

Subject pronoun drop in informal English

Andrew Weir

March 2008

The phenomenon of subject pronoun drop in informal English is investigated, both in spoken English and written English (“diary drop”). In particular, Haegeman and Ihsane (1999, 2001)’s reports of dropped subject pronouns in embedded clauses in written English are considered. It is determined that the dropping of subject pronouns in spoken English is best analysed as a phonological phenomenon, following work by Gerken (1991); specifically, that the first syllable in an intonational phrase, if weak, can be deleted in English. It is also determined that this analysis cannot be extended to subject pronoun drop in written English, and that “diary drop” is best analysed as a syntactic phenomenon — specifically, an instantiation of “topic drop”, as proposed by Haegeman (1990); it is also proposed that a topic drop analysis can satisfyingly account for embedded null subjects, *contra* Haegeman and Ihsane (1999). There is also brief discussion of the broader implications of proposing separate analyses for spoken and written English for our theory of language.

Submitted in partial fulfilment of the requirements for the degree of

M.A. (Hons) Linguistics

University of Edinburgh

College of Humanities and Social Sciences

School of Philosophy, Psychology and Language Sciences

Supervisors: Professor Caroline Heycock and Dr Peter Ackema

Contents

1	Introduction	5
2	Restrictions on subject pronoun drop	7
2.1	Spoken English	7
2.1.1	Motivating evidence	7
2.1.2	Forbidden configurations	8
2.1.3	Additional examples of drop	10
2.2	Written English	11
2.2.1	Unambiguously permitted configurations	11
2.2.2	Marginal configurations — embedded null subjects . . .	12
2.2.3	Forbidden configurations	13
2.2.4	Fiction versus non-fiction	13
2.3	Differences between spoken and written English	15
3	Spoken English	16
3.1	Domain	16
3.2	For a phonological analysis	17
3.2.1	A metrical analysis	17
3.2.2	How it works	18
3.2.3	Evidence from prosody	19
3.3	Against a syntactic analysis	20
3.3.1	Syntactic explanations are <i>ad hoc</i>	22
3.3.2	WEAK- σ_1 -DEL is not <i>ad hoc</i>	22
3.3.3	One phonological rule versus several syntactic rules . .	23
3.3.4	Salience	24
3.4	Conclusion	24
4	Written English	25

4.1	Domain	25
4.1.1	Phonological	25
4.1.2	Pragmatic	27
4.1.3	Syntactic	29
4.2	Nature of the empty category	30
4.2.1	<i>pro</i>	30
4.2.2	PRO	32
4.2.3	A-trace	32
4.2.4	A'-trace	32
4.3	Analysis	33
4.3.1	The phenomenon of topic drop	33
4.3.2	Topic drop in diary English	34
4.3.3	An objection: null objects and double topics	38
4.3.4	A second objection: dropping of expletive subjects	42
4.3.5	Two diary dialects?	43
4.4	Conclusion	44
5	Implications for theory of language	45
6	Conclusion	47

Acknowledgements

Great thanks go to my dissertation supervisors, Caroline Heycock and Peter Ackema, for their helpful supervision and comments, and in particular their willingness to take so much time out of their schedules to discuss this dissertation with me.

This dissertation was composed using LyX, a front-end to the L^AT_EX typesetting engine. These programs made putting the dissertation together far easier than it would otherwise have been and I am grateful to their creators.

Thanks to my flatmates, for keeping me healthy (physically and mentally) throughout writing this.

Thanks to my mum and dad for everything.

1 Introduction

Sentences in English are generally considered to require overt subjects. In a standard description of English, (1 a) below would be considered grammatical, while (1 b) would be considered ungrammatical:

- (1) (a) I walked the dog yesterday.
 (b) Walked the dog yesterday.

But in fact this is not true. (1 b), and sentences like it lacking subject pronouns, are in fact common in colloquial spoken English. This phenomenon has a written counterpart in “diary drop” (Haegeman, 1990, 1997; Haegeman and Ihsane, 1999, 2001); subject pronouns can frequently be dropped in certain registers of written English, such as diaries but also including text messages, emails, and other forms of informal communication¹, as below.

- (2) Should really go to the gym tomorrow.
 (3) Don’t think I can make it tonight.

As sentences are required to have subjects, following the Extended Projection Principle (Haegeman, 1994, p. 68), this phenomenon needs to be accounted for. The most obvious solution is to suggest that the pronoun in these sentences is the non-overt pronoun *pro*, used in (for example) Spanish sentences such as the below:

- (4) *pro Compré unas manzanas.*
pro bought-1SG some apples
 ‘I bought some apples.’

However, we will see in the course of this dissertation that this cannot be the correct solution and that a different analysis is required.

¹Although subjects can be dropped from many forms of written communication, throughout this dissertation I shall generally make reference only to diaries; it is to be assumed that all the above-mentioned forms of written communication are meant throughout, unless specifically stated otherwise.

We will also see that, contrary to how it may at first appear, subject pronoun drop in spoken and written English are different phenomena; contrary to what Horsey (1998) suggests, they cannot both be explained with the same analysis. In section 2 I will lay out the conditions under which subject pronoun drop is permitted in spoken and written English, and demonstrate that the conditions for subject pronoun drop are very different between the two. I argue in section 3 that subject pronoun drop in spoken English is a phonological phenomenon — specifically, a metrical one — while in section 4 I will argue that subject pronoun drop in written English is a syntactic phenomenon, building on Haegeman (1990)'s analysis of subject pronoun drop as topic drop. Finally, in section 5, I will discuss the broader ramifications of such a split analysis on our analysis of the differences between spoken and written language in general.

2 Restrictions on subject pronoun drop

2.1 Spoken English

Overt subject pronouns are not required in certain configurations in spoken English. In this section I will lay out the permitted and restricted configurations.

2.1.1 Motivating evidence

The following sentences are all fully grammatical in a colloquial style (*e* represents the gap where a subject would ordinarily be present):

- (5) (a) *e* Won't be in the office tomorrow.
 (b) (Why didn't you and your flatmates go to the party?)
 — *e* Didn't fancy it.
 (c) (Am I invited to the party?)
 — *e* Must be, surely.
 (d) (Why didn't [he/she/they] come to the party?)
 — *e* Didn't fancy it, I suppose.
 (e) *e* Seems to be quite noisy over there.
 (f) *e* Always rains on Mondays.

These examples show that subject pronouns can be easily dropped in simple declarative sentences. Examples (5 a, 5 b) show that first person subject pronouns, both singular and plural, can be dropped; example (5 c) shows that second person subject pronouns can be dropped; example (5 d) shows that third person subject pronouns, both singular and plural, can be dropped. Examples (5 e, 5 f) show that pleonastic *it*, both expletive *it* as in (5 e) and “weather” *it* as in (5 f), can be dropped.

2.1.2 Forbidden configurations

However, there are several configurations in which subject pronoun drop is not permitted, for example in questions (both yes/no and *wh*-questions), as examples (6 a, 6 b) show. Note that in both these cases *e* could stand in for either *you* or *they*; similar sentences could be constructed to show that first person and third-person singular subject pronouns are equally undroppable in these configurations.

- (6) (a) *Are *e* going to the party?
 (b) *Why are *e* not going to the party?

Contrastive subjects may not be dropped:

- (7) (Who runs this place?)
 (a) — I run this place.
 (b) — **e* Run this place.

Subject pronoun drop is also not permitted in embedded clauses, whether the root subject is overt or not, and whether or not the subject of the embedded clause is co-referential with the subject of the root clause, as shown by the following examples (8 a–8 d). In these examples, *e* can be read as either being co-referential with the root subject or not, without changing the ungrammaticality.

- (8) (a) *I don't think *e* should go.
 (b) **e* Don't think *e* should go.
 (c) *I don't know whom *e* should see.
 (d) **e* Don't know whom *e* should see.

Subject pronoun drop is in fact generally impossible in cases where there is preposed material:

(9) *Tomorrow, *e* won't be in the office.

In addition, the verbs *be*, *have*, *will*, *would* and *had*, in their affirmative form, seem to be unable to stand without a subject:

- (10) (a) **e* Is going to the party.
 (b) **e* Have been to Turkey.
 (c) **e* Will rain tomorrow.²
 (d) ??*e* Would go to the party if I could.
 (e) **e* Had met John before that.

However, all these verbs can be affixed with *-n't* to create a negative form, and in this case, examples (10 a–10 e) suddenly hugely improve:

- (11) (a) ?*e* Isn't going to the party.
 (b) *e* Haven't been to Turkey.
 (c) *e* Won't rain tomorrow.
 (d) *e* Wouldn't go to the party (even) if I could.
 (e) *e* Hadn't met John before that.

In addition, in situations where the cliticisable verbs are contrastive, subject pronoun drop is possible:³

- (12) (a) (— Turkey?
 e Haven't been there.
 — Egypt?
 e Haven't been there.
 — Cyprus?)
 e Have been there!

²There seems to be an exception to this in the exclamation *Will do!* as an enthusiastic response to a request. I suggest that this particular chunk of language is a fixed and unanalysable expression — although how *Will do!* might have achieved this status in the first place is not clear.

³I am indebted to a friend of mine, Neil Bennet, for pointing out this possibility.

(b) (— The meeting on Tuesday?

e Won't go to that.

— The lecture on Thursday?

e Won't go to that.

— The party on Saturday?)

e Will go to that!

(and so on for other verbs).

2.1.3 Additional examples of drop

There are further examples of dropped elements in spoken English beyond subjects. Although these do not have a direct bearing on the phenomenon at issue, investigating these dropped elements can nevertheless provide evidence for our analysis of subject pronoun drop.

While initial subjects before cliticisable verbs (as presented in 10a–10e) cannot be dropped, if the verbs are in fact cliticised to the subject, the whole subject + clitic element can be dropped:⁴

(13) (a) I'm having a party on Saturday.

→ Having a party on Saturday.

(b) I've been to Turkey before.

→ Been to Turkey before.

(c) I'd met John before that.

→ Met John before that.

Furthermore, in yes-no questions with subject-auxiliary inversion, the auxiliary can be deleted:

⁴It seems that this is only possible if it is obvious from the context which auxiliary has been dropped, and there is not a more salient reading which blocks the example with a dropped subject/auxiliary. Compare *I'd go to the party* with [#]*go to the party*; the latter is more easily read as an imperative, so subject/auxiliary dropping is blocked in this case.

- (14) (a) Are you having a party on Saturday?
 → You having a party on Saturday?
- (b) Have you been to Turkey before?
 → You been to Turkey before?
- (c) Had you met John before that?
 → ?You met John before that?

2.2 Written English

Many of the constraints set forward in Section 2.1 also apply to informal written English, but some do not. I shall again set out the permitted and restricted configurations.

2.2.1 Unambiguously permitted configurations

In simple declarative sentences such as those in examples (5 a–5 f), subject pronoun drop is permissible in written English.

Subject pronoun drop is also possible with preposed material, in contrast to spoken English:

- (15) (a) Tomorrow *e* won't be in the office.
- (b) So *e* shall now stop writing for a day. (Haegeman, 1990, p. 164, quoting a Virginia Woolf diary)

and also with affirmative cliticisable verbs, also in contrast to spoken English:

- (16) (a) *e* Am going to the gym tomorrow.
- (b) *e* Have been feeling a bit peaky lately.
- (c) *e* Will go to the gym tomorrow.
- (d) *e* Would go to the party if I could.
- (e) *e* Had been feeling a bit peaky.

2.2.2 Marginal configurations — embedded null subjects

Until recently it was considered that the configurations listed in section 2.2.1 were the only permitted configurations in diary English. That is, subject pronoun drop was (a) a root phenomenon, not permitted in embedded clauses, and (b) not permitted in (root) yes/no or wh-questions.

However, recent work has suggested that (a) above is not true; subject pronoun drop is also permissible in embedded clauses, although it is a marginal construction. Haegeman and Ihsane (1999) report sentences such as (17 a), from a real (but published) diary *Journals 1954–1958* by Allen Ginsberg, and (17 b–17 f), from Helen Fielding’s fictional *Bridget Jones’s Diary*:

- (17) (a) When *e* saw him at noon, he’d been in North Beach all last night
 (b) *e* Think *e* will cross that bit out as *e* contains mild accusation
 (c) *e* Cannot believe *e* have not realised this before
 (d) *e* Understand where *e* have been going wrong
 (e) *e* Give all clothes which *e* have not worn for two years or more to
 homeless
 (f) but only string *e* have got is blue

In addition to the investigation of diaries, the author of this dissertation has collected and examined many (200+) text messages that he has received over the course of writing this dissertation. Examples of root null subjects in text messages are very common, while examples of embedded null subjects are very rare — but one does occur:

- (18) Hey i can’t come as *e* will be at work but *e* hope it goes well!

It is possible that the rareness of embedded null subjects in text messages is due to a rareness of embedded clauses in text messages in general, but I have not run any statistical analysis on this collection of text messages to confirm or refute this possibility.

2.2.3 Forbidden configurations

Haegeman and Ihsane (1999, p. 129) also discover, however, that condition (b) above — dropped subjects are not found in root yes/no or wh-questions — **does** appear to hold in the diaries that they investigate.⁵ The below sentences, for example, are ungrammatical (and are all equally ungrammatical whether the empty category is taken to represent the first, second or third persons):

- (19) (a) *Should *e* go to the gym on Saturday?
 (b) *Do *e* go to the gym often enough?
 (c) *Who did *e* see last night?
 (d) *Where should *e* go for my birthday?

These seem to be the only contexts in which null subjects are absolutely ruled out.

2.2.4 Fiction versus non-fiction

It is worth noting that null subjects in the configurations set out in section 2.2.2 seem relatively rare. In the research prior to Haegeman and Ihsane (1999) — for example Haegeman (1990, 1997); Horsey (1998) — embedded null subjects were assumed to be ungrammatical. Much of Haegeman and Ihsane (1999)'s data comes, not from genuine diaries or other contexts in which subject pronoun drop seems licit (text messages, computer mediated communication in general), but from the fictional *Bridget Jones's Diary* by Helen Fielding, as well as other fictional diaries. Indeed, Haegeman and Ihsane (2001) suggest that a register allowing embedded null subjects constitutes a (minority)

⁵They do acknowledge (p. 134) that “given the low numbers of sentences with *wh*-fronting in our extracts, the absence of the null subject may be a sampling accident”. I will suggest later in this dissertation that the lack of empty subjects in root *wh*-fronting contexts provides crucial evidence for my proposed analysis; an effective means of falsifying my analysis would be to find a null subject in a root *wh*-fronting context.

separate dialect, “represented by some recent British fictional diary writing” (p. 334).

These “diaries” are not natural expressions of written language, but rather attempts to fictionally **imitate** such written language. One could therefore argue the authors of these fictional diaries have overgeneralised the rule that “really” exists in informal writing — i.e. that root subjects may be null — to one that says that embedded subjects may also be null. In essence, in writing a fictional diary, they are making the style more “diary-like” than that which real diaries actually exhibit.

There is certainly nothing in principle barring the possibility of sentences which are ungrammatical being used in a literary register; an author can use the language as he or she wishes. However this cannot be the full story, for two reasons:

- Embedded null subjects do occur in “genuine” written language, as the text message in (18) shows (although not necessarily in all the contexts in which they appear in the fictionalised diaries).
- There are still **some** syntactic restrictions on null subjects even in the fictional diaries, as shown in section 2.2.3; it is not simply a case of “drop any subject”. In particular, the asymmetry between root questions with null subjects (ungrammatical) and embedded/indirect questions with null subjects (grammatical) is not accounted for.

We must therefore take the examples of embedded null subjects in written English at face value, although without ruling out the possibility (as proposed by Haegeman and Ihsane, 2001) that there are two “dialects” of written communication in contention; ramifications of this possibility on the proposed analysis will be discussed in section 4.3.5.

2.3 Differences between spoken and written English

Horsey (1998, section 5) suggests that subject pronoun drop in spoken and written English can be accounted for by a unified analysis. Based on the facts outlined above, I do not agree. Any attempt to explain subject pronoun drop has to account for various differences in the distribution of subject pronoun drop between spoken and written English, some of which are shown in Table 1.

Context	Spoken	Written
Subject placed initially	OK	OK
Preposed material	Ungrammatical	OK
Before cliticisable verb	Ungrammatical	OK
Embedded clauses	Ungrammatical	OK but rare
Root yes/no questions	Ungrammatical	Ungrammatical
Root <i>wh</i> -questions	Ungrammatical	Ungrammatical

Table 1: Some differences in grammaticality of subject pronoun drop in different contexts in spoken and written English.

Given the stark difference depending on context of acceptability between spoken and written English, I argue that the two merit separate analyses. In the rest of this dissertation I shall set out and evaluate various analyses for both phenomena.

3 Spoken English

3.1 Domain

In order to determine what mechanisms can account for subject pronoun drop in spoken English, we must first consider what the **domain** of the phenomenon is — that is, whether it is **phonological**, **syntactic**, or **pragmatic**.⁶

From the data presented in section 2.1, one thing is immediately striking; subject pronoun drop is only licit in one location: initially.

(20) *Utterance-initiality*

Subject pronouns which are dropped in spoken English must be **utterance-initial**.

This rules out a pragmatic explanation, under which one would not expect the position of the dropped subject to be relevant — if a pragmatic explanation rules a subject pronoun out in one position, it should rule them out in all positions, as a pragmatic analysis would not be sensitive to linguistic structure. In principle, it leaves both a phonological and a syntactic explanation possible; “initial” position can be defined as one of two things:

- initial in a **phonological phrase**
- **clause-initial**, i.e. leftmost in a **syntactic tree**

We therefore need to investigate further to determine which of these is the relevant domain for subject pronoun drop.

⁶By a phenomenon having a “pragmatic domain” is meant that the phenomenon is due to actual limiting (extra-linguistic) factors in the world, such as time, money, shortness of breath, a desire to be taciturn etc.

3.2 For a phonological analysis

I propose that the domain is **phonological**, based on work by Gerken (1991).

3.2.1 A metrical analysis

Gerken (1991) investigates the phenomenon of subject pronoun drop in infant English. She concludes that subject pronoun drop in infant English is not the result of an early “mis-setting” of the *pro*-drop parameter, as Rizzi (1994) claims; in fact it is not a syntactic phenomenon at all, but rather a **metrical** one. Gerken suggests that, in utterances that commence with an iambic foot (that is, an unstressed syllable followed by a stressed one), the first syllable can be dropped from the utterance; she supports this with experimental infant data.

We can generalise this metrical account to adult speech:

- (21) *Initial weak syllable deletion in informal spoken English* (WEAK- σ_1 -DEL)

If the first syllable in an English phonological phrase is weak (= does not bear phonological stress), it can be deleted in an informal register of spoken English (which we shall call INFORMALS)

One other rule is required to fully explain the scope of the phenomenon.

- (22) *Obligatory cliticisation in INFORMALS* (CLITICISE)

In INFORMALS, those verbs which **can** cliticise (*am, are, is, have, had, will, would*) **must** do so, unless the verb is contrastive and so bears phonological stress.

The reasons why this rule is required shall be set out in detail below. It is not an *ad hoc* rule; there is independent reason to support the existence of CLITICISE. It does seem to be the case that *I am going to the cinema* belongs to a more formal register in spoken English than *I'm going to the cinema*; CLITICISE provides one way of accounting for this fact.

3.2.2 How it works

The above two rules account for the facts laid out in section 2.1. “Subject pronoun drop” is not in fact subject pronoun drop, but rather “first syllable drop”; it happens, however, that as English is an SVO language, subject pronouns very often appear as the first syllable of an utterance. This explains why subject pronoun drop is not possible with preposed material (6 a, 6 b, 9) or in embedded clauses (8 a–d).

The proposed rules also explain why subjects may not be dropped before affirmative cliticisable verbs (10 a–e). We can examine two possible derivations to determine why this is the case.

The two rules CLITICISE and WEAK- σ_1 -DEL can in principle occur in any order. We will assume that these rules work on a surface structure like *I am going to the cinema*. Let us first consider a derivation where CLITICISE occurs before WEAK- σ_1 -DEL (underlining indicates phonological stress):

- (23) I am going to the cinema
 → I’m going to the cinema (CLITICISE)
 → I’m going to the cinema (WEAK- σ_1 -DEL)

This form is grammatical (although it is unclear whether the phrase *Going to the cinema* is actually a result of these rules, or rather the generation of a syntactic fragment). Let us also consider a derivation where WEAK- σ_1 -DEL occurs before CLITICISE:

- (24) I am going to the cinema
 → I am going to the cinema (WEAK- σ_1 -DEL)

At this point the verb *am* should cliticise by application of CLITICISE. However it has nothing to cliticise to. The derivation therefore crashes. Neither of these derivations can, therefore, generate the form **Am going to the cinema*.

This analysis also explains why contrastive subjects cannot be dropped (example 7). Contrastive pronouns receive phonological stress in English, whereas non-contrastive pronouns do not (Wells, 2007, p. 124). As such, under the metrical analysis, contrastive pronouns are not candidates for subject pronoun drop, while non-contrastive pronouns are.

Furthermore, it satisfactorily explains why, before cliticisable verbs, subjects can be dropped just in case the verb is stressed (examples 12 a, 12 b) — the application of CLITICISE is blocked in this case, rendering an utterance such as *Have been there!* grammatical.

It also explains the difference in grammaticality between **Have been there* (with normal, non-contrastive stress) and *Haven't been there*. In a sentence such as *I haven't been there*, CLITICISE cannot apply (as *haven't* cannot cliticise — Zwicky and Pullum, 1983). As *I* would only receive stress if contrastive (Wells, 2007, p. 124), the (normal) position for the first stressed syllable will be on *have*, and the subject *I* can drop by application of WEAK- σ_1 -DEL, resulting in *Haven't been there*.

3.2.3 Evidence from prosody

It is not, in fact, strictly accurate to claim that subject pronoun drop in English is solely an utterance-initial phenomenon. Consider (25) below. (I have marked intonation using the system of Wells (2007), with ´ signifying a rise, ` a fall, ˇ a fall-rise, and underlining signifying nuclear stress.)

(25) (— Are you going to the party?)

 `Yeah, e said I would.

Compare the ungrammatical⁷ example (26):

(26) (— But you said he was a doctor!)

?* *Yeah, e* said he was a doctor of phi^hlosophy, not of *medicine*.

How do these examples differ? They do not obviously differ in their syntax. They do however differ in their phonology, and specifically their intonation. In (25), the intonation is “reset” after the fall on *Yeah*, while in (26) it is not; the high tone created by the rise on *Yeah* continues until the fall on phi^hlosophy. We can interpret this in terms of intonational phrase (IP) boundaries; intonation “resets” after an IP boundary, as below:

(27) (a) || *Yeah*, || *e* thought I *would*. ||

(b) *|| *Yeah*, | *e* said he was a doctor of phi^hlosophy, || not of *medicine*. ||

We can see that there is a generalisation here. The subject can disappear from an utterance just in case it is right-adjacent to a major phonological break (||), that is, it occupies the first position in an intonational phrase.

3.3 Against a syntactic analysis

Gerken (1991) provides evidence to believe that subject pronoun drop is a phonological phenomenon, but not a watertight proof that subject pronoun drop is not a syntactic phenomenon. In principle there is no particular reason to suggest that one could not have an analysis where subject pronouns drop (or are realised by an empty category) just in case they are in the highest possible syntactic position in their tree — and there are many analyses suggesting just that (such as Rizzi (1994); Haegeman (1997); Horsey (1998)). In

⁷Ungrammatical in my judgement, that is. I freely admit that intuitionistic data start to get hazy at this point. It would be interesting to investigate corpora of informal spoken English to confirm or refute the analysis I present here.

fact, as we will see later in this dissertation, we will have to adopt a syntactic analysis for written English, so the possibility certainly cannot be dismissed.

One might assume that the impossibility of dropping subjects before affirmative cliticisable verbs — but not negative cliticisable verbs — suggests that subject pronoun drop is a phonological phenomenon. After all, it stretches credibility to suggest that the addition of a negative affix to a verb could possibly **syntactically** affect the presence or absence of a subject — and in fact we know that in general it does not; modals such as *should* can also take the affix *-n't*, but both sentences below are equally grammatical:

(28) (a) *e* Should really go to the gym tomorrow.

(b) *e* Shouldn't really go to the gym tomorrow.

We do, however, know that adding *-n't* blocks cliticisation of the verb to the subject — a phonological phenomenon. We might therefore assume that subject pronoun drop is therefore licensed by phonological considerations, not syntactic ones, as the affixation of *-n't* only changes the phonological situation in an utterance, not the syntactic situation (or at least, not in a way relevant to subjects).

However, this would not be correct.⁸ Assuming that CLITICISE is a valid rule (at least with respect to INFORMALS), then it does not matter whether the subject disappears by a syntactic or phonological process — CLITICISE can still cause a derivation such as (24) to crash. Adding *-n't* to the verb does alter the grammaticality of a sentence with a dropped subject — but this is a consequence of CLITICISE; it does not reflect on whether the dropping of the subject itself is syntactic or phonological.

There are, however, various grounds on which a phonological explanation can be preferred to a syntactic one. I set these grounds out below.

⁸Thanks to Peter Ackema for pointing this out.

3.3.1 Syntactic explanations are *ad hoc*

Any syntactic rule which can account for the facts in section 2.1 — i.e. any syntactic rule that can account for a subject pronoun being non-overt just in case it is sentence-initial — will necessarily be somewhat *ad hoc*, in that it will require at least one rule to be introduced which cannot be generalised to pronouns in non-sentence-initial positions.

This is not automatically a problem. We should not refuse to introduce a rule simply because it does not generalise, if we do not have any alternative explanation. However, by application of Occam's Razor, we would rather not create more rules than are absolutely necessary to explain a given phenomenon; and of those rules, we prefer those which generalise to those which do not. I consider that a phonological explanation generalises better than a syntactic explanation, as I shall explain in the next section.⁹

3.3.2 WEAK- σ_1 -DEL is not *ad hoc*

If the rule WEAK- σ_1 -DEL does exist, we would expect it to act in situations other than just subject pronoun drop. This does indeed appear to be the case.

For example, we have seen that WEAK- σ_1 -DEL can account for the production of sentences where both the subject and clitic auxiliary drop, such as *I'm going to the cinema* → *Going to the cinema*. Although it could be argued that this is simply the production of a syntactic fragment, no such explanation can be made for cases where the auxiliary drops in a yes/no question, such as *Are you going to the cinema?* → *You going to the cinema?*, a grammatical sentence also predicted by application of WEAK- σ_1 -DEL.

⁹The counter-argument may be raised at this point that, as we have admitted that we will require a syntactic analysis for subject pronoun drop in written English, it would actually be preferable to posit a syntactic analysis for spoken English as well, in an attempt to unify the two phenomena. As I state in section 2.3, I do not agree with this; I consider the differences between the distribution of subject pronoun drop in spoken and written English to be too great to support a unified analysis.

We also see deletion of non-constituents, as in the following example (from Gerken, 1991, p. 438):

(29) (Is this restaurant any good?)

— Man over there seems to think so.

The initial article *the* has been dropped, in exactly the same fashion as subject pronouns. It is difficult, although not impossible, to suggest a syntactic explanation for the deletion of the article in (29); examples such as (29) strengthen the claim that WEAK- σ_1 -DEL is a rule which generally holds true in English speech.

3.3.3 One phonological rule versus several syntactic rules

We have seen that the sentences below are all grammatical in INFORMAL_S:

(30) (a) I don't think so.

(b) I'm going to the cinema.

(c) Are you going to the cinema?

(d) The man over there seems to think so.

In theory we can still posit a syntactic explanation for these deletions. However, all of the deleted elements above inhabit different positions in the syntactic tree. In a standard description of English, *I* in (30 a) occupies [Spec, IP]; *I'm* in (30 b) is split over [Spec, IP] and I^0 ; *Are* in (30 c) occupies C^0 ; and *The* in (30 d) occupies D^0 (Haegeman, 1994). If we are determined to posit a syntactic analysis, we would either require (at least) four separate rules to deal with the four cases above, or else require a completely ad-hoc rule like “the first lexical item in a clause can optionally be null”. If we do this we are clearly missing a generalisation. WEAK- σ_1 -DEL captures the generalisation neatly in only one rule.

3.3.4 Saliency

If we accept WEAK- σ_1 -DEL as an explanation for subject pronoun drop, we do not need to posit any syntactic mechanism for determining the saliency of a subject. Contrastive subjects cannot be dropped, as shown in example (7) on page 8. A metrical analysis allows us to account for (7) simply by appealing to the known facts about English phonology; as stated above, contrastive pronouns receive phonological stress in English, whereas non-contrastive pronouns do not (Wells, 2007, p. 124).

3.4 Conclusion

Based on the above, I conclude that 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”, which we have seen generally holds true in English. I will now proceed to consider the case of subject pronoun drop in written English.

4 Written English

I have argued above that the phenomena of subject pronoun drop in spoken and written English merit separate analyses. In this section I will discuss how we can account for subject pronoun drop in written English.

4.1 Domain

As with spoken English, our first concern is to establish the domain of the phenomenon — phonological, pragmatic or syntactic.

4.1.1 Phonological

I have argued that subject pronoun drop in spoken English is a phonological phenomenon, so it may make sense to start our enquiry into subject pronoun drop in written English by assuming that the same holds for written English.

However, we run into an immediate problem with this line of analysis; namely, what does it mean to talk of phonology in written language? If what is meant is that the written language reflects the phonological form (PF) of the spoken language, so that phonological processes are reflected in writing, then we cannot claim that the domain of subject pronoun drop in written English is phonological. We have established that there are sentences like *So e shall now stop writing for a day* which are grammatical in written English but have no counterpart in the spoken phonology. This cannot be the phonological domain we seek.

However, one possible analysis of written English in general is that the written register of English is a register which does not surface in the normal spoken language, but which nonetheless exists and has spellout rules which generate a “PF” which corresponds to what is written down. If we accept this, can we find a way of explaining subject pronoun drop as a phonological

phenomenon?

Based on observational data, there does not seem to be any evidence to suggest that subject pronoun drop in written English is a phonological phenomenon. Even if we discount embedded null subjects, there is no indication that phonological processes have any bearing on the distribution of null subjects, or that the phenomenon's domain is the intonational phrase (IP), as shown for spoken English in section 3.2.3. The following examples illustrate this point. (The intonation markings come from my intuition, based on how the sentence would sound if the subject were present.)

(31) || So *e* shall ˘now | stop ˘writing for a day. ||

(32) || ˘Tomorrow | *e* won't be in the ˘office. ||

(33) || *e* Will go to the ˘gym tomorrow. ||

(34) || *e* ˘Won't go to the gym tomorrow. ||

Only (34) is grammatical in spoken English, as the empty subject occurs at the start of an intonational phrase (and CLITICISE does not cause the derivation to crash, as in (33)). This is consistent with the analysis put forward in section (3.2.3).

In contrast, however, all of these sentences are grammatical in written English. From a phonological point of view, the empty subject is not adjacent to any phonological break in (31), adjacent to a minor break in (32), and adjacent to an IP boundary/major break in both (33) and (34). There does not appear to be any phonological constraint on where a null subject can appear. It seems safe to say that subject pronoun drop in written English is not a phonological phenomenon.

4.1.2 Pragmatic

It is possible that subject pronoun drop in written English can be explained due to purely pragmatic factors. That is, actual constraints in the world, whether lack of time or space, can force certain elements — those that are understood in the discourse — to be dropped.

At first glance this seems an appealing explanation. For example text messaging, one possible environment for subject pronoun drop, bears constraints on the length of an individual message which could lead to the dropping of “redundant” elements such as subject pronouns. It is well established that constraints on length can lead to registers such as “telegamese”; “unnecessary” words in telegrams are dropped so that a charge (levied per word) is not paid. Pressure of time, such as when taking lecture notes, also results in a severely “truncated” register of English, with many elements being dropped (see Janda (1985) for discussion).

There is clearly also a “pragmatic” or discourse-related element in the dropping of subjects. Subjects need a prominent referent in order to be dropped; very often this referent is not explicitly presented in the discourse, but rather assumed pragmatically (in the case of text messaging and diaries, a dropped subject seems to generally be equated with the author, unless there is a discourse-driven reason to suppose otherwise).

However, we cannot rely — at least, not solely — on pragmatic constraints to explain subject pronoun drop. Consider the text message in example (18) on page 12, repeated here as (35):

(35) Hey i can't come as *e* will be at work but *e* hope it goes well!

A single SMS text message can contain up to 160 characters (Wikipedia, 2008). The above message contains 58 characters. Adding in the two missing pronouns, adding four characters (the pronoun *I* plus a following space in

two places), would not have pushed the message over the limit. In general terms, subject pronoun drop cannot be due to objective constraints on length of communication; diary entries, for example, are not generally restricted in length.¹⁰

One could argue that constraints of time rather than space are at issue. If this were true, then we would expect communications containing null subjects to look rather like the note-taking register investigated by Janda (1985). But in general they do not. For example, Janda provides the below example (p. 443):

(36) What *e* did was take “sha” and. . .

In note-taking, we do see a situation where elements seemingly drop anywhere, with no syntactic restrictions on where this may happen or what may drop (Janda discusses the dropping of many other elements, including object pronouns, articles, copulas etc.); the only restriction is that dropped elements can be reconstructed from discourse. We can presume that this dropping is due to a pragmatic pressure of time.¹¹ But it is sentences precisely of the form above — preposed *wh*-material along with a dropped subject — that we have seen are **never** found in diary registers. This fact could not be explained by a pragmatic analysis. Examine the two sentences below:

(37) (a) *e* Went to the gym on Saturday.

(b) What *e* did on Saturday was go to the gym.

The dropped pronouns are equally interpretable in both sentences. There

¹⁰Interestingly, **fictional** diary entries often **are** restricted in length. Haegeman and Ihsane (2001, p. 337) note that *Bridget Jones’s Diary* was originally written as a 1,000-word newspaper column. This could potentially indicate that the forms of subject pronoun drop which have up till now only been found in fictional diaries could indeed be due to pragmatic factors — although the question still remains of why subject pronoun drop is forbidden in e.g. root *wh*-questions.

¹¹Although having said that, this clearly isn’t the full story; there is no semantic need for the *wh*-cleft (*What e did. . .*) in (36). If constraints on time are so important, why is the sentence not paraphrased further to *e took sha and. . .*, which would be quicker to write?

can be no purely pragmatic explanation why (37 a) is acceptable in the diary register, while (37 b) is not.

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 — see section 4.2 for discussion) can be dropped. But relying solely on pragmatic factors cannot explain the distribution of the phenomenon which we see.

4.1.3 Syntactic

If subject pronoun drop in written English is neither phonological nor pragmatic, only a syntactic explanation is left. There is also positive evidence to suggest that subject pronoun drop in written English is a syntactic phenomenon. Haegeman (1997) claims that there is an asymmetry in the nature of possible preposed material before subject pronoun drop. According to Haegeman, adjuncts can be preposed but arguments cannot:

- (38) (a) Tomorrow, *e* will go to gym.
(b) *More problems *e* don't need.

Such an asymmetry suggests that subject pronoun drop in written English is sensitive to syntactic structure and, as such, we are best placed looking for a syntactic explanation, where subject pronouns are replaced with an empty category. Having determined that, our next task is to determine the nature of that empty category.

4.2 Nature of the empty category

Haegeman and Ihsane (1999, pp. 118ff.) consider the nature of the empty category that replaces overt subjects in diary registers of English.¹² They consider the four types of empty category distinguished by Chomsky (1982), namely

- *pro*;
- PRO;
- A-traces;
- A'-traces.

Haegeman and Ihsane consider each in turn; in the sections below I provide a précis of their arguments. The definitions of the empty categories which I provide below are based on Haegeman (1994).

4.2.1 *pro*

pro as a possible candidate for the empty category is rejected for a variety of reasons. Firstly, *pro* is supposed to be possible in a language only if that language has either rich verbal agreement morphology, that is, verbal agreement sufficient to identify an argument even if that argument is missing (in principle, this could be applied to objects as well as subjects, if verb-object agreement is present; see Huang (1984, pp. 535ff.) for analysis of data from Pashto which shows *pro* in object position); or no verbal agreement morphology at all, as with Chinese (again, see Huang (1984) for discussion). English is in neither of these positions; although verbal

¹²Haegeman presents the same arguments in various places, for example Haegeman (1990, pp. 169ff.), and to some extent in Haegeman (1997, pp. 235ff.). They do not differ in outcome, although some aspects are more fully discussed in some papers. I consider Haegeman and Ihsane (1999) to be the fullest overall discussion.

agreement morphology is in general very poor in regular verbs in English, it does exist in the form of the third person singular present ending *-s*.

Furthermore, in *pro*-drop languages, “the definite subject can occupy a postverbal position” (Haegeman, 1997, p. 235). The Spanish sentence below, for example, is grammatical (Spanish being a *pro*-drop language), while the English equivalent is not:

- (39) (a) *Apareció el profesor.*
 appeared-3SG the professor
 ‘The professor appeared.’
 (b) *Appeared the professor.

If the *pro*-drop parameter were to be “reset” in the diary register of English, one would expect sentences like (39 b) to appear in English diaries; but these sentences are as ungrammatical in the diary register as in any other register of English.

In addition, *pro* can appear in root yes/no and *wh*-questions, although empty categories in English diaries cannot; compare the below examples, again from Spanish:

- (40) (a) *¿Estás pro seguro?*
 are-2SG *pro* sure
 ‘Are you sure?’
 (b) *Are *e* sure?
- (41) (a) *¿Qué viste pro?*
 what saw-2SG *pro*
 ‘What did you see?’
 (b) *What did *e* see?
- (42) (a) *¿Dónde está pro?*
 where is *pro*
 ‘Where is he/she/it?’
 (b) *Where is *e*?

Based on the conditions above, it seems that we cannot categorise subject pronoun drop in the English diary register as an example of *pro*-drop.

4.2.2 PRO

PRO is that empty category which appears as the subject of nonfinite clauses such as the below (from Haegeman and Ihsane, 1999, p. 119):

- (43) (a) The cat expects [PRO to get regular meals].
 (b) [PRO To be fed regularly] is important for the cat.
 (c) When [PRO hungry], the cat comes into the kitchen.

PRO does not alternate with an overt subject and does not appear in finite clauses. Subjects dropped in diary registers of English **do** alternate with overt subjects and **do** appear in finite clauses. As such, PRO is not a candidate for explaining subject pronoun drop in the English diary register.

4.2.3 A-trace

An A-trace is an empty category whose antecedents occupy an argument position, i.e. an A-position. It is the empty category in the non-finite clauses below.

- (44) (a) John_i wants [_{t_i} to go to Turkey].
 (b) John_i seems [_{t_i} to be angry].

Again, A-traces do not alternate with overt subjects or appear in finite clauses, so like PRO, they are not candidates for explaining the phenomenon at issue.

4.2.4 A'-trace

An A'-trace is the trace of movement to a non-argument position, i.e. an A'-position. It is the trace left after the operations of *wh*-movement and topicalisation, as can be seen below.

- (45) (a) Who_i did he see t_i yesterday?
 (b) Chips_i, he likes t_i.

This seems like a promising candidate for the empty category at issue here. A'-traces alternate with overt subjects and can appear in finite clauses. They do have to be governed by an co-indexed antecedent, which does not seem to be the case with subject pronoun drop. However, I will argue below that the empty category which replaces subjects in written English is in fact an A'-trace, and provide an explanation for its apparent lack of government.

4.3 Analysis

We have seen that subject pronoun drop in written English is most satisfactorily analysed as a syntactic phenomenon and that the empty category which replaces the dropped subject is best analysed as an A'-trace. In this section I suggest an **topic drop** analysis for subject pronoun drop in written English. This analysis was first considered by Haegeman (1990), but later (in Haegeman (1997) and subsequent papers) Haegeman rejected this analysis. I contest, however, that with some alterations a topic drop analysis is the best way to account for the facts of subject pronoun drop in written English.

4.3.1 The phenomenon of topic drop

Haegeman (1990) explains topic drop with reference to an example from Portuguese (adapted from Haegeman, 1990, p. 175, who is citing Rizzi):

- (46) (a) *A Juana viu e na televisão*
 Juana saw on television
 'Juana saw him/her/it on television.'
 (b) TOP_i A Juana viu t_i na televisão.

(46a) shows a typical Portuguese sentence, without an overt object. (46b) is a proposed analysis; that the object of the sentence is an A'-trace which is

bound by a **null topic operator**; that is, the object has topicalised, leaving an A'-trace, but the topicalised constituent is itself phonetically null, and receives its referent from the discourse. Haegeman claims that there is good evidence for supposing that this phenomenon can be best analysed in terms of non-overt topic operators. For example, one would expect only one topic to be possible in any given sentence; that is, if a sentence already has an overt topic in Portuguese, an empty category (with a different referent than the overt topic) should not be possible. Haegeman claims that this is indeed the case, although does not give examples. Furthermore, topicalisation and *wh*-movement are incompatible, at least in English:

(47) *That guy_i, who kissed t_i?

and Haegeman (1990, p. 175) claims that *wh*-movement is also incompatible both with overt topicalisation and with empty categories in Portuguese (although, again, she does not give examples).

4.3.2 Topic drop in diary English

Haegeman (1990) suggests that this model of topic drop — an empty category, bound by a non-overt topic operator located higher in the clause — can be extended to diary English. Sentences would have structures as proposed below.

(48) *e* Saw Jim yesterday.

[TOP_i [t_i Saw Jim yesterday.]]

In diaries or text messages the referent of the null topic operator would normally be the author of the communication, but this is not necessarily the case; a topic can be established by the preceding discourse, as (49) below shows:

(49) *e* Saw Jim yesterday. *e* Was not a happy man.

[TOP_i [t_i Saw Jim yesterday.]] [TOP_{jim} [t_{jim} Was not a happy man.]]

At first glance there is much to recommend this analysis. It explains the asymmetry between argument and adjunct preposing; **More problems, e don't need* is bad in diary English because the topicalised *More problems* and the operator TOP binding the empty category cannot co-exist (as one sentence cannot have two separate topics). On the other hand, *Tomorrow e will go to gym* does not create a problem, as *Tomorrow* is not a topic.

The analysis also explains why null subjects are incompatible with yes/no questions and root *wh*-questions; overt topics are incompatible with these structures, so one would expect null topic operators to be equally incompatible:

(50) (a) **That man_i, do you see t_i?*
 (b) **Do e go to the gym enough?*

(51) (a) **That man_i, where did you see t_i?*
 (b) **Where should e go tomorrow?*

Haegeman (1990) considered embedded null subjects impossible. Does a topic drop analysis extend to registers where embedded null subjects are possible? Haegeman and Ihsane (1999, p. 135) argue that topic drop is incompatible with embedded null subjects, arguing that null subjects occur in positions where topicalisation is impossible and providing the below examples:

(52) (a) Was worried that *e* might split (from *Bridget Jones's Diary*)
 (b) **My blue skirt_i I was worried [that t_i might split]*

(53) (a) Understand where *e* have gone wrong (also *Bridget Jones's Diary*)
 (b) **This relationship_i I wonder [where t_i went wrong]*

(52b) is indeed ungrammatical, due to a violation of the *that*-trace filter in English., i.e. “the sequence of an overt complementizer followed by a trace is ungrammatical” (Haegeman, 1994, p. 423); and (53b) is ungrammatical, possibly due to a Subjacency violation (*this relationship* crossing two IP nodes). However, there is not any particular restriction on topicalising phrases in embedded clauses (compare the acceptable *My blue skirt I was worried might split*); and Haegeman and Ihsane seem to be assuming that topicalisation can raise only to the very highest position in the clause. This does not seem correct, as seen from the examples below:

- (54) (a) I think [that [that book]_i, he read t_i.]
 (b) ?I don't know [why [people like that]_i you put up with t_i.]

Topicalisation can raise a constituent to a high position in an embedded clause. This operation is vacuous for (overt) topicalisation of subjects in embedded clauses, which would not raise above any overt elements; but can be shown for object topicalisation, as in (54 a). Assuming this analysis, embedded null subjects can perfectly well be bound by a null topic operator, assuming that that operator itself is to be found in the same position that overt embedded topicalised material occupies in, for instance, (54 a).

In general it seems to be the case that contexts that are genuinely infelicitous for topicalisation are also infelicitous for null subjects, adding further weight to a null topic analysis. I would tentatively suggest that the unifying factor in all such cases is subject-auxiliary inversion; this covers yes/no questions and root *wh*-questions (but not embedded/indirect questions), as well as other cases which seem to be infelicitous, as shown below:

- (55) (a) *That party_i, should you attend t_i?
 (b) *Should *e* attend the lecture?

- (56) (a) *That guy_i, when did you see t_i?
 (b) *When did *e* see that guy? (or *When saw that guy?)
- (57) (a) ??That book_i, under no circumstances will I read t_i.
 (b) ??Under no circumstances will *e* read that book.
- (58) (a) *That programme_i, only then did I watch t_i.
 (b) *Only then did *e* watch that programme. (or *Only then watched that programme.)¹³
- (59) (a) *That man_i had I seen t_i, I would have shaken him by the hand.
 (b) *Had *e* seen that man, [I/*e*] would have shaken him by the hand.

This is consistent with the findings of Haegeman and Ihsane (1999), who in their investigation of diary English found no instances of subject pronoun drop in contexts of subject-auxiliary inversion (p. 129).

One may ask at this point which syntactic position topicalised material occupies. We have some hints in this regard: for example, topicalised material in embedded clauses cannot occupy [Spec, CP], as the topicalised material would then appear before the complementiser, when in fact it appears after the complementiser, as in (54 a). Neither could topicalised material co-exist with *wh*-words, which also occupy [Spec, CP] (Haegeman, 1994, p. 382), although (54 b) shows that it can. Furthermore, wherever topicalised material does appear, it would seem that subject-auxiliary inversion — i.e. I-to-C movement (Haegeman, 1994, p. 591) — has a deleterious effect, suggesting that topicalised material occupies a position such that I-to-C movement may break a binding relationship between topicalised material and its trace.

¹³It may be of interest to further exploration of this phenomenon that the latter of these sentences, while still badly ungrammatical, feels to me to be a considerable improvement over the former. There is not the space here to explore this in detail, but this difference in acceptability, if genuine, seems to support the view that the blocking of topicalisation is dependent on subject-auxiliary inversion.

Adjunction to IP is one possible position for topicalised material:

(60) [_{IP} That book_i [_{IP} he read t_i]]

(61) I think [_{CP} that [_{IP} that book_i [_{IP} he read t_i]]]

with a bar on subject-auxiliary inversion being explained as a Subjacency violation — I-to-C movement would involve I⁰ movement across two IP nodes, which is forbidden in English (Haegeman, 1994, p. 423). However, this creates problems with sentences such as the below:

(62) I don't know [_{CP} why_i [_{IP} people like that_j [_{IP} you put up with t_j t_i]]]

Here *why* has also had to cross two IP nodes, so the sentence should be ungrammatical, contrary to fact.¹⁴

I will not here attempt to determine which syntactic position in particular topicalised elements (null or overt) occupy; for the present purposes, it is sufficient simply to note that it is **possible** to topicalise elements in both root and embedded clauses, and it is therefore possible to postulate a null topic operator occupying the same position as overt topicalised material, whatever that position is.

4.3.3 An objection: null objects and double topics

A topic drop analysis therefore seems to satisfyingly capture the distribution of null subjects in informal written English; null subjects can be found only in those cases where topicalisation is possible. However, the analysis does have problems. One problem noted by Haegeman (1990, p. 177) is that the presence of a null topic operator should in principle allow null objects, as in the Portuguese case. Haegeman raises the objection that these are not in fact allowed in English diaries, citing the below example:

¹⁴Thanks to Caroline Heycock for pointing this out.

(63) *Bill did not greet *e*.

TOP_i Bill did not greet t_i.

I would contest that it is not in fact generally impossible to drop objects in the English diary register. I would suggest that (64) below is acceptable in such a register:

(64) *e*_i Received my bank statement_j today; *e*_i will shred *e*_j later.

And Haegeman herself (1987; 1990, p. 159) has noted the phenomenon of object drop in instructional registers such as recipes, in phrases such as *Boil e for ten minutes*.

I would suggest that subject pronoun drop in diary register, and object drop in instructional registers, are essentially instantiations of the same phenomenon. Subject pronoun drop in diaries is clearly more common than object drop in diaries; I would suggest that this is due to the fact that subjects are in general more prone to be interpreted as topics, and therefore to drop. In instructional registers, composed as they are mainly of imperatives, there are no subjects to be topicalised in the first place, and objects are therefore at liberty to be replaced with an empty category bound by a null topic operator.

However, in (64) we have two empty categories, with different referents, in the one sentence. If there is only one position for a null topic operator in any given sentence, then at least one of the empty categories in (64) cannot be bound. One solution is to propose, as Huang (1984, p. 555) does, that multiple topics are in fact possible; Huang supports this assumption with evidence from Chinese, a topic-prominent language which allows topic drop:

(65) *Zhangsan, neiben shu, ta hen xihuan.*
 Zhangsan that book he very like
 '(As for Zhangsan), that book, he likes very much.'

And certainly in Japanese — also a topic-prominent language, according to Huang — a sentence containing two dropped arguments is permissible and indeed normal:

(66) *(John-wa keiki-o tabeta ka?*
 John-TOP cake-OBJ ate Q
 ‘Did John eat the cake?’)

Hai, e e tabeta.
 Yes, ate
 ‘Yes, he ate it.’

Presumably an analysis of (64) would look something like the below:

(67) TOP_i [TOP_j [t_j Will shred t_i later]]

However, earlier in this dissertation, it was claimed that English sentences can contain only one topic, in order to explain the asymmetry in preposing arguments and adjuncts (**More problems e don't need* versus *Tomorrow e will go to gym*). If we assume that English sentences can have more than one topic, there is nothing to rule out the below:

(68) More problems_i [TOP_j [t_j don't need t_i]]

In general, it seems that English sentences indeed cannot have two topics. This is difficult to test as subject topicalisation would be a vacuous operation in English. However, we can see from the below that topicalising two separate constituents seems to lead to a very bad result:

- (69) (a) ??London_i, I saw John in t_i¹⁵
 (b) John_i, I saw t_i in London.
 (c) **London_j, John_i, I saw t_i in t_j

It is perhaps notable that (69 c) seems to be far less acceptable than (68). This may be because (69 a) is in itself quite marginal. It may, however, indicate that in general non-overt topicalised material is more acceptable than overt topicalised material in diary English,¹⁶ with two overt topics (69 c) being very bad, one overt topic and one non-overt topic (68) being fairly bad, and two non-overt topics (64) being relatively acceptable.

Although we can appeal to intuition judgements such as the above to go some way towards explaining the phenomenon, it is clear that sentences like (64) present a problem for the topic-drop analysis. As the topic-drop analysis is satisfying in other respects — such as explaining the syntactic constraints on subject pronoun drop, and providing a natural explanation for the fact that dropped subjects are recovered from the discourse — I am loath to abandon it on the grounds of sentences such as (64). It is possible that future work on the nature of the empty categories in Japanese sentences such as (66) will provide an explanation.¹⁷

¹⁵I suspect this construction by itself may be ungrammatical for many speakers. It is (very) marginal for the author (and feels something of a Scotticism, my native language being Scottish Standard English), but it is just possible for me. It is improved for me by adding material like *not Paris*, making the contrastive nature clear.

¹⁶We are assuming here that the option of null topics is simply not available in other registers of English.

¹⁷Neeleman and Szendrői (2007) argue that such empty arguments (“radical *pro*-drop”, as they describe it) are possible in languages whose pronouns agglutinate for case, number, or other features. Although this is the case for Japanese (where case is marked by particles, and where (e.g.) the first person plural pronoun *watasitati* is derived from the first person singular *watasi* and the plural marker *tati*), it clearly is not the case for English (compare *I/we, he/him* etc.)

4.3.4 A second objection: dropping of expletive subjects

Another objection to the topic-drop analysis, raised by Haegeman and Ihsane (1999), is that expletive subjects can drop in diary English (e.g. *Rained yesterday; Seems like there'll be trouble ahead; Weatherman says will rain tomorrow* etc.), despite not being subject to topicalisation (as they cannot be topics).

It is indeed unclear why sentences without expletives are possible in diary English under a topic-drop analysis. One possible analysis is that there is simply no requirement for them to be present. As both subjectless sentences and sentences containing subjects are permissible in diary English, we can view diary English as containing two competing grammars — one which permits topic drop and one which does not. With this assumption, we could suggest that the grammar that permits topic drop does not require expletive pronouns as standard English does. That is, in sentences such as *Rained yesterday*, expletive subjects are not bound by a null topic operator; they are simply **not present**. There is possibly some evidence for this from Japanese. Japanese — a topic-drop language — does not require (and indeed does not permit) expletive subjects in sentences like (70) below:

- (70) *Samui.*
 is-cold
 'It's cold.'

However, there is no reason to believe that what is missing in (70) is an **expletive** subject. Japanese being a topic-drop language, (70) could have easily derived from the sentences below:

- (71) (a) *Watasi-wa samui.*
 I-TOP is-cold
 'I'm cold.'
- (b) *Heya-wa samui.*
 room-TOP is-cold
 'The room is cold.'

- (c) *Tenki-wa samui.*
 weather-TOP is-cold
 'The weather is cold.'

That is, the non-present subject could well be a semantically contentful element which is dropped by the normal procedures of topic drop in Japanese.¹⁸ The dropping of expletive subjects certainly presents a problem for a topic-drop analysis; further research into the existence and/or nature of expletive pronouns in languages known to exhibit topic drop could provide further insight here. Interestingly, topic drop in Germanic languages such as German or Dutch does not permit the dropping of expletive subjects (Ackema, personal communication). This fact may be counted as an argument against a topic-drop analysis for English subject pronoun drop, as one would expect English, as a Germanic language itself, to behave in a fashion more similar to German or Dutch than to Japanese or Chinese.

4.3.5 Two diary dialects?

One interesting possibility is flagged by Haegeman and Ihsane (2001). Having investigated both American and British (albeit fictional) diaries, Haegeman and Ihsane find that embedded null subjects are attested only in (a subset of) the British diaries; no embedded null subjects are found in any American diaries.

It would be interesting to investigate if sentences with overt topicalisation in embedded clauses — such as *I think that that book I have read* — are licit in American English. The author of the present intuition has no intuitions about American English, but it could prove an interesting avenue for further work.

¹⁸Thanks to Caroline Heycock for pointing this out.

4.4 Conclusion

We have seen that there are convincing arguments for suggesting that subject pronoun drop in written English is not the same phenomenon as initial syllable deletion in spoken English. Rather, the fact that it is subject to constraints such as subject-auxiliary inversion suggests that subject pronoun drop in written English is a syntactic phenomenon. I have proposed a topic drop analysis to explain the phenomenon of pronoun drop; and have suggested that not only subject drop in diaries, but also object drop in instructional registers, can be explained in this way. However, there do remain problems with this analysis, and further research into this phenomenon is still necessary. I now intend to briefly consider the wider implications of these findings, and in particular consider what implications it has for our theory of language to propose such a great difference between the grammars of the spoken and written varieties of the same language.

5 Implications for theory of language

The above findings allow us some deeper insights for our theory of language. As was briefly touched upon in section 4.1.1, the analyses above reveal that there is quite a large distinction between spoken and written English — at least in the written English found in diary registers.

What we see is that there are sentences which are grammatical in written English but not in spoken English, such as for example *So e will now stop writing for a day*. That is, written English is not a subset of spoken English, but rather *vice versa*; although all spoken sentences are grammatical in written English,¹⁹ not all written sentences are grammatical in spoken English. This is notable, because it indicates that writing cannot simply be an iconic representation of a given phonological form.

The questions then arise — what is writing if not an iconic representation of PF; and how does the human language capacity treat writing distinctly from speaking? A suggestion — that written English has its own, completely separate grammar, with its own “PF” and “spellout rules”, was put forward in section 4.1.1. If the topic-drop analysis for dropped subjects in diary dialects of English is correct, then the grammar of diary English is really very distinct from spoken English — arguably changing from a subject-prominent to a topic-prominent language.

However, this seems unsatisfying, for various reasons. In general, written English does not differ hugely from spoken English. If this dissertation were read aloud, it would not be strictly ungrammatical (although it may sound stilted). It does not seem likely that the grammars are completely separate. Why, then, does the grammar of the diary register seem to differ so sharply from that of other registers? Furthermore, how is the grammar of the diary

¹⁹This notion is a little tricky to define, but it seems reasonable to state that any spoken utterance could be written down with the claim that the resulting writing represents a colloquial register.

register acquired? One is not exposed to writing as an infant; however, one has grammatical intuitions about what is and is not acceptable in written diary English in the same way as one has similar intuitions about spoken English. Where do these intuitions come from?

There is not enough space in this dissertation to discuss these issues fully, but they would provide an interesting avenue for further research.

6 Conclusion

In this dissertation I have investigated the phenomenon of subject pronoun drop in English. I have determined that, while subject pronouns are dropped both in colloquial spoken English and in certain informal registers of written English, these two phenomena are distinct and require different analyses. I have provided evidence to support a metrical analysis of subject pronoun drop in spoken English. In fact, in spoken English, “subject pronoun drop” is a somewhat misleading name, as the element that can be dropped is the first syllable in an intonational phrase, if that syllable is unstressed. It simply happens that that element is often a subject pronoun.

I have shown, however, that in written English the phenomenon is syntactic, and “subject pronoun drop” is an accurate name. I have proposed an analysis suggesting that subject pronoun drop is a case of topic drop, and that non-overt pronouns in English diaries are bound by a null topic operator occupying the same position as topicalised elements.

I have also discussed some of the ramifications of forming separate analyses of spoken and written English for this phenomenon on our analysis of the language capacity as a whole, and specifically the distinctions between spoken and written language more generally.

References

- Chomsky, N. (1982). *Lectures on Government and Binding*. Dordrecht: Foris.
- Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30(4):431–451.
- Haegeman, L. (1987). Complement ellipsis in English: or how to cook without objects. In Simon-Vandenberghe, A., editor, *Studies in Honour of René Derolez*, pages 248–261. Ghent: Seminarie voor Engelse en Oud-Germaanse Taalkunde R.U.G.
- Haegeman, L. (1990). Understood subjects in English diaries. *Multilingua*, 9(2):157–199.
- Haegeman, L. (1994). *Introduction to Government and Binding Theory*. Oxford: Blackwell Publishers, 2nd edition.
- Haegeman, L. (1997). Register variation, truncation, and subject omission in English and in French. *English Language and Linguistics*, 1(2):233–270.
- Haegeman, L. and Ihsane, T. (1999). Subject ellipsis in embedded clauses in English. *English Language and Linguistics*, 3(1):117–145.
- Haegeman, L. and Ihsane, T. (2001). Adult null subjects in the non-*pro*-drop languages: two diary dialects. *Language Acquisition*, 9(4):329–346.
- Horse, R. (1998). Null arguments in English registers: A Minimalist account. BA thesis, La Trobe University, Bundoora, Victoria, Australia.
- Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15:531–574.
- Janda, R. D. (1985). Note-taking English as a simplified register. *Discourse Processes*, 8:437–454.

- Neeleman, A. and Szendrői, K. (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry*, 38(4):671–714.
- Rizzi, L. (1994). Early null subjects and root null subjects. In Teun Hoekstra and Bonnie Schwartz, editors, *Language Acquisition Studies in Generative Grammar*, number 8 in Language Acquisition and Language Disorders, pages 151–176. Amsterdam: John Benjamins Publishing Company.
- Wells, J. C. (2007). *English Intonation: An Introduction*. Cambridge: Cambridge University Press, 3rd edition.
- Wikipedia (2008). Short message service — Wikipedia, the free encyclopedia. Online at http://en.wikipedia.org/w/index.php?title=Short_message_service&oldid=191435598; accessed 15 February 2008.
- Zwicky, A. M. and Pullum, G. K. (1983). Cliticization vs. inflection: English N'T. *Language*, 59(3):502–513.