</td>
<td>389</td>
</tr>
<tr>
<td>Coordinated with roots</td>
<td>29</td>
<td>29.29%</td>
<td>70</td>
<td>70.71%</td>
<td>99</td>
</tr>
<tr>
<td>Coordinated with embedded</td>
<td>7</td>
<td>46.67%</td>
<td>8</td>
<td>53.33%</td>
<td>15</td>
</tr>
<tr>
<td>Expletive</td>
<td>8</td>
<td>61.5%</td>
<td>5</td>
<td>38.5%</td>
<td>13 (2.52%)</td>
</tr>
<tr>
<td>Roots</td>
<td>7</td>
<td>58.33%</td>
<td>5</td>
<td>41.67%</td>
<td>12</td>
</tr>
<tr>
<td>Coordinated with roots</td>
<td>1</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>1</td>
</tr>
<tr>
<td>Coordinated with embedded</td>
<td>0</td>
<td>0%</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
</tbody>
</table>
4.4.1.1 Comparing the results with Ihsane: referential and expletive subjects

Among the total of 111 null subjects Ihsane (1998: 21) identifies 1 (0.90%) null expletive subject. She concludes that expletives in diary registers are rare. According to the data in my corpus, there are 5 (2.27%) null expletive subjects among the total of 220 null subjects which means that the rate of omission of expletive subjects in my corpus is relatively higher than in Ihsane’s corpus. It is essential to notice, however, that we don’t know how many clauses with an overt expletive subject Ihsane’s corpus contains because after providing the general figures for the clause types, Ihsane focuses only on the null subjects. Observe, however, that the proportions of the overtly realized subjects would be relevant for drawing more precise conclusions with respect to the frequency of occurrence of subjects in diaries. Ihsane demonstrates the proportions of the referential and expletive subjects in general terms and does not refine her analysis to consider the distribution of the expletive subjects vs. the third person pronominal subjects. Given this it is hard to draw any sufficient conclusions with regard to the relative rate of deletion of expletive subjects in her corpus.

The comparative analysis based on the general figures shows that Truman’s Diary demonstrates a higher rate of expletive subject omission than Virginia Woolf’s Diary, though, again it should be mentioned that taking into account the small figures, all conclusions can be tentative.

Further I will not discuss the expletive subjects and will confine the analysis to referential subjects. I will only revisit the discussion of expletives in Section 4.4.2.1 to compare their rate of deletion with that of the third person pronominal subjects, as already mentioned above. The next section deals with the interpretation of null subjects. Then, I focus on the distribution patterns of the referential null and overt subjects according to the grammatical categories of person and number.

4.4.2 The interpretation of null subjects

This section deals with the general figures of null subjects according to the grammatical categories of person and number. In the following sections I will provide a more nuanced analysis in terms of these categories giving the figures for both overt and null subjects according to the clause types.
On closer observation of the grammatical features of person and number of the null subjects attested in the corpus, it becomes obvious that the majority of the null subjects can be interpreted as being first person singular. As shown in Table 15, out of 214 referential null subjects 150 (70.1%) can be interpreted as I. 30 (14.02%) are understood as we. There are 26 (12.15%) instances which refer to third person singular and 8 (3.73%) cases understood as third person plural. Observe that these are the overall figures illustrating the distribution of the null subjects with respect to the grammatical categories of person and number; I will give a breakdown of these numbers for the clause types and will provide examples for each pattern in the later sections.

Table 15. Referential null subjects: person and number

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>% 214</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person sg.</td>
<td>150</td>
<td>70.1%</td>
</tr>
<tr>
<td>1st person pl.</td>
<td>30</td>
<td>14.02%</td>
</tr>
<tr>
<td>3rd person sg.</td>
<td>26</td>
<td>12.15%</td>
</tr>
<tr>
<td>3rd person pl.</td>
<td>8</td>
<td>3.73%</td>
</tr>
</tbody>
</table>

The observed high rate of I-omissions seems rather natural in view of the material which is being studied: in a diary the author writes about him/herself: the diary writer is, in fact, the default topic of the diary. The omission of the first person pronoun I can, thus, be motivated by functional reasons: since the diary is about the narrator/writer, I does not need to be expressed overtly, as already discussed in Chapter 1.

Observe that in the present tense the inflectional ending of the verb -(e)s would suffice to identify third person singular subjects. However, no such cases have been identified in my corpus when a third person singular subject is followed by a verb in present simple which carries the only inflectional ending in English denoting the third person singular. In 24 instances the third person singular null subject is followed by a verb in past tense, in the remaining 2 cases the third person implicit subject is followed by the auxiliaries will and must. In all 26 cases the third person singular subject can be recovered through a contextually salient antecedent. Similarly, the context is helpful when identifying we and they. I will return to this later in Section 4.7.

There are no cases of null subjects which can be interpreted as second person singular or plural but there are two cases of overtly realized second person subjects attested in the
entries. I will discuss both cases of you in the next section. The very rare occurrence of you can be explained by the fact that unlike letters, diaries have no addressee.

To conclude, the majority of the null subjects found in the corpus refer to the narrator, there are no instances of null subjects referring to the addressee.

In the following sections I discuss the categories of person and number of the attested overt and non-overt subjects relating them to the clause type they occur in. For all the categories examined I consider how many of the overt third person subjects are pronouns and how many are lexical NPs. This seems essential because, as discussed above, the subjects represented by lexical NPs cannot be deleted as they are not recoverable in the discourse context while pronouns can be recovered, as already discussed above. So I will contrast the distribution of the third person null subjects both with that of the overtly realized pronominal subjects, and the total number of the overt subjects, i.e. both the lexical NPs and the pronouns.

4.4.2.1 The categories of person and number in root clauses

This section provides an overview of the categories of person and number of the overt and non-overt subjects attested in root clauses in the corpus. The examples in (140) illustrate the omitted subjects identified as first person singular (a), first person plural (b), third person singular (c). There are no cases of they-omissions attested in root clauses in the corpus.

(140) a. Ø Returned to Washington. (Truman Diary, 1947, 29July)

b. Ø Saw all of Brazil from Belem to Rio de Janeiro. (idem, 1Sept.)

c. Ø Said Argentine [sic] wanted to get along with us, etc. (idem, March 31)

As shown in Table 16, among the total of 389 overt and non-overt subjects in root clauses 169 are understood as first person singular. 109 (64.5%) out of these are implicit. This means that about two thirds of the subjects identified as first person singular are non-overt. As already mentioned above, the high rate of I-omission is not very surprising as the author writes about him/herself and appears to be the ‘topic’ of the diary.

In root clauses there are 39 overt and non-overt subjects, interpreted as first person plural. Among these 19 (48.72%) are non-overt. This means that in the material studied, the ratio of non-overt vs. overt subjects is much higher with the first person singular than with the first person plural.
The diary contains 6 instances when the subject is expressed by a lexical NP+I, e.g. the President and I, Byrnes & I. It seems logical to treat these cases as equal to we. If we add these 6 instances to the total number of the first person plural subjects, we get that 19 (42.22%) out of 45 we-subjects are implicit. But given that DNS are treated on a par with a pronoun, as discussed in the preceding sections, the ratio of 19/39 seems to be more relevant.

There are 2 cases of the second person subject you in the corpus; it would be natural to expect no subjects denoting the second person singular or plural since a diary lacks an obvious addressee. The two instances are represented below. The context they occur in is provided, too, in order to make the interpretation more comprehensible (141a, b):

(141) a. But they all walk up and down the halls of this place and moan about what they should have done and didn't. So-you see. I've only named a few. (Truman Diary, 1947, 6 Jan.)

b. The rule around here is that no one may speak to the President. I break it every day and make 'em speak to me. So-you see what I get. But I still want 'em to tell me. (idem. 16 Jan.)

As we can see in (141), both cases of you are attested in the same sentence pattern: So-you see. Prima facie this sounds as an address, not a narration. In both cases the writer seemingly expresses discontent over the existing circumstances. I assume that you might refer to a potential reader: Truman might have expected that someone will read his diary in the future and addresses him/her to share his thoughts.

Then, as we can see in the table, there are 131 third person singular subjects, 8 (6.1%) of which are non-overt. A further breakdown of the overt third person singular subjects into lexical NPs and pronouns shows that there are 65 lexical NPs (49.62% of the total 131) and 58 pronouns (44.28%). Given that I treat DNS as a deleted pronoun, as discussed above, further it would be necessary to compare the third person singular overt pronouns with the non-overt ones. Thus, we get 58 third person singular overt pronouns and 8 non-overt ones, that is 8 (12.12%) out of 66 third person pronouns are non-overt.
Table 16. The categories of person and number of overt and null subjects in root clauses

<table>
<thead>
<tr>
<th>Category</th>
<th>Overt subjects</th>
<th>%Total</th>
<th>Non-overt subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person sg.</td>
<td>60</td>
<td>35.5%</td>
<td>109</td>
<td>64.5%</td>
<td>169</td>
</tr>
<tr>
<td>1st person pl.</td>
<td>20</td>
<td>51.28%</td>
<td>19</td>
<td>48.72%</td>
<td>39</td>
</tr>
<tr>
<td>Lexical NP+1st person sg.</td>
<td>6</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td>2nd person</td>
<td>2</td>
<td>100%</td>
<td>0</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td>3rd person sg. lexical NP</td>
<td>65</td>
<td>49.62%</td>
<td>0</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td>3rd person sg. pronouns</td>
<td>58</td>
<td>44.28%</td>
<td>8</td>
<td>6.1%</td>
<td>131</td>
</tr>
<tr>
<td>3rd person pl. lexical NP</td>
<td>25</td>
<td>59.52%</td>
<td>0</td>
<td>0%</td>
<td>42</td>
</tr>
<tr>
<td>3rd person pl. pronouns</td>
<td>17</td>
<td>40.48%</td>
<td>0</td>
<td>0%</td>
<td>42</td>
</tr>
</tbody>
</table>

As already mentioned above, given that the expletive belongs to the third person category, it seems logical to do a further breakdown and look at the proportions of the expletive subjects and the referential subjects which can be interpreted as third person singular. Since null subjects are interpreted as pronouns, I will consider only the third person pronouns. Moreover, since omission in root contexts is different from that in coordination, and as there are no null expletive subjects in coordinate clauses in the corpus as Table 14 shows, I compare only the subject patterns attested in the roots.

The figures are illustrated in Table 17. As we can see, there are 66 instances of third person referential subjects, 8 (12.12%) of which are implicit. Out of the total of 12 expletive subjects 5 (41.67%) are non-overt. We get 8/66 against 5/12 or, to be more precise, 12.12% against 41.67%. As noted in Section 4.4.2, for drawing more accurate conclusions it would be necessary to remove all the cases when a subject has a contracted verb. There are 9 cases of a third person referential subject + a contracted verb and 1 case of an expletive subject + a contracted verb. Thus, recalculation of the data shows that out of 57 third person referential subjects 8 (14.04%) are non-overt while among 11 expletive subjects 5 (45.45%) are implicit. This suggests that null expletives are more frequently attested than the third person referential null subjects. However, let me notice that the figures presented here are not high enough to allow me to draw conclusions.

---

67 I use the term verb as among them there are both lexical and auxiliary verbs, as well as copula be. For the classification of verb types see Section 4.4.3.
Table 17. The distribution of referential third person pronoun subjects and expletive subjects in roots

<table>
<thead>
<tr>
<th></th>
<th>Null subjects</th>
<th>%Total</th>
<th>Overt subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd person sg. pronouns</td>
<td>8</td>
<td>12.12%</td>
<td>58</td>
<td>87.88%</td>
<td>66</td>
</tr>
<tr>
<td>Expletive</td>
<td>5</td>
<td>41.67%</td>
<td>7</td>
<td>58.33%</td>
<td>12</td>
</tr>
</tbody>
</table>

Finally, there are 42 instances of third person plural subjects. Among those there are no cases of subject omission. Again, the breakdown of the third person overt subjects into lexical NPs and pronouns reveals that 25 of them (59.42%) are lexical NPs, while in 17 (40.48%) instances the subjects is expressed by a pronoun.

If we compare the omission rate of the first person pronouns and that of the third person pronouns, the ratio of non-overt vs. overt for third person singular subjects (8/66) is much lower than it is for the first person singular subjects (109/169). Similarly, there are relatively fewer cases of non-overt third person singular subjects than of the first person plurals (19/39).

Given that DNS are non-overt pronouns, the next logical step will be to observe the ratio of pronoun omissions in relation to the person/number category. To understand the distribution of the pronominal overt and null subjects in roots with respect to the categories of person and number it will be necessary to subtract the number of the overt third person singular and plural subjects which are represented by a lexical NP from the total of the overt subjects. Thus, if 96 out of the total of 253 overt subjects are lexical NPs, we get 136 null subjects vs. 157 pronominal overt subjects, which means that 46.42% of the pronominal subjects are implicit.

Next, I compare the results of Table 16 with Ihsane’s findings.

---

68 The figures provided in the table include the cases with contracted verbs.
4.4.2.2  The categories of person and number: comparison with Ihsane’s findings

Given that Ihsane (1998) does not focus on the distribution of subject patterns according to the grammatical categories of the person and number in coordinate clauses and confines her analysis to roots, it would be appropriate to compare the findings of both corpuses here. According to Ihsane (1998: 24), in her corpus 53 (47.75%) out of 111 implicit subjects refer to the narrator. In my corpus I-omissions in roots form a vast majority – 109 (80.15%) out of 136. Ihsane reports 17 (15.32%) instances of null subjects which are understood as we. In Truman’s Diary there are 19 (13.97%) we-omissions. For third person singular omissions we get 36 cases (32.43%) in Ihsane’s corpus and 8 cases (5.88%) in my corpus. Ihsane identifies 5 (4.5%) third person plural null subject patterns while in my corpus there are no they-omissions attested in roots.

To conclude, in Truman’s Diary the rate of I-omissions in roots is considerably higher than in V. Woolf’s Diary. First person plural subjects demonstrate roughly an equal rate of deletion. For third person singular and plural subjects, the figures in my material are relatively lower as compared with those in Ihsane’s corpus.

4.4.2.3  The categories of person and number in clauses coordinated with roots

This section deals with the distribution of overt and non-overt referential subjects attested in coordinate root clauses, taking into consideration the grammatical categories of person and number of the subjects. First, I present the numbers for this clause type, then, I compare them with the results of root clauses. In (142) examples of null subjects occurring in coordinate root clauses are presented. They have been identified as first person singular (142a), first person plural (142b), third person singular (142c) and third person plural (142d).

(142) a. Arose at 5:45 A.M.,[.] Ø read the papers and at 7:10, Ø walked to the station to meet the family. (Truman Diary, 1947, 6 Jan.)
   b. We get into my big open Lincoln car and Ø start for American Embassy. (idem, 4 March)
   c. Mr. Byrnes called at 5 P.M. and Ø said he’d like to see me. (idem, 7 Jan.)
   d. The new pilots were rattled on account of the passenger and Ø were careful & conservative. (idem, 29 July)
In all the examples in (142) the subject of the second (and third in case of (142a)) conjunct is co-referential with that of the first one. Recall that subject omission in the second conjunct of a coordinate clause is also licit in core grammar if the subjects of both conjuncts are co-referential. However, the rule of co-referentiality does not always apply in diaries. As the corpus data show, in diaries the subject of the second conjunct can be implicit even if it is not co-referential with that of the preceding clause. Specifically, 2 out of 70 coordinate null subjects are different from that of the first conjunct. All the four clauses are illustrated in (143). The clauses are given in the context they appear in and are marked in bold. The coordinate null subject in (143a) can be interpreted as I, the one in (143b) can be understood as we. I will discuss them in Section 4.6.

(143) a. Doc tell's [sic] me I have Cardiac Asthma! Ain[']t that hell. Well it makes no difference[,] Ø will go on as before. I've sworn him to secrecy! So What! (idem, 7 March)

b. Arrived in Grandview about 3:30 CST[,] went to the house and met sister & brother. Went to Belton with them and Ø picked a casket. (idem, 26 July)

Given that the omission of co-referential coordinate subjects is allowed in Standard English, too, it will be interesting to see whether co-referentiality plays a role with respect to the rate of subject omission in diaries. For this purpose I compare the distribution ratio of overt and non-overt subjects in roots with those in coordinate clauses. I do not include the 2 non-co-referential coordinate null subjects given in (143) above in my discussion.

As Table 18 makes it clear, in the clauses coordinated with roots the majority of the null subjects can be identified as first person singular. Among the total of 99 cases there are 52 instances of I, 39 (75%) of which are implicit. When comparing the first-person-singular subject ellipsis in terms of the grammatical categories of person and number in root and coordinate root clauses, we get 109/169 vs. 41/54. In both clause types I is non-overt in most of the cases as compared with the other person/number categories; in coordinate root clauses the rate of I-omission is higher (75%) than in roots (64.5%). Observe that if in case of root clauses one could assume that the high rate of the first person singular subject omissions might be determined by pragmatic reasons as postulated in Chapter 1, in case of coordinate roots the recoverability of a subject has syntactic foundations. This is because the subject ellipsis in the second conjunct of a coordinate clause is licit in the core grammar of English and is not register-specific. Specifically, the subject of the second conjunct can be deleted provided it is
co-referential with the subject of the preceding clause. I will discuss this restriction with respect to the diary material in Section 4.6.

There are 11 cases of first person plural, 9 (81.82%) of which are non-overt. Comparing the ratio of the subjects understood as the first person singular and plural, one can see that in coordinate roots we-omissions are slightly more frequently attested than I-omissions in my material, though it is difficult to draw conclusions from such small numbers. Observe that if the subject omission were just a matter of pragmatics then one would expect that, as is the case in roots, the first person singular would have the highest rate of omission due to its high recoverability status. But it is not so. Comparison of the results with roots reveals that for first person plural subjects in roots we have nearly as many overt subjects as non-overt ones: 19/39, i.e. we is non-overt in 48.72% of the cases whereas in coordinate root clauses the picture is completely different: 9/11: the subject is null in about 80% of the cases. Why can this be? I will consider this discrepancy below.

Further, 15 (65.22%) among the attested 23 third person singular subject patterns are implicit. A further breakdown of the third person singular overt subjects into pronouns and lexical NPs reveals that there are 5 lexical NPs and 3 pronouns. Comparing the ratio of the non-overt vs. overt pronominal subjects we get that more than two thirds (83.33%) of the attested subjects are non-overt: 15/18. The picture is quite different for the same category in root clauses: there are about 8 times as many overt subjects as null ones attested - 8/66 (subject is null in 12.12% of the cases). Again, such a big difference in the rate of subject omission in coordinate clauses as compared with roots cannot be just a matter of pragmatics. This discrepancy will be considered later.

As to third person plural subjects, in this clause type there are a total of 11 cases and 5 (45.46%) of them are non-overt. A breakdown of the third person plural overt subjects into lexical NPs and pronouns shows that 2 out of 6 third person plural overt subjects are lexical NPs, and the other 4 are pronouns. Thus, the comparison of the pronominal null vs. overt subjects reveals: 4/9. So, while there are no third person plural implicit subjects attested in root clauses, in coordinate root clauses nearly the half of the attested they-subjects are non-overt. I will consider this discrepancy below.
Table 18. The categories of person and number of overt and null subjects in coordinate root clauses

<table>
<thead>
<tr>
<th></th>
<th>Overt subjects</th>
<th>%Total</th>
<th>Non-overt subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; person sg.</td>
<td>13</td>
<td>25%</td>
<td>39</td>
<td>75%</td>
<td>52</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; person pl.</td>
<td>2</td>
<td>18.18%</td>
<td>9</td>
<td>81.82%</td>
<td>11</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; person sg. lexical NP</td>
<td>5</td>
<td>21.74%</td>
<td>15</td>
<td>65.22%</td>
<td>23</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; person sg. pronouns</td>
<td>3</td>
<td>13.04%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; person pl. lexical NP</td>
<td>2</td>
<td>18.18%</td>
<td>5</td>
<td>45.46%</td>
<td>11</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; person pl. pronouns</td>
<td>4</td>
<td>36.36%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Next, it will be necessary to observe the ratio of pronominal subject omissions in this clause type, like it was done for the roots. So, out of 29 overt subjects 7 are lexical NPs which means: 68 null subjects vs. 22 overt subjects: 68/90, i.e. 75.55% of the pronominal subjects are implicit. Observe that this is much higher than it was observed in case of roots: 136/293 (46.42%).

As it became obvious throughout the section, this clause type demonstrates a higher rate of subject omission with respect to the grammatical category of person and number as compared with simple roots. As already suggested above, this can be determined by the fact that subject omission in a coordinate clause is grammatical in Standard English and is not register-specific while root subject ellipsis is not licit in core grammar and is restricted to diary registers. This factor can account for the differences of the frequency of subject deletion in the two types of clauses.

4.4.2.4 The categories of person and number in coordinate embedded clauses

This section presents the analysis of the categories of person and number of the subject patterns attested in the clauses coordinated with embedded clauses. (144) illustrates examples of omitted subjects attested in this clause type. (144a) illustrates a first person plural subject, (144b) presents an example of third person singular and the empty category in (144c) can be interpreted as third person plural.
(144) a. Looks as if we've lost a grand, honest man & wife of the same caliber and Ø
have gained a good man and a baby talking, henna haired lady. (Truman
Diary, 1947, 23 July)

b. Said he'd got his checks mixed up, Ø had lied to the Secret Service and he
wanted to tell me before his boss did. (idem, 6 Jan.)

c. Mrs. Roosevelt came in at 3 P.M. to assure me that Jimmy & Elliott had
nothing against me and Ø intended no disparagement of me in their recent
non-edited remarks. (idem, 3 Jan.)

The data are illustrated in Table 19. Given the low numbers the percentages are not
presented. As we can see, there is only one case of I which is overt. There are 2 instances of
the first person plural one of which is non-overt. In roots the picture is similar; nearly the half
of the first person plural subjects are null, in coordinate roots 8 out of 10 subjects are null,
which is relatively higher.

Among the 7 third person singular subjects 3 are non-overt. All 4 overt subjects are
pronouns. Comparing the omission frequency of this category in the given clause type with
those of roots and coordinate roots we get that in this clause type 3 out of the 7 examples of
the third person singular subjects are implicit whereas in the root clauses the ratio of null and
overt subjects is 1 vs. 8 and in coordinate root clauses more than two thirds of the subjects in
this category are non-overt. Thus, with respect to third person singular, coordinate roots
demonstrate a higher rate of subject deletion, too.

Among the 5 third person plural subjects 4 are implicit. The only overt subject is
represented by a lexical NP. In root clauses there are no cases of implicit they among the total
of 48 instances attested in the corpus. In coordinate root clauses nearly the half of the attested
third person plural subjects are implicit. In the given category the rate of the pronominal
subject omission attested in coordinate embedded clauses overweighs the deletion rate
identified in the two other clause types.

As it has been already mentioned, none of the embedded clauses with a null subject has
an overt subordinator.69

---

69 See also Sections 4.2 and 4.3.
Table 19. The categories of person and number of overt and null subjects in coordinate embedded clauses

<table>
<thead>
<tr>
<th>Category</th>
<th>Overt subjects</th>
<th>Non-overt subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person sg.</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1st person pl.</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3rd person sg. lexical NP</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd person sg. pronouns</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3rd person pl. lexical NP</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd person pl. pronouns</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

In this clause type there are no subjects introduced by a lexical NP, so no further breakdown of figures is needed to observe the distribution of the pronominal overt and non-overt subjects. Note that given the small numbers any conclusions established will be tentative.

4.4.2.5 Concluding remarks

The analysis above reveals that the majority of the null subjects attested in the sample corpus can be interpreted as first person singular. I tentatively conclude that this is conditioned by the nature of the material under study: since the diary is about the narrator/writer, and hence, I is the default ‘topic’ of the diary, it does not need to be expressed overtly.

Then, although diaries do not have an addressee, there are 2 cases of you-subjects in the corpus; this has implications for the ‘expected audience’ hypothesis.

The analysis also shows that subject omission is more frequent in coordinate clauses as this type of ellipsis is also licit and widely applicable in core grammar. Recall, however, that coordinate clauses demonstrate a register-specific property: the rule of co-referentiality of subjects is not always respected in diaries, as noted above: there are 2 cases among 70 when the null subject of the coordinate clause is not co-referential with the subject of the preceding clause. I refer the reader to Section 4.6 for discussion of non-coreferential coordinate null subjects. In the following section I will observe the tense of the verbs the null and overt
subjects are attested with in the corpus. The analysis is conducted according to the clause types.

4.4.3 The type of verbs

This section deals with the type of the verbs the overt and implicit subjects occur with in the corpus. The distribution of the subject patterns is analyzed according to the clause types the patterns occur in.

As was done in Ihsane (1998: 26), I classify the types of verbs the overt and null subjects occur with into three different subclasses: lexical verbs, auxiliary verbs, copula be. Verbs like come, do(=perform), have\textsuperscript{70}, read, live, meet, jump, dare\textsuperscript{71}, need are lexical (145), while may, will, can, have\textsuperscript{72}, dare, need, be\textsuperscript{73}, do(‘dummy’ operator)\textsuperscript{74} are auxiliary verbs\textsuperscript{75} (146).

Copula be and auxiliary be are differentiated by the type of complement they take. Basically, when be takes a VP complement, it is considered as an auxiliary, if it takes an NP, AP, PP complement, it is classified as a copula be (147).

(145) a. Ø Began getting things in order. (Truman Diary, 1947, 26 July)  
    b. Ø Returned to Washington. (idem, 29 July)

(146) a. Ø Was driven to Charlottesville, Virginia at 2 P.M. (idem, 3 July)  
    b. Ø Have seen pictures of Franz Joseph, Marcus Aurelius & Napoleon doing it. (idem, 4 March)

(147) They seem to like us but Ø are suspicious of the British. (idem, 16 Jan.)

As seen in Chapter 3, Ihsane (1998: 27-28) reports that the majority of null subjects in her corpus are followed by a lexical verb whereas with copula be the null subjects are attested least frequently. It will be interesting to observe whether the same distribution applies to

\textsuperscript{70} This is the transitive lexical verb “with DO-periphrasis” (Quirk et al., 1972: 80)

\textsuperscript{71} “DARE and NEED can be constructed either as modal auxiliaries (with the bare infinitive and without any inflected –s form) or as lexical verbs (with the to-infinitive and with the inflected –s forms)” (Quirk et al., 1972: 82-83). I have taken into account these criteria for classification.

\textsuperscript{72} This is the “aspect auxiliary HAVE” used “to form perfective complex verb phrases” and the transitive verb “without DO-periphrasis” (Quirk et al., 1972: 80)

\textsuperscript{73} For further classification of be see below.

\textsuperscript{74} See also Quirk et al. (1972: 77) for classification of do.

\textsuperscript{75} All ‘primary auxiliaries’ and modal verbs have been classified as auxiliaries. See Quirk et al. 1972: 69.
Truman’s diary. One reason why auxiliaries and copula be are less frequently attested with DNS than lexical verbs might be that auxiliaries/copula be can delete with a subject, as already mentioned before (cf 124), while lexical verbs cannot. Let us see what the figures in my corpus tell.

4.4.3.1 The verb type: root clauses

In this section I illustrate the distribution of lexical and auxiliary verbs as well as the copula be in root clauses. (148) provides examples of null subjects attested with lexical and auxiliary verbs, as well as the copula be in root clauses. Lexical verbs are illustrated in (148a), auxiliaries in (148b) and the only copula be in (148c), (149) is an example with a contracted auxiliary.

(148) a. Ø Spent the day working on messages. (Truman Diary, 1947, 2 Jan.)
    b. Ø Was sitting at my desk just before dinner tonight when [name of person and
       staff position restricted] came up …(idem, 16 Jan.)
    c. Ø Was aboard the Williamsburg with secretaries and military [sic] and naval
       aides, and Adm[iral] Leahy. (idem, 1 Jan.)

(149) He's a nice boy and will go places. (idem, 6 Jan.)

Table 20 provides the details on the occurrence of the various verb types with overt and non-overt subject patterns in root clauses. As we can see, 274 of the attested 389 finite verbs are lexical verbs and 128 (46.72%) of these display a null subject pattern. 55 clauses have a finite auxiliary among which 7 (12.73%) occur with an implicit subject. Finally, only 1 (1.67%) out of the total of 60 clauses with a finite copula be has a non-overt subject. The findings show that in this clause type null subject patterns are most frequently attested with lexical verbs. As already noted above, the small numbers of auxiliaries and copula be attested with null subject patterns can be determined by the fact that these verb types can delete with a subject.
Table 20. The verb type: root clauses

<table>
<thead>
<tr>
<th></th>
<th>Overt</th>
<th>%Total</th>
<th>Null subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical verbs</td>
<td>146</td>
<td>53.28%</td>
<td>128</td>
<td>46.72%</td>
<td>274</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>48</td>
<td>87.27%</td>
<td>7</td>
<td>12.73%</td>
<td>55</td>
</tr>
<tr>
<td>Copula be</td>
<td>59</td>
<td>98.33%</td>
<td>1</td>
<td>1.67%</td>
<td>60</td>
</tr>
</tbody>
</table>

Given that a contracted verb form would not allow a subject drop\textsuperscript{76} I have also examined whether any of the subjects occurs with a contracted verb. There are 2 cases when the lexical verb have has been contracted (1 in present tense, 1 in past tense), 14 cases of a subject + a contracted auxiliary and 8 cases of a subject + a contracted copula be in roots in the corpus. Given that the subject cannot be null with a contracted verb, it will be appropriate to refine the figures to make a more precise idea with regard to the distribution of the subject patterns. Removing all the cases with contracted verbs, we get: 144 overt lexical verbs, 34 overt auxiliaries, 51 overt copula be. As Table 21 illustrates, even after recalculation of the data, the majority of the null subjects occur with a lexical verb and implicit subjects are least attested with copula be.

Table 21. The verb type: root clauses: refined numbers

<table>
<thead>
<tr>
<th></th>
<th>Overt</th>
<th>%Total</th>
<th>Null subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical verbs</td>
<td>144</td>
<td>53.14%</td>
<td>128</td>
<td>47.06%</td>
<td>272</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>34</td>
<td>82.93%</td>
<td>7</td>
<td>17.07%</td>
<td>41</td>
</tr>
<tr>
<td>Copula be</td>
<td>51</td>
<td>98.08%</td>
<td>1</td>
<td>1.92%</td>
<td>52</td>
</tr>
</tbody>
</table>

Although contracted auxiliaries might be of interest for the analysis of DNS, I will not focus on them in my thesis; this might be an avenue for future research.

\textsuperscript{76} See also Section 4.4.1.
4.4.3.1.1 Comparison with Ihsane’s findings

Given that Ihsane (1998) only looks at root clauses in her analysis, it would be logical to compare the findings of both corpuses for this clause type here. Before comparing the data it is relevant to calculate the distribution figures for the null subject patterns attested in my corpus. Thus, among 136 null subjects 128 (94.11%) occur with a lexical verb, there are 7 (5.15%) null subjects attested with an auxiliary and only one copula be (0.74%). According to Ihsane, in her corpus out of 111 null subjects 96 (86.49%) are attested with a lexical verb, 13 (11.71%) occur with an auxiliary and only 2 of them (1.8%) with a copula be. As we can see, the distribution rate according to the verb types is not relevantly different for the two corpuses. It should be noted that in Truman’s Diary the rate of omission of null subjects occurring with lexical verbs is higher than in V. Woolf’s Diary while in case of the auxiliaries and copula be the numbers in my corpus are lower as opposed to Ihsane’s corpus.

4.4.3.2 The verb type: coordinate root clauses

This section focuses on the distribution of overt and null subjects with different verb types attested in the coordinate root clauses. (150) provides examples of the three verb types with a null subject in the clauses coordinated with roots. Lexical, auxiliary verbs and copula be are presented in (150a), (150b) and (150c) respectively.

(150) a. Ø Visited Monroe’s Ashland after the festivities and Ø enjoyed it very much. (Truman Diary, 1947, 3 July)
   b. Ø Confirmed him by unanimous consent and Ø did not even refer his nomination to a committee. (idem, 8 Jan.)
   c. The new pilots were rattled on account of the passenger and Ø were careful & conservative. (idem, 29 July)

Table 22 provides the details of the subject patterns occurring with the three verb types in the clauses coordinated with roots. As it was the case with the root clauses, in this clause

---

77 As already noted, Ihsane also examines two categories which I do not look at: verb tense and whether a null subject is attested with a positive or a negative verb. I do not consider these categories as it is not quite clear in what way they can affect the DNS distribution.
type again null subjects attested with lexical verbs form a majority. Among the total of 99 finite verbs 78 are lexical, 62 out of which (79.49%), have a null subject. There are 14 auxiliaries 5 (35.71%) of which are attested with a null subject. Among the 7 cases of copula be 3 (42.86%) occur with a null subject. Thus, in this clause type again the lexical verbs are attested with an implicit subject more frequently than with the two other verb types. In coordinate roots the ratio of null subjects attested with auxiliaries and copula be is higher as compared with the simple root clauses.

Table 22. The verb type: coordinate root clauses

<table>
<thead>
<tr>
<th></th>
<th>Overt</th>
<th>%Total</th>
<th>Null subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical verbs</td>
<td>16</td>
<td>20.51%</td>
<td>62</td>
<td>79.49%</td>
<td>78</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>9</td>
<td>64.29%</td>
<td>5</td>
<td>35.71%</td>
<td>14</td>
</tr>
<tr>
<td>Copula be</td>
<td>4</td>
<td>57.14%</td>
<td>3</td>
<td>42.86%</td>
<td>7</td>
</tr>
</tbody>
</table>

As already mentioned above, contracted auxiliaries may play a significant role in that they do not allow a subject to be null. There are two cases of a contracted copula be attested in this clause type. Table 23 makes it obvious that after recalculation of the figures the rate of subject omission with copula be is considerably higher than that of the implicit subjects which occur with an auxiliary. Nevertheless, even after refining the data, the null subjects attested with lexical verbs in this clause type prevail.

Table 23. The verb type: coordinate root clauses: refined data

<table>
<thead>
<tr>
<th></th>
<th>Overt</th>
<th>%Total</th>
<th>Null subjects</th>
<th>%Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical verbs</td>
<td>16</td>
<td>20.51%</td>
<td>62</td>
<td>79.49%</td>
<td>78</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>9</td>
<td>64.29%</td>
<td>5</td>
<td>35.71%</td>
<td>14</td>
</tr>
<tr>
<td>Copula be</td>
<td>2</td>
<td>40%</td>
<td>3</td>
<td>60%</td>
<td>5</td>
</tr>
</tbody>
</table>
4.4.3.3 The verb types: coordinate embedded clauses

This section provides the details on the distribution of null and overt subjects attested with the three verb types in coordinate embedded clauses. (151) provides a number of examples of null subjects with different verb types attested in the coordinate embedded clauses. Lexical verbs are presented in (151a), auxiliaries in (151b) and copula be in (151c).

(151) a. Mrs. Roosevelt came in at 3 P.M. to assure me that Jimmy & Elliot had nothing against me and Ø intended no disparagement of me in their recent non-edited remarks. (Truman Diary, 1947, 3 Jan.)
   b. Ø Said he'd got his checks mixed up, Ø had lied to the Secret Service and he wanted to tell me before his boss did. (idem, 16 Jan.)
   c. Ø Said she didn't like Byrnes and Ø was sure he was not reporting Elliott correctly. (idem, 3 Jan.)

Table 24 illustrates that among the coordinate embedded clauses 7 have a lexical verb, 3 out of which occur with a null subject. 6 have auxiliaries; 4 of them have a null subject. 1 of the remaining two auxiliaries attested with an overt subject is contracted. There are 2 instances of copula be, 1 occurs with an overt subject and is not contracted, the other one is attested with a null subject. Given the small numbers I do not give the percentages.

In this clause type, as opposed to the roots and coordinate roots, auxiliaries attested with a null subject prevail. Null subjects are least attested with lexical verbs.

After removing the contracted auxiliary the refined figures further prove that in this clause type the rate of the distribution of null subjects is highest with auxiliaries.

Table 24. The verb type: coordinate embedded clauses

<table>
<thead>
<tr>
<th></th>
<th>Overt</th>
<th>Null subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical verbs</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Copula be</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
4.4.3.4 Concluding remarks

To sum up, the data in the corpus show that in roots and coordinate roots null subject patterns are mostly attested with lexical verbs while in coordinate embedded clauses auxiliaries occurring with implicit subjects prevail. Only 5 instances of copula be out of the total of 68 are attested with a null subject. The findings in both my and Ihsane’s corpus point in the same direction: null subjects are mostly attested with lexical verbs. On closer examination of the corpus, we can see that, by and large, in the sample diary the lexical verbs overweigh auxiliaries and copula be irrespective of the fact whether the sentence has a subject or lacks one. It follows that the higher frequency of occurrence of lexical verbs with null subjects cannot be relevant for the subject ellipsis and I assume that the verb type cannot be relevant for subject omission.

4.5 Preposing

This section investigates whether the null and overt subject patterns have preposed material and observes whether the DNS in the given diary are sensitive to the type of material which can precede them. Specifically, it examines whether the data in the corpus are compatible with Haegeman’s proposals discussed in Chapter 2, i.e. whether there is an adjunct/argument asymmetry with respect to the preposed material of the diary null subjects. With this in mind, I study the distribution of the referential and expletive overt and null subjects in the following contexts when the subject occurs with:

- no preposed material (152a)
- a preposed adjunct (152b)
- a preposed argument (152c)
- a preposed wh- element (152d)
- subject-auxiliary inversion (SAI) (cf. 152d)

Below I give examples with an overt subject.

(152) a.  I bought this book yesterday.
    b.  Yesterday I bought this book.

114

d. *What* did I buy yesterday?

As it has been done above, the discussion is conducted according to the three clause types: roots, coordinate roots, coordinate embedded clauses. Given that null subjects in diary registers are treated on a par with a non-overt pronoun, as already discussed previously, for all the abovementioned categories I consider how many of the overt subjects are pronouns. This is important for observing the distribution of the pronominal subjects with respect to the preposed material.

4.5.1 Preposed constituents: roots

In this section first I discuss the cases when a null subject in a root clause in the corpus co-occurs with a preposed constituent. Then I give an overview of both overt and null subjects which are attested with fronted material. For this clause type I have observed cases like (152) given above. After studying the whole corpus, among 136 null subjects 12 cases have been identified when a null subject in a root clause co-occurs with what might be considered fronted material. (153) illustrates all 12 examples:

(153) a. *At 12:45* had the G. D. message in shape. (Truman Diary, 1947, 5 Jan.)
b. *Sure* is fine to have them back. (idem, 6 Jan.)
c. *About nine or nine thirty* see Popocatetepel… (idem, 4 March)
d. *Never* saw such crowds—such enthusiasm. (idem, 4 March)
e. *Tuesday morning* lay a wreath on soldiers monument with lots and lots of ceremony. (idem, 4 March)
f. *Never* saw anything lie it… (idem, 6 March)
g. *Of course on the 4th* had to take the plaudits of the populous [sic] but outside that no in convenience. (idem, 3 July)
h. *At 3:30 today* had a very interesting conversation with Gen[eral] Eisenhower. (idem, 25 July)
i. *Finally* recieved [sic] bill at air port. (idem, 26 July)
j. *At 1:30 Washington* time recieved [sic] message my mother has passed on. (idem, 26 July)
k. *Then* went in to the Cabinet Food Committee meeting. (idem, 30 Sept.)

l. *Then* had to decide the argument between Charlie & John and Clark & Matt. (idem, 1 Oct.)

Observe, however, that in the examples above it is not obvious that all the italicized constituents occupy a position to the left of the canonical subject position. In some cases the adjunct might actually occur clause-medially, which would mean that, in fact, there is no fronted material. Specifically, if the clauses given in (153 b, d, f, I, k, l) had an overt subject, then we might have patterns as illustrated in (154). I will not include the examples given in (154) in my analysis.

(154) a. I _never_ saw such crowds-such enthusiasm.
    b. It sure is fine to have them back.
    c. I _never_ saw anything like it.
    d. I _finally_ received [sic] bill at air port.
    e. I _then_ went in to the Cabinet Food Committee meeting.
    f. I _then_ had to decide the argument between Charlie & John and Clark & Matt.

Table 25 presents the findings on the realization of subjects with preposed elements in root clauses attested in the sample corpus. Given the small numbers of the attested subject patterns which definitely occur with preposed material, I do not give the percentages for this or for the two other clause types. As one can see, there are no cases with preposed arguments or with preposed *wh*-elements with null subjects in the corpus. There are 2 instances of subject-auxiliary inversion; both of them occur with an overt subject. Thus, as it becomes clear, the data attested in the corpus are fully in line with Haegeman’s analysis discussed in Chapter 2, i. e. there are no instances of fronted arguments and *wh*-elements and no cases of SAI attested with DNS.

Looking at the preposed adjuncts, one can observe that 128 out of 360 clauses which occur with no fronted material\(^78\) have an implicit subject, the remaining 231 clauses have an overt subject 146 of which are pronouns. The general figures show that, roughly, 1 out of 3 clause-initial subjects are null (128/360). A further breakdown of the overt subjects into

---

\(^{78}\) In the sample corpus there are 12 instances when an overt subject is preceded by an element, such as *well, of course, anyway, well of course, so*. There is only 1 case of *so* attested with a null subject. Given that there is only one example of a null subject, I do not consider these cases and do not include them in the analysis presented in this section. Observe that the 6 examples in (154) have been left out from the data presented in Table 25 as well.
lexical NPs and pronouns shows that the distribution ratio of null subjects in the clauses with no preposed material is 128/275 which means that nearly the half of the pronominal subjects which have no preposed constituents are implicit. Then, among the 20 examples with a preposed constituent, there are 6 instances in which a null subject occurs in a sentence with a preposed adjunct. 14 clauses have an overt subject of which 6 are pronouns. It turns out that 1 out of 3 subjects with a preposed adjunct is non-overt (6/20). Refining the results and considering only overt pronominal subjects we get that there are 6 null subjects and 6 overt pronominal subjects. It appears that 1 out of 2 pronominal subjects with a preposed adjunct is null (6/12). Thus, again in line with Haegeman’s findings, I conclude that adjunct preposing does not influence on the realization of a subject: the subject can be null with a fronted adjunct.

Table 25. Preposed constituents: roots

<table>
<thead>
<tr>
<th></th>
<th>Lexical NPs</th>
<th>Overt pronouns</th>
<th>Null subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>With no preposing</td>
<td>85</td>
<td>146</td>
<td>128</td>
<td>360</td>
</tr>
<tr>
<td>with a preposed adjunct</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>with a preposed argument</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>with a preposed wh- element</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Subject-auxiliary inversion</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

4.5.1.1 Comparing the results with Ihsane

Again as it was done in the previous sections, given that in her analysis Ihsane focuses only on root clauses it would be appropriate to compare the findings of the two diaries here. Observing whether the null subjects are the first elements in the clause or if there are any constituents to their left Ihsane finds that among the 111 clauses only 11 (9.91%) occur with a preposed constituent, all of which are adjuncts.

To compare the results of both corpora it is necessary to look at the general figures: 6 (4.26%) out of 141 null subjects attested in roots have a fronted adjunct. Obviously, the null
subjects with a preposed adjunct have a slightly higher rate of occurrence in V. Woolf’s Diary as compared with Truman’s Diary. Note, however, that given the low figures no conclusions can be drawn from them.

As before I would like to point out that Ihsane’s data remain incomplete, since there is no information on the occurrence of overt subjects in the contexts with preposing. We also do not have information concerning the relative frequency of adjunct preposing and argument preposing with overt pronominal subjects, so there is no way we can assess her findings.

4.5.2 Preposed constituents: coordinate root clauses

To examine whether implicit subjects in coordinate root clauses demonstrate certain constraints with respect to the material to the left of the canonical subject position. For this purpose, I have collected examples like (155), which illustrates an example of a preposed adjunct (155a), a preposed argument (155b), a preposed wh-element (155c).

(155) a. I talked to Mary and after she left (I) went to bed.
   b. I met Mary but Tom *(I) did not meet.
   c. What will you do and where will *(you) go?

As discussed in Section 2.5, in the core grammar of English adjunct preposing is compatible with subject omission in a coordinate clause but argument preposing and subject-auxiliary inversion are not. In this section I examine whether any patterns like (155 a, b, c) are attested in my corpus and whether coordinate clauses in Standard English and in diary registers differ with respect to the material which can precede a null subject. As it was done for roots, first I examine the instances of preposed constituents which co-occur with null subjects, then I provide the figures both for overt and null subject patterns. Among the 70 coordinate null subjects in the corpus there are 9 cases with what might be considered fronted material. The examples are illustrated in (156).

(156) a. R[ea]r Adm[iral] Foskett came to the House with me and then went home.
   (Truman Diary, 1947, 2 Jan.)

See also Haegeman 2011: 25-26.
b. Took an "electric" shave (practically none) and then went walking at 8 A.M. with … (idem, 5 Jan.)
c. Arose at 5:45 A.M. [,] read the papers and at 7:10 walked to the station to meet the family. (idem, 6 Jan)
d. Never saw anything like it and never expect to again[.] (idem, 6 March)
e. Finally the Chaplain pointed to his insignia and informed the tough Nazi cop that he belonged to the religious section of the army and finally remarked… (idem, 27 June)
f. Held a reception before speaking time and then signed some programs for those who had helped with the arrangements. (idem, 4 July)
g. Arose at 2 A.M., shaved, took a bath and then called Bess & Margaret. (idem, 1 Sept.)
h. Listened to some commentators and then called Dr. Steelman… (idem, 1 Oct.)
i. Went down to the W[hite] H[ouse] garage to see the tree and then ate a tall dinner… (idem, 25 Dec.)

As it becomes apparent from the examples given above all the fronted constituents are adjuncts. Observe that again among these examples there are instances when the adjunct might in fact be taken to occupy a medial position. Close examination of the data reveals that in all the cases except (156c) the italicized adjunct may also be taken to occur clause-medially. The alternative clauses with an overt subject are given in (157).

(157) a. Rear Adm[iral] Foskett came to the House with me and he then went home.
b. Took an "electric" shave (practically none) and I then went walking at 8 A.M. with …
c. Never saw anything like it and I never expect to again[.]d. Finally the Chaplain pointed to his insignia and informed the tough Nazi cop that he belonged to the religious section of the army and he finally remarked…
e. Held a reception before speaking time and I then signed some programs for those who had helped with the arrangements.
f. Arose at 2 A.M., shaved, took a bath and I then called Bess & Margaret.
g. Listened to some commentators and I then called Dr. Steelman…
h. Went down to the W[hite] H[ouse] garage to see the tree and I then ate a tall dinner…
Thus, if I leave out all these clauses, it will mean that in the corpus there is only one clause in which the null subject has a preposed adjunct.

Table 26 shows that in coordinate root clauses there are 2 examples with a preposed adjunct: 1 occurs with an overt subject which is a pronoun and there is 1 instance with a null subject. There are no instances of fronted arguments and wh-elements. There is 1 case with subject-auxiliary inversion which occurs in a clause with an overt subject. In the core register subject omission in such examples would lead to an unacceptable result (cf. 155c)\(^80\).

Given the small figures any conclusions with respect to the null subject preposing in coordinate clauses can be tentative. It is essential to note, however, that with respect to preposing the attested data in coordinate clauses do not show any properties specific to diary registers: the attested data are grammatical in Standard English.

**Table 26. Preposed constituents: coordinate root clauses**

<table>
<thead>
<tr>
<th></th>
<th>Lexical NPs</th>
<th>Overt pronouns</th>
<th>Null subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>With no preposing</td>
<td>7</td>
<td>21</td>
<td>69</td>
<td>97</td>
</tr>
<tr>
<td>with a preposed adjunct</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>with a preposed argument</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>with a preposed wh-element</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>with subject auxiliary inversion</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

4.5.3 Preposed constituents: coordinate embedded clauses

For embedded coordinate clauses I have looked at the following patterns:

- instances of coordination in which the second conjunct is introduced by a subordinating element such as *that* (158 a, b, c)

\(^80\) See also Section 2.5.
- instances of coordination in which the second conjunct is an embedded \textit{wh} - clause or an \textit{if} clause (158d)
- instances which do not have a subordinating element (159)

(158) a. She said that she talked to Mary and that* (she) left at four.
    b. She said that she talked to Mary and that after she left *(I) went to bed.
    c. She said that she met Mary but that Tom *(I) did not meet.
    d. She asked what I would do and if/where *(I) would go.

In the core grammar of English none of these would allow for a null subject. As I have already discussed in Section 4.2, the examples in (158) are ungrammatical because the subject cannot be implicit with an overt complementizer (cf. (116c))\textsuperscript{81}. Besides, in case of (158c) the subject should be overt with a preposed argument as discussed in Chapter 2. Observe, however, that in the core grammar of English (158a, b) would be grammatical with a non-overt complementizer (159 a, b) but (158c) wouldn’t because a preposed complement requires an overt subject (159c).

(159) a. She said that she talked to Mary and (she) left at four.
    b. She said that she talked to Mary and after she left (I) went to bed.
    c. She said that she met Mary but Tom *(I) did not meet.

Table 27 makes it clear that in my sample corpus among the 15 attested coordinate embedded clauses 7 have an overt subject all of which are pronouns. 5 out of the 7 clauses are introduced by an overt complementizer \textit{that}, in the remaining 2 clauses \textit{that} is implicit. There is one instance when an overt subject has a fronted adjunct, the clause has a non-overt complementizer. There are no cases with a preposed argument and \textit{wh}-element. Among the 8 clauses with a null subject 4 are \textit{that}-clauses, 3 are \textit{wh}-clauses and 1 is introduced by \textit{as if}. In all 8 cases the subordinator is \textbf{non-overt}. There are no instances of preposed adjuncts, complements or \textit{wh}-elements attested with non-overt subjects. Thus, as the data show, none of the constraints illustrated in (158) and (159c) have been found in Truman’s Diary: in the corpus there are no attested patterns which would not be allowed in Standard English.

\textsuperscript{81} See also Quirk et al. (1972: 555-56, 574-76).
Table 27. Preposed constituents: coordinate embedded clauses

<table>
<thead>
<tr>
<th></th>
<th>Overt pronouns</th>
<th>Null subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>With no preposing</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>With a preposed adjunct</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>With a preposed complement</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>wh-/if clause with an overt subordinator</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>wh-/if clause with a non-overt subordinator</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>That-clause with an overt subordinator</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>That-clause with a non-overt subordinator</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

4.5.4 Concluding remarks

The data in my corpus are in line with Haegeman’s (1997, 2007, 2011) proposals presented in Chapter 2: DNS are compatible with adjunct preposing but do not occur with a fronted argument, wh-element and SAI. Furthermore, both coordinate roots and coordinate embedded clauses do not demonstrate specific patterns with respect to preposing: the patterns attested in the diary and in core grammar are similar.

However, as it has been already mentioned above, subject omission in coordinate clauses in diaries and in core grammar do differ in terms of the principle of co-referentiality: in core grammar the second conjunct of a coordinate clause can have a null subject if the subject is co-referential with that of the first conjunct; in diaries the principle of co-referentiality does not always apply: the subject of the second conjunct can be implicit even if it is not co-referential with that of the first conjunct. I discuss this in the following section.

4.6 Coreferentiality of subjects in Standard English and in diaries

The data in my corpus have proved that subject ellipsis in coordinate clauses in diaries and in Standard English are not identical. While the core register allows for a coordinate subject to
be null only if it is co-referential with the subject of the first conjunct, diaries demonstrate a more liberal pattern of subject ellipsis. The examples in (160) are from the Guardian. The common feature of all these sentences is that the null subject in the coordinate clause is co-referential with the subject of the first conjunct and can always be recovered (it’s always got an antecedent in the first clause). Co-reference with another salient antecedent in the discourse would lead to ungrammaticality. In (160 d), e.g., the null subject cannot be co-referential with the NP *opposition* but rather should be interpreted as identical to the move.

(160) a. China immediately questioned the US move and φi said it deserved further scrutiny.  
http://www.guardian.co.uk/world/2011/nov/17/obama-asia-pacific-address-australia-parliament?INTCMP=SRCH

b. The Observatory defends human rights in Syria and φi does not adopt any political stances.  

c. Immigration halls will be full from 7am and φi might exceed capacity under fire regulations.  
http://www.guardian.co.uk/politics/2011/nov/25/border-agency-passport-checks?INTCMP=SRCH

d. The opposition said the move represented a U-turn and φi/*j* was in effect a revised version of Labour's future jobs fund, cancelled by the coalition government when it came to office.  

e. Each subsidy is worth half the current youth national minimum wage and φi will last for six months.  

---

82 See also Quirk et al. (1972: 555).
f. The ECB is independent and $\phi$ is responsible for monetary policy alone.


While its overall appearance is in line with the core grammar of English, subject omission in coordinate clauses does present some specific properties in the diary register. A detailed analysis of the corpus reveals that the non-coreferential coordinate null subjects are not limited to fictional diaries as discussed in Chapter 2, though remaining a minority, they occasionally do occur in non-fiction diaries, too, as it was made obvious in Section 4.4.2.3. Among the total of 70 coordinate clauses attested with a null subject in my corpus 2 cases have been identified in which the subject of the second conjunct is omitted even though it is different from that of the first conjunct. These examples presented in (143) are repeated here in (161).

(161) a. Doc tell's [sic] me I have Cardiac Asthma! Ain'\['\]t that hell. Well it makes no difference,] $\varnothing$ (I) will go on as before. I've sworn him to secrecy! So What! (idem, 7 March)

b. Arrived in Grandview about 3:30 CST[,] went to the house and met sister & brother. Went to Belton with them and $\varnothing$ (we) picked a casket. (idem, 26 July)

Let us discuss them in turn.

In (161a) the subject in the first conjunct is $it$ and that in the coordinate clause can be interpreted as the first person singular. The omitted $I$ can be retrieved from salient referents which are available both in the preceding and the following clauses. Again, in the core grammar this kind of subject omission is not possible. As noted for (160d), in Standard English the null subject must be co-referential with the subject of the preceding conjunct while co-reference with another salient referent in the discourse is not allowed. This rule is not respected in my example.

In (161b) both conjuncts lack a subject. Given the information in the preceding clause [...] went to the house and met sister & brother, and the string with them following the
omitted subject, the empty category in the first conjunct can be identified as being the first person singular, *I*. In the second clause, though, it is clear from the context that the omitted subject should be recovered as *we*. The implied subject *we* can be retrieved due to the salient referent in the discourse context, *sister & brother*, which occurs in the preceding sentence: [*…* went to the house and met sister & brother]. Besides, the string *with them* which occurs in the first conjunct allows one to understand that the following action should involve a group of people, in this case, *we*. Observe that in Standard English the first conjunct always has an overt subject while in diaries, as the example shows, this rule does not always apply. Besides, in core grammar the subjects of both conjuncts should be identical for the subject of the coordinate clause to be null. In diaries this rule is not respected either.

These two cases display a marked pattern of subject omission in coordinate clauses which is not acceptable in the core grammar. In the diary style subject omission in coordination is apparently possible even when the subject is not co-referential with that of the preceding conjunct. However, the omission is not completely free in that in both cases in (161) the subject in the coordinate clause can be recovered from a broader context. This is a condition which we also identify in the case of root subject ellipsis. In all other instances of subject omission in coordinate clauses attested in my corpus, the subject of the second conjunct is overt whenever it is different from the one of the first conjunct.

As mentioned in Chapter 2, Haegeman (2002b: 141-2) reports that such cases have also been observed by Becquet (2000). In her analysis of *Bridget Jones’s Diary* Becquet gives 8 examples out of 184 with a non-coreferential null subject in the second conjunct. What is of interest here is that, according to Haegeman (2002b: 142), Becquet observes that in diaries where no embedded null subjects are attested, patterns like SU&Ø and Ø& Ø do not occur. SU means that the subject is overt and Ø means that it is omitted. My corpus findings challenge Becquet’s observation and reveal that the two patterns DO occur in non-fiction diaries which lack embedded null subjects. Note that the relative proportion of such cases that I have found (2/70) is lower than Becquet’s ratio of 8/184.

To sum up, in the diary material I have examined subject omission in coordination mostly respects the general co-referentiality condition on the subject ellipsis typical of the core grammar of English. There are only 2 exceptions out of 70 to this generalization. In these cases I assume that the subject ellipsis in coordinate clauses where the null subject is not co-referential with that of the first conjunct is generally available as a register-specific property of diaries, which can occur both in fictional and in non-fiction diaries, contrary to Haegeman’s (2002b) claim.
Observe also that coordinate subject ellipsis in diary registers demonstrates an extra possibility which is not available in core grammar. It follows that the analysis for ordinary coordination cannot apply to these cases and a new account should be elaborated to suffice for the new pattern. This can be an avenue for future work.

4.7 Discourse factors of DNS

4.7.1 The position of the subject in the diary entry

a) The first position in the entry

The relative position of the sentence within the diary entry seems to play a part in determining subject omission. In my sample there are 42 entries. The inspection of the first finite sentences of the entries makes it clear that first person singular subjects are always omitted at the beginning of an entry, first person plural subjects can be both overt and non-overt whereas the third person subjects are never omitted. More precisely, in 25 out of the total of 42 diary entries that constitute my corpus, the omitted subject of the opening clause can be recovered as I. (162). There are 2 entries with an implicit we (163). There are no examples of the first finite sentences with an overt I as the subject, 3 entries have an overt we in the first finite clause of the entry (164) and 8 start with an overt third person subject (165). 7 out of these 8 third-person subjects are NPs and 1 is realized by the demonstrative pronoun this (166). 3 January entry begins with Byrnes and I but observe that the non-overt realization of I, like Byrnes and Ø, would be ungrammatical (167). The openings of the remaining 4 entries aren’t finite clauses, in all 4 cases, subject + auxiliary ellipsis has applied83 (168 a, b, c, d).

(162) Ø (I) Went to Bethesda to see Bill Hassett. (Truman Diary, 1947, 25 Dec.)
(163) Ø (We) Landed in Washington aboard the Williamsburg at 8:00 A.M. (idem, 20 Sept.)
(164) We have had a grand cruise aboard the Battleship Missouri. (idem, 19 Sept.)
(165) My sister, Mary Jane, called & said that mamma is sinking swiftly. (idem, 26 July)

83 Recall that I have not included these patterns in my analysis.
This was a terrible day. (idem, 28 July)

Byrnes & I discussed General Marshall’s last letter and decided to ask him to come home. (idem, 3 Jan)

A terrific day. (idem, 7 Jan.)

Fiesta! (idem, 6 March)

Meeting with "big six" in study at White House. (idem, 24 March)

Meeting with Argentine Ambassador. (idem, 31 March)

As it turns out, in about 60% of the total number of the entries I is identified as the subject of the opening clause. Looking only at the entries which have a null subject it appears that in 25 among 27 entries which have an opening clause with a null subject, the implicit subject refers to the narrator. Why can this be? I will return to this point below.

First, let us consider (169) which illustrates the entry for 5 January. Even not having read the whole entry, it is obvious that the subject of the clause Spent all morning with the State of the Union Message should be recovered as I because there are no other salient referents in the context. The only potential antecedents in the discourse are the NP Jim Rowley[,] Chief of the White House S[ecret] S[ervice] detail or the NP a couple of more men but the reader will agree that they can hardly qualify as the subject of the first and the subsequent clauses given the fact that diaries refer to the narrator, i.e. I.

Spent all morning with the State of the Union Message. Ø Went to sleep at 12:15 last night or this morning reading it. Ø Slept until 7:30-most unusual. Ø Get up nearly every morning at 5:30 or five minutes to six. Ø Took an "electric" shave (practically none) and then Ø went walking at 8 A.M. with Jim Rowley[,] Chief of the White House S[ecret] S[ervice] detail and a couple of more men following. And some in a car following along behind. I'm not supposed to know about the car. Ø Went down F St[.] and back G. Ø Like to look in merchants['] windows. Ø Had breakfast at 9 A.M. At 12:45 Ø had the G. D. message in shape. Ø Read & Ø reread. Ø Spent the afternoon in study on the same message and the Economic one too.

It seems plausible to suggest that I has a high rate of deletion because it is the default topic of a diary and it is actively involved in the discourse, as suggested in Chapter 1. Hence,
it is easily understood in the context even if non-overt. Besides, basing on the fact that the diary is about the narrator, i.e. I, the latter might be considered a non-informative element, which, thus, can be omitted. This might have an implication for Grice’s Maxim of Quantity\textsuperscript{84} given in (59) and repeated here as (170). Refering to the discussion in Section 1.5.2, I suggest that diary null subjects are omitted to avoid redundancy.

(170) 1. Make your contribution as informative as is required (for the current purposes of the exchange)

2. Do not make your contribution more informative than is required.

In the 2 cases when the first clause of the entry lacks a subject identified as we the subject can be recovered due to a discourse antecedent. In one of them (the entry for March 4) we can be recovered from a salient referent available in the preceding entry. (171 a, b) present the entries.

(171) a. Spend a pleasant day.

Go to bed and get called at 2:30 A.M. Tuesday.

It is a nice morning. But we run into clouds over Texas and Okla[homa].

(Truman Diary, 1947, 3 March)

b. Ø Come into sight of Monterey [sic] after the sun had been up an hour or two.

Country looks like a map. About nine or nine thirty Ø see Popocatetepitl and Ø try to see Orizaba-haze too thick[,]Ø can’t see it. Approach rim around Tenochtitlan Valley-up 11000 feet. No discomfort. Beautiful valley. Ø Must have been lovely when a lake. Too bad the Spaniards drained it. Ø Made a lot of dust.

Perfect day. Ø Land at 10:00 on the dot. My pilot never misses a schedule. Ø Step down from plane. Mexican President comes down steps of observation tower at same time. We meet. I like him at once. He introduces his Cabinet, I introduce my secretaries and aides. (idem, 4 March)

As we can see from the discussion above, since the diary is about the writer (=I), the first person singular omissions can “arise out of the blue” (Haegeman 2011: 6) and though non-overt, I can be easily understood. The pronominal subject denoting other person categories

\textsuperscript{84} See Section 1.5.2.
(in this case *we* as we have seen), on the other hand, is omitted because it can be recovered from a salient referent available in the discourse context. Thus, recoverability plays a significant role for the subject ellipsis in diaries.

Next, I observe the cases when *I* is overt in the middle part of the entry. Given that the first person singular is the default omission pattern, as the data show, it will be interesting to see what determines its overt realization.

**b) The middle part of the entry**

In this section I observe the cases when *I* is overt in the middle part of the diary entries. The data in my corpus show\(^{85}\) that *I* is the most frequently omitted category both in the first and in the in the middle parts of the entries. What is of interest here is when Truman DOES use *I*.

In the whole corpus among 647 finite clauses, which occur in the middle part of an entry, 454 have an overt subject and there are 193 clauses with a null subject. Out of 454 overt subjects 101 can be identified as *I*. I will be studying these 101 instances here.

Among the 101 instances I consider how many:

- occur in an embedded clause
- have a preposed argument
- occur with SAI
- are attested with a contracted verb form
- are overt to avoid ambiguity.

Observe that the occurrence of overt subjects in the first four categories would be conditioned by syntactic factors. Recall that, as the discussion in Chapter 2 makes it obvious, diary null subjects are not available in these environments\(^{86}\) and Truman’s Diary further confirms Haegeman’s proposal that DNS are a root phenomenon: DNS are not compatible with embedded clauses; there are no null subjects occurring with a preposed argument, SAI or a *wh*-element. With a contracted verb the subject cannot be omitted as there will be no host for the clitic (*'ve, 'll, 's*)\(^{87}\).

\(^{85}\) See Sections 4.4.2, 4.4.3.

\(^{86}\) See Section 1.3.1.3 and the discussion in Chapter 2.

\(^{87}\) See Section 4.4.1.
Observation reveals that in 23 cases I is overt because it occurs in an embedded clause (172). There no cases when I has a preposed argument. There is one case when I is attested with SAI (173).

In 18 cases I is followed by a contracted verb (168).

(172) Both were up when I called them. (Truman Diary, 1947, 1 Sept.)
(173) So am I. (idem, 20 Sept.)
(174) I've read thousands of messages from all over the world in the White House study and can shed tears as I please-no one's looking. (idem, 28 July)

There are two instances when the matrix clause occurs medially in the subordinate clause. I is the subject of the matrix clause and in one of the instances it follows the subject+predicate of the subordinate clause (175a), in the second case it is preceded by the subject of the subordinating clause (175b).

(175) a. Marshall is, I think[,] the greatest man of the World War II. (Truman Diary, 1947, 8 Jan.)
   b. The Jews, I find are very, very selfish. (idem, 21 July)

Apparently, in the 48 cases mentioned above (nearly the half of the cases when the first person singular subject is overt) the overt realization of I is determined by syntactic factors. Let us see what conditions the overt use of I in the remaining 53 cases.

After analyzing the corpus, it becomes clear that in 39 cases the first person pronoun is used when the subject of the following/preceding sentence changes and, hence, the subject is not co-referential with the subject of the following/preceding clause. Hence, I tentatively conclude that the pronoun is used to avoid ambiguity (176).

(176) a. The papers this morning are full of Marshall's appointment and Mr. B[yrne]'s resignation. I am very sorry Mr. Byrnes decided to quit. (Truman Diary, 1947, 8 Jan.)
   b. My sister, Mary Jane, called & said that mamma is sinking swiftly. Dr. Greene was at home in Grandview and said she'd not last long. Call was at 9 A.M. Washington time. I ordered plane set up at 12:30. (idem, 26 July)
Finally, we get 13 instances when Truman uses *I* when it, in fact, could be omitted. In (177a) the reflexive pronoun *myself* implies that the subject of the clause should be *I*: the omitted subject binds the reflexive pronoun. One can suggest that Truman uses the first person pronoun for emphasis. In the second example (177b) both the coordinate and the preceding clause have an overt co-referential subject. It is not clear why the subject of the coordinate clause might be overt.

(177) a. Foreign Minister and the Commandant of the Cadets wept—so did news men and photographers. *I* almost did myself. (Truman Diary, 1947, 4 March)

b. *I've* turned them over to Steelman[,] Harriman, Snyder and Schwellenbach, and *I* hope for the best. (idem, 4 Jan.)

Examination of the data makes it obvious that first person singular subject is deleted because being the default topic of the diary, it can be easily recovered. Cases like (172-75) make it obvious that the recoverability of a null subject is not only conditioned by pragmatic factors and that syntax plays a crucial role, too. With respect to the other categories it can be suggested that they are deleted because they can be retrieved due to a salient antecedent in the discourse context, as observed in case of *we*-omissions. For third person singular subjects the morphological ending in present simple -*e(s)* can be decisive. Finally, the overt realization of *I* pronouns is conditioned both by syntactic and pragmatic factors. Hence, I can again reiterate that both syntax and pragmatics are relevant for understanding the phenomenon of DNS.

4.8 Concluding remarks

The analysis of the Truman Diary entries (1947) has revealed that the corpus data are in line with Haegeman’s (1990, 1997, 2007, 2011) analysis. The findings further prove Haegeman’s claim that DNS are a root phenomenon: DNS are attested in simple roots while the entire corpus data do not show availability for subject ellipsis in embedded and in interrogative clauses.

Then, as it became apparent, the null subjects in the sample corpus are sensitive to the material which can precede them: there is an apparent adjunct/argument asymmetry in all the clause types observed.
The distribution of subject patterns has been studied according to the clause type they occur in. I have identified two types of subjects: referential and expletive. Although the overall number of the referential subjects prevails that of the expletives, the observation reveals that in Truman’s Diary the relative rate of deletion of the expletive subjects is higher as opposed to the referential subjects.

On closer examination of the sample corpus it becomes clear that 220 out of 685 finite clauses have an implicit subject, which means that roughly 1 out of 3 clauses is attested with a null subject. The vast majority of implicit subjects are attested in simple roots: 141/220, while coordinate clauses rank second: 79/220.

Given that the DNS are treated on a par with a pronoun, I have further refined the data and looked at the relative proportions of the null subjects vs. the overt pronominal subjects. Accordingly, in roots we get 141/305, i.e. roughly nearly the half of the pronominal subjects is implicit.

As compared with roots, coordinate root clauses demonstrate a higher rate of subject ellipsis. It has been tentatively concluded that this can be conditioned by the grammaticality of subject omission in the second conjunct of coordinate clauses in Standard English. Specifically, 70 out of the total of 93 pronominal coordinate subjects are null.

In coordinate embedded clauses the null vs. overt subjects are distributed as follows: 8 vs. 7.

Furthermore, it has been concluded that subject omission in coordinate clauses in diary registers in the large majority of cases is in line with the general rule on subject ellipsis in the core grammar of English: the subject in the second conjunct can be null when it is co-referential with the subject of the first conjunct, both when the clause is coordinated with a root clause and when it is coordinated with an embedded clause. Coordinate roots, however, DO display a register-specific pattern: the rule of co-referentiality does not always apply and in diaries the subject of the second conjunct can be null even if it is not identical with that of the preceding clause. Although this pattern had been attested in a diary register which also licenses embedded null subjects (Becquet 2000), Truman’s Diary has, in fact, proved that the feature DOES also occur in a diary ‘dialect’ where embedded implicit subjects are not attested. This implies that the analysis elaborated for ordinary coordination cannot fully account for DNS: diary registers demonstrate an extra possibility for coordinate null subjects.

The figures both for referential and expletive subjects are presented here.
The corpus does not show availability of other register-specific patterns with respect to coordinate clauses.

It has also become evident that mostly null subjects can be identified as first person singular. I have intuitively concluded that I –omissions are natural in view of the material under study: the diary refers to the narrator and the non-overt I can be easily recovered; it has been assumed that somehow in our mental representation the first person subject is ‘always available’ as an antecedent.

The other person/number categories can in most cases be identified through a salient antecedent in the discourse. Hence, recoverability of subjects is a relevant factor for subject omission.

Based on Grice’s Maxim of Quantity, it has been observed that by and large in diaries subjects are omitted to avoid redundancy and are overtly realized to avoid ambiguity. Based on the findings and the distribution patterns identified, it has been concluded that the phenomenon of subject ellipsis in diaries has both syntactic and pragmatic foundations.

The comparison of Ihsane’s results with my findings reveals that the rate of subject ellipsis in Truman’s Diary is much higher as compared with V. Woolf’s Diary. While Ihsane concludes that null expletive subjects in her corpus are rare, the same generalization does not hold for Truman’s Diary: the rate of omission of the expletives is higher as compared with the referential ones, as noted above. In general terms, both corpuses demonstrate similar DNS distribution patterns:

- null subjects are more frequently attested in coordinate clauses as compared with roots;
- the omission of first person singular category is attested most frequently;
- the majority of null subjects occur with a lexical verb;
- implicit subjects can be preceded by an adjunct but not by an argument,

The findings of the two diaries have allowed me to tentatively conclude that in American diaries the rate of subject omission is higher than in British diaries. Note, however, that, as mentioned throughout the chapter, Ihsane only discusses the properties of the clauses with null subjects and she fails to compare the distribution of null subjects with that of overt pronominal subjects. This did not allow me to conduct more accurate and fine-grained comparative analysis of the two corpuses.
Observe also that given the small figures any conclusions arrived at in this thesis can be tentative.

Chapter 5. Conclusion

In this thesis I have examined the phenomenon of subject ellipsis in diary registers in English. Based on Haegeman’s (1997, 2007, 2011) account for the diary null subjects and Ihsane’s (1998) methodology I have studied Truman’s Diary (1947) and have concluded that:

- DNS are a root phenomenon (the corpus data investigated does not show availability of null subjects in embedded environments);
- diary subject ellipsis is not compatible with *wh*-movement;
- there is an apparent adjunct/argument asymmetry with respect to the material to the left of the canonical subject position: DNS can have a preposed adjunct, but not a topicalized argument.

All the conclusions above indicate that the data in Truman’s Diary are in line with Haegeman’s analysis and do not demonstrate any so-far-unidentified, marked patterns. Hence, Haegeman’s (1997: 233) proposals that diary null subject is “an antecedentless empty category in the A-specifier of the root” and “[n]ull subjects depend on the truncation of CP, which turns the specifier of IP into the highest specifier of the clause” are fully applicable for my corpus data.

With respect to coordinate clauses it has been found that by and large the subject ellipsis in this clause type does not deviate from the patterns which are licit in core grammar. Nevertheless, subject omission in diary coordinate clauses can be more liberal, as the observation reveals: the principle of co-referentiality, which is an obligatory condition for the coordinate subject omission to apply in Standard English, does not always hold for diaries. A coordinate subject in diaries can be implicit even if it is not co-referential with that of the preceding clause. As this finding shows, my prediction that the diary style seems to explore and generalize an option which is already available in core grammar, namely, the subject
ellipsis in coordination, is not borne out. Apparently, diary registers do divert from core grammar and do demonstrate specific patterns which are not licit in Standard English.

As it also appears, the analysis of ordinary coordination cannot account for these data. The diary register has an extra possibility and to suffice for this register-specific pattern a new account for coordinate subject omission would be needed: this can be an avenue for future work.

Furthermore, the comparison of the data of Truman’s and V. Woolf’s diaries has made it clear that the subject ellipsis in American diaries is more frequently attested as compared with British diaries. Note, however, that the comparison of only two diaries cannot be sufficient for drawing such conclusions and for reaching more accurate conclusions further research is required.

In line of the data studied it has been observed that although DNS are subject to syntactic constraints, the role of pragmatics should not be undermined.
### Appendix: Null subjects in Harry S. Truman 1947 Diary

<table>
<thead>
<tr>
<th>Clause</th>
<th>Person</th>
<th>Ref./expl.</th>
<th>Verb</th>
<th>Preposing type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Spent the day at work.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>2. Was aboard the Williamsburg with secretaries …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>cop.</td>
</tr>
<tr>
<td>3. Had a most pleasant evening …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>4. Went to bed at 1:30 A.M. tomorrow …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>5. Arose at 7:30 …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>6. … shaved …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>7. … dressed …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>8. … and had breakfast at 8:15…</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>9. Arrived at the White House about 9 A.M.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>10. Had a beautiful snow the night before.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>11. and then went home.</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>12. Spent the day working on messages.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>13. Called all the members of the Cabinet[,] …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>14. … wished them a happy New Year.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>15. Called Henry Stimson, Miss Perkins …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>16. … and decided to ask him to come home.</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>17. … and intended no disparagement of me …</td>
<td>c.-e.</td>
<td>3pl.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>18. Said she was for me.</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>19. Said she didn't like Byrnes …</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>20. … and was sure he was not reporting Elliott correctly.</td>
<td>c.-e.</td>
<td>3sg.</td>
<td>ref.</td>
<td>cop.</td>
</tr>
<tr>
<td>21. Said Byrnes was always for Byrnes …</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>22. Spent the day working on State of the Union …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>23. Spent all morning with the State of the Union Message.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>24. Went to sleep at 12:15 last night …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>25. Slept until 7:30-most unusual.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>26. Get up nearly every morning at 5:30 …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>27. Took an &quot;electric&quot; shave</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>28. and then went walking at A.M. with Jim Rowley …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>29. Went down F St[,] and back G.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>30. Like to look in merchants['] windows.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>31. Had breakfast at 9 A.M.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>32. At 12:45 had the G. D. message in shape.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
<tr>
<td>33. Read</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
</tr>
</tbody>
</table>

89 The diary is available at [http://www.trumanlibrary.org/diary/transcript.htm](http://www.trumanlibrary.org/diary/transcript.htm).
<table>
<thead>
<tr>
<th>Line</th>
<th>Text</th>
<th>Verb</th>
<th>Subject</th>
<th>Object</th>
<th>Action</th>
<th>Time</th>
<th>Location</th>
<th>Attribution</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.</td>
<td>&amp; reread.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td>Spent the afternoon in study on the same message …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td>Arose at 5:45 A.M.[.]</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td>read the papers</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38.</td>
<td>and at 7:10 walked to the station to meet the family.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td>Took 35 minutes.</td>
<td>root</td>
<td>3sg.</td>
<td>expl.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td>Sure is fine to have them back.</td>
<td>root</td>
<td>3sg.</td>
<td>expl.</td>
<td>cop.</td>
<td>be</td>
<td>adj.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td>and crack all night long.</td>
<td>c.-r.</td>
<td>3pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td>and didn't.</td>
<td>c.-e.</td>
<td>3pl.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43.</td>
<td>and are misrepresented in history …</td>
<td>c.-e.</td>
<td>3pl.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44.</td>
<td>Read my annual message.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.</td>
<td>and will go places.</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td>and said</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td>and told me</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.</td>
<td>and asked</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49.</td>
<td>and told me</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50.</td>
<td>Confirmed him by unanimous consent</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51.</td>
<td>and did not even refer his nomination to a committee.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52.</td>
<td>Went to Nat[ional] Press Club dinner …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53.</td>
<td>Had quite a day.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54.</td>
<td>and discussed Mid East Affairs at some length.</td>
<td>c.-r.</td>
<td>3pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55.</td>
<td>but are suspicious of the British.</td>
<td>c.-r.</td>
<td>3pl.</td>
<td>ref.</td>
<td>cop.</td>
<td>be</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56.</td>
<td>Was sitting at my desk just before dinner tonight…</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57.</td>
<td>… when … came in and asked</td>
<td>c.-e.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58.</td>
<td>Said he'd got his checks mixed up,</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59.</td>
<td>had lied to the Secret Service</td>
<td>c.-e.</td>
<td>3sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.</td>
<td>and promised to help him out.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61.</td>
<td>and like to help 'em …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62.</td>
<td>and help 'em out …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63.</td>
<td>and make 'em speak to me.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64.</td>
<td>See mamma[.]</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65.</td>
<td>Spend a pleasant day.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66.</td>
<td>Go to bed</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67.</td>
<td>and get called a[t] 2:30 A.M. Tuesday.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.</td>
<td>Come into sight of Monterey …</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69.</td>
<td>About nine or nine thirty see Popocatepetl</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.</td>
<td>and try to see Orizaba …</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71.</td>
<td>can't see it.</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72.</td>
<td>Approach rim around Tenochtitlan Valley …</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73.</td>
<td>Must have been lovely when a lake.</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Text</td>
<td>Verb</td>
<td>Number</td>
<td>Text</td>
<td>Verb</td>
<td>Number</td>
<td>Text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------</td>
<td>------</td>
<td>--------</td>
<td>---------------------------------------------------------------------</td>
<td>------</td>
<td>--------</td>
<td>---------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>Made a lot of dust.</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Land at 10:00 on the dot.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>Step down from plane.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>and mispronounce Tenochtitlan …</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>and start for American Embassy.</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Never saw such crowds-such enthusiasm.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Arrive at Embassy[,]</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>bid President goodbye.</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Have dinner at Palace</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Shake hands with some two or three thousand.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>and wave to a sea of people</td>
<td>c.-r.</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Have seen pictures of Franz Joseph …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Tuesday morning lay a wreath on soldiers monument …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>and jumped 200 feet to his death.</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>Had all the cadets lined up</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>Never saw anything like it</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>and never expect to again[,]</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Left Mexico City at 6 A.M.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>Land at Waco in the rain at 11 A.M.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>will go on as before.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>Hope for renewal.</td>
<td>root</td>
<td>1pl.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Invited Dean Atcheson [sic], Tom Connolly [sic] …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>Told the Ambassador …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>Informed him …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Said Argentine [sic] wanted to get along with us, etc.</td>
<td>root</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Called in Sec[retary] of State, Gen[eral] Marshall…</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Lunched with Marshall &amp; Att[orne]y. Gen[eral] Clark…</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>and tried his level best to be decent.</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>and informed the tough Nazi cop …</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>and finally remarked…</td>
<td>c.-r.</td>
<td>3sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>Was driven to Charlottesville, V[irginia] at 2 P.M.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>aux.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>Stayed at Stanley Woodward's farm…</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>Had a most delightful week end.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>Of course on the 4th had to take the plaudits of …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>adj.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>Visited Monroe's Ashland after the festivities …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>and enjoyed it very much.</td>
<td>c.-r.</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>Went back to the N[orth] of V[irginia] at 5 P.M.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>Had a very pleasant time.</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>Had most cordial reception at Jefferson's home…</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>Held a reception before speaking time …</td>
<td>root</td>
<td>1sg.</td>
<td>ref.</td>
<td>lex.</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
114. and then signed some programs...  
115. Spent a quiet pleasant day at Stanley Woodward's...  
116. Haven't had a more pleasant week end...  
117. Drove an open car from Charlottesville to Washington...  
118. Had a Virginia Highway Policeman in a car ahead...  
119. Made the drive in 3 hours.  
120. Had Secretary of Treasury Snyder, Admiral Leahy...  
121. and felt they needed no extra accident coverage!  
122. Had ten minutes conversation with Henry Morgenthau...  
123. Told him...  
124. Had the usual hectic day...  
125. Lectured eleven junior Democratic Congressmen...  
126. Had four Republicans not long ago-nice young men...  
127. Talked to young Franklin for almost thirty minutes...  
128. Said...  
129. Went to Jim Forestal's [sic] house to a party...  
130. Don't know [why] Leahy, Nimitz & House members...  
131. Looks...  
132. and have gained a good man and a baby talking...  
133. At 3:30 today had a very interesting conversation...  
134. Sent for him to discuss the new Secretary...  
135. Asked him...  
136. Told him...  
137. and talked politics.  
138. and wanted to quit.  
139. and said...  
140. and did not even ask him about it!  
141. & said...  
142. and said...  
143. Began getting things in order.  
144. Asked...  
145. Finally received [sic] bill at air port.  
146. Signed it...  
147. and appointed Forestal [sic].  
148. Took off at 12:30 Washington time.  
149. At 1:30 Washington time received [sic] message...  
150. Arrived in Grandview about 3:30 CST...
| 151. went to the house | c.-r. | 1sg. | ref. | lex. | no |
| 152. and met sister & brother. | c.-r. | 1sg. | ref. | lex. | no |
| 153. Went to Belton with them | root | 1sg. | ref. | lex. | no |
| 154. and picked a casket. | c.-r. | 1pl. | ref. | lex. | no |
| 155. but stayed in Grandview… | c.-r. | 3pl. | ref. | lex. | no |
| 156. Spent Sunday morning and afternoon at Grandview. | root | 1sg. | ref. | lex. | no |
| 157. Arose at 6:15 | root | 1sg. | ref. | lex. | no |
| 158. had breakfast, | c.-r. | 1sg. | ref. | lex. | no |
| 159. Didn't sleep much Saturday night or Sunday night. | root | 1sg. | ref. | aux. | no |
| 160. So took a nap after breakfast. | root | 1sg. | ref. | lex. | adj. |
| 161. Had a time doing it. | root | 1sg. | ref. | lex. | no |
| 162. and said … | c.-r. | 3sg. | ref. | lex. | no |
| 163. had lunch at 12:00 | c.-r. | 1sg. | ref. | lex. | no |
| 164. and went to Grandview | c.-r. | 1sg. | ref. | lex. | no |
| 165. Returned to Washington. | root | 1sg. | ref. | lex. | no |
| 166. Had the new crew on the Sacred Cow… | root | 1sg. | ref. | lex. | no |
| 167. Landed in Washington at 4:16. | root | 1sg. | ref. | lex. | no |
| 168. Called Bess from White House. | root | 1sg. | ref. | lex. | no |
| 169. and were careful & conservative. | c.-r. | 3pl. | ref. | cop. be | no |
| 170. Stay at the house of the American C[ommanding]... | root | 1pl. | ref. | lex. | no |
| 171. Arose at 2 A.M., | root | 1sg. | ref. | lex. | no |
| 172. shaved, | c.-r. | 1sg. | ref. | lex. | no |
| 173. took a bath | c.-r. | 1sg. | ref. | lex. | no |
| 174. and then called Bess & Margaret. | c.-r. | 1sg. | ref. | lex. | adj. |
| 175. and signed autographs for the commandant … | c.-r. | 1pl. | ref. | lex. | no |
| 176. and autographed dollar bills for all the help. | c.-r. | 1sg. | ref. | lex. | no |
| 177. Arrived at Belem at 8:45. | root | 1pl. | ref. | lex. | no |
| 178. Saw the Guiana jungle | root | 1sg. | ref. | lex. | no |
| 179. and was greatly impressed at the Amazon. | c.-r. | 1sg. | ref. | cop. be | no |
| 180. Saw all of Brazil from Belem to Rio de Janeiro. | root | 1pl. | ref. | lex. | no |
| 181. Circled the city, | root | 1pl. | ref. | lex. | no |
| 182. landed on an island in the harbour. | c.-r. | 1pl. | ref. | lex. | no |
| 183. Took a power boat to the city. | root | 1pl. | ref. | lex. | no |
| 184. Recieved [sic] all the commanders… | root | 1sg. | ref. | lex. | no |
| 185. and then went overside with all the honors. | c.-r. | 1sg. | ref. | lex. | adj. |
| 186. Took about 21 hours to go down and twelve days … | root | 3sg. | expl. | lex. | no |
| 187. Landed in Washington aboard the Williamsburg … | root | 1pl. | ref. | lex. | no |
| 188. Have all sorts of things facing me. | root | 1sg. | ref. | lex. | no |
| 189. Had General Fleming in at 3 P.M… | root | 1sg. | ref. | lex. | no |
| 190. Had Sec[retary] of State, Agriculture… | root | 1sg. | ref. | lex. | no |
| 191. Went to 1st Baptist Church at 9:45[.] | root | 1sg. | ref. | lex. | no |
| 192. Spoke to the Sunday School Graduating Classes. | root | 1sg. | ref. | lex. | no |
| 193. Walked both ways. | root | 1sg. | ref. | lex. | no |
| 194. Had my new chairman of the Citizens Food Committee… | root | 1sg. | ref. | lex. | no |
| 195. Told him … | root | 1sg. | ref. | lex. | no |
| 196. Then went in to the Cabinet Food Committee meeting. | root | 1sg. | ref. | lex. | no |
| 197. Told them … | root | 1sg. | ref. | lex. | no |
| 198. Had an acrimonious meeting of my secretaries … | root | 1sg. | ref. | lex. | no |
| 199. and had the usual go around until 1 P.M. | c.-r. | 1sg. | ref. | lex. | no |
| 200. Then had to decide the argument between … | root | 1sg. | ref. | lex. | adj. - |
| 201. Listened to some commentators … | root | 1sg. | ref. | lex. | no |
| 202. and then called Dr. Steelman | c.-r. | 1sg. | ref. | lex. | adj. - |
| 203. & told him | c.-r. | 1sg. | ref. | lex. | no |
| 204. Sent to Congress the European Recovery Plan… | root | 1sg. | ref. | lex. | no |
| 205. Had the Dep[artmen]t of Interior survey our assets… | root | 1sg. | ref. | lex. | no |
| 206. Had the Treasury look into the financing. | root | 1sg. | ref. | lex. | no |
| 207. And finally had State, Defence [sic] and the White House… | s. coord. | 1sg. | ref. | lex. | adj. - |
| 208. Went to Bethesda to see Bill Hassett. | root | 1sg. | ref. | lex. | no |
| 209. Took him a poinsetta [sic] from the base… | root | 1sg. | ref. | lex. | no |
| 210. and charged him with receiving [sic] flowers … | c.-r. | 1sg. | ref. | lex. | no |
| 211. Saw Cordell Hull and Adm[iral] King. | root | 1sg. | ref. | lex. | no |
| 212. Went through 3 or 4 wards … | root | 1sg. | ref. | lex. | no |
| 213. and shook hands with patients … | c.-r. | 1sg. | ref. | lex. | no |
| 214. Went over to Walter Reed from Bethesda … | root | 1sg. | ref. | lex. | no |
| 215. and went through the bed fast wards with Dr. Graham… | c.-r. | 1sg. | ref. | lex. | no |
| 216. Met forty or fifty patients … | root | 1sg. | ref. | lex. | no |
| 217. Makes a person ashamed to be gloomy … | root | 3sg. | expl. | lex. | no |
| 218. Went down to the W[hite] H[ouse] garage … | root | 1sg. | ref. | lex. | no |
| 219. and then ate a tall dinner[,] | c.-r. | 1sg. | ref. | lex. | adj. - |
| 220. gained a pound and a half… | c.-r. | 1sg. | ref. | lex. | no |
Notes

Person

1sg. - 1\textsuperscript{st} person singular \hspace{1cm} 3sg. – 3\textsuperscript{rd} person singular
1pl. – 1\textsuperscript{st} person plural \hspace{1cm} 3pl. – 3\textsuperscript{rd} person plural

Clause type

root
\text{c.-r.} - a clause coordinated with a root
\text{c.-e.} - a clause coordinated with an embedded clause

ref. – a referential subject
expl. – an expletive subject

Verb type

lex. – a lexical verb
aux. – auxiliary
cop. \textit{be} - copula \textit{be}

Preposing

adj. – a preposed adjunct
adj.- - a non-preposed element
no – no preposing
References


Chomsky 1982


Haegeman, Liliane. 1997. Register variation, truncation and subject omission in English and in French. English Language and Linguistics 1, 233-270.


Haegeman, Liliane. 2011. The syntax of registers: Diary subject omission and the privilege of the root. Ms, GIST, Ghent University


Napoli, D. 1982. Initial material deletion in English. Glossa 16
Scott, K. 2010. The Relevance of Referring Expressions: the Case of Diary Drop in English, PhD, University College London
Diaries

*Carolyn’s diary.* The diary life of an American woman. Online
*President Harry S. Truman’s 1947 Diary Book, 1947,* Diary and Manual of the Real Estate
Board of New York, Inc. Transcribed by Raymond H. Geselbracht, Education and