

Subject pronoun drop in informal English¹

ANDREW WEIR

University College London

ABSTRACT

The phenomenon of subject pronoun drop (SPD) in informal English is investigated. SPD in spoken English is discussed and compared to SPD in written English, or ‘diary drop’ as discussed by Haegeman (1990; 1997), Horsey (1998) and Haegeman & Ihsane (1999; 2001). It is argued that these two phenomena cannot be unified and in fact result from completely different linguistic processes. The dropping of subject pronouns in spoken English is analysed as a phonological phenomenon, following work by Gerken (1991); specifically, that the first syllable in an intonational phrase, if unstressed, can be deleted in English. It is also argued that this analysis cannot be extended to subject pronoun drop in written English, and that ‘diary drop’ is best analysed as a syntactic phenomenon, although one for which a definitive account is still lacking. There is also brief discussion of the broader implications of proposing separate analyses for spoken and written English for a theory of language.

1 INTRODUCTION

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

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

But in fact this is not true. (1b), 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 & 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 communication,² as below.

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

I will argue that subject pronoun drop (henceforth SPD) in spoken and in 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 SPD is permitted in spoken and written English, and demonstrate that the

conditions for SPD are very different between the two. I argue in section 3 that SPD in spoken English is a phonological phenomenon – specifically, a metrical one – while in section 4 I will argue that SPD in written English is a syntactic phenomenon, although one for which a definitive theoretical account is still lacking. In section 5 I will briefly discuss the ramifications of these findings on a broader theory of language, and in particular, the relationship between spoken and written language.

2 THE DATA

Although SPD definitely exists, it is not a free procedure in either spoken or written English; not all subjects can be dropped. This section will examine the relevant data, first for spoken English and then for written English.

2.1 *Spoken English*

2.1.1 *Permitted configurations*

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

- (4) (a) *e* Won't be in the office tomorrow.
- (b) A: Why didn't you and your flatmates go to the party?
B: *e* Didn't fancy it.
- (c) A: Am I invited to the party?
B: *e* Must be, surely.

- (d) A: Why didn't (he/she/they) come to the party?
 B: *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 (4a, b) show that first person subject pronouns, both singular and plural, can be dropped; example (4c) shows that second person subject pronouns can be dropped; example (4d) shows that third person subject pronouns, both singular and plural, can be dropped. Examples (4e, f) show that pleonastic *it*, both expletive *it* as in (4e) and 'weather' *it* as in (4f), can be dropped.

2.1.2 *Forbidden configurations*

However, there are several configurations in which SPD is not permitted, for example in questions (both yes/no and *wh*-questions), as examples (5a, b) show. Note that in both these cases *e* could replace 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.

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

Focused (that is, stressed) subjects may not be dropped:

- (6) A: Who runs this place?
 (a) B: **I** run this place.
 (b) B: **e* Run this place.
- (7) (a) I for one don't recall seeing him
 (b) **e* For one don't recall seeing him.³

SPD 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 (8a–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 who *e* should see.
 (d) **e* Don't know who *e* should see.

SPD is in fact generally impossible in cases where there is preposed material:

- (9) (a) *Tomorrow, *e* won't be in the office.
 (b) *When I was in Paris, *e* visited the Louvre.

In addition, the verbs *be*, *have*, *will*, *would* and *had*, in their affirmative form, seem 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.
 (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, in which case examples (10a–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, SPD is possible:⁴

- (12) (a) A: Turkey?
 B: *e* Haven't been there.
 A: Egypt?
 B: *e* Haven't been there.
 A: Cyprus?
 B: *e* Have been there!
- (b) A: The meeting on Tuesday?
 B: *e* Won't go to that.
 A: The lecture on Thursday?
 B: *e* Won't go to that.
 A: The party on Saturday?
 B: *e* Will go to 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 (4a–f), subject pronoun drop is permissible in written English.

In contrast to spoken English, however, SPD is also possible with certain forms of preposed material:

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

SPD is also possible with affirmative cliticisable verbs, again in contrast to spoken English:

- (14) (a) *e* Am going to the gym tomorrow.
 (b) *e* Have been feeling a bit ill 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, SPD 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; SPD is also permissible in embedded clauses, although it is a marginal construction. Haegeman & Ihsane (1999) report sentences such as (15a), from a real (but published) diary *Journals 1954–1958* by Allen Ginsberg, and (15b–f), from Helen Fielding’s fictional *Bridget Jones’s Diary*:

- (15) (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

We can see, therefore, that in at least some dialects of diary writing, SPD is licensed in embedded clauses, whether subordinate clauses or relative clauses.

2.2.3 *Forbidden configurations*

Haegeman & Ihsane (1999: 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 sentences below, for example, are ungrammatical (and are all equally ungrammatical whether the empty category is taken to represent the first, second or third persons):

- (16) (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?

I would add that instances of subject-verb inversion more generally are also infelicitous for SPD:

- (17) (a) *Never have *e* seen such a spectacle.
 (b) *Only then did *e* watch it.
 (c) *Had *e* seen that man, I would have shaken him by the hand.

In addition, Haegeman (1997) points out that there is an asymmetry in the nature of possible preposed material before subject pronoun drop; non-arguments can be preposed but arguments cannot:

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

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 & Ihsane (1999) – for example Haegeman (1990; 1997), Horsey (1998) – embedded null subjects were assumed to be

ungrammatical. Much of Haegeman & Ihsane (1999)'s data comes, not from genuine diaries or other contexts in which subject pronoun drop seems licit (for example text messages, or computer mediated communication in general), but from the fictional *Bridget Jones's Diary* by Helen Fielding, as well as other fictional diaries. Indeed, Haegeman & Ihsane suggest that a register allowing embedded null subjects constitutes a (minority) separate dialect, 'represented by some recent British fictional diary writing' (2001: 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. 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 & Ihsane 2001) that there are two 'dialects' of written communication in contention, one

permitting SPD in embedded clauses and one not. This possibility does not substantially change the task at hand; if any dialect permits SPD in embedded clauses, a theoretical account must be presented for this, whether it is a minority dialect or not.

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 SPD has to account for various differences in the distribution of SPD between spoken and written English, some of which are recapped in Table 1. Given the stark difference depending on context of acceptability between spoken and written English, I argue that the two merit separate analyses.

Context	Spoken	Written
Subject placed initially	OK	OK
Before cliticisable verb	Ungrammatical (if verb unfocused)	OK
Embedded clauses	Ungrammatical	OK but rare
Preposed adjuncts	Ungrammatical	OK
Preposed arguments	Ungrammatical	Ungrammatical
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

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 whether the phenomenon is phonological, syntactic, or pragmatic.⁶ From the data presented in section 2.1, one thing is immediately striking. SPD is only licit in one location: initially.

(19) *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.

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*

Subject pronoun drop is quite characteristic of infant English; Gerken (1991) investigates this phenomenon. She concludes that it is not the result of an early mis-setting of the pro-drop parameter, as Hyams (1986) 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, 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:

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

If the first syllable in an English phonological phrase does not bear phonological stress, it can be deleted in informal registers of spoken English.

There is significant independent evidence for this; there are many examples of initial syllable deletion in spoken English which does not apply solely to subjects. For example, while initial subjects before cliticisable verbs (as presented in 10a–e) cannot be dropped, if the verbs are in fact cliticised to the subject, the whole complex [subject + clitic] can be dropped:⁷

- (21) (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:

- (22) (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?

Initial articles can also be deleted (example from Gerken 1991: 438):

- (23) (Is this restaurant any good?)
 Man over there seems to think so.

This rule of initial weak syllable deletion appears sufficiently independently motivated. One other phonological rule is required to fully explain the scope of the phenomenon. It is intuitively correct that informal varieties of spoken English require verbs which can cliticise to do so; saying *It is raining* is stilted when one could say simply *It's raining*. This is not applicable when the verb is contrastive, and so bears phonological stress; saying *It IS raining!* is fine. This is formalised in (24) below.

(24) *Obligatory cliticisation in informal spoken English (CLITICISE)*

In informal registers of spoken English, 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.

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 SPD is not possible with preposed material (5a, b; 9a) or in embedded clauses (8a–d).

The proposed rules also explain why subjects may not be dropped before affirmative cliticisable verbs (10a–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 string like *I am going to the cinema*.

Let us first consider a derivation where CLITICISE occurs before WEAK- σ_1 -DEL:

- (25) I am going to the cinema
 → I'm going to the cinema (by CLITICISE)
 → I'm going to the cinema (by 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:

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

At this point CLITICISE should apply. We can make two different suggestions for its application: either the application of the rule fails (as the clitic 'm does not have a host to cliticise to), and so the whole derivation fails; or that the application of the rule succeeds, regardless of the lack of a host, and the verb appears in its clitic (or reduced) form. Judgement of the grammaticality of a sentence like *'m going to the cinema* is left to the reader, but cases such as *'S gonna rain tomorrow* seem acceptable to me.⁸ The point, however, is that no possible application of these rules can result in sentences with unreduced verbs but no subject, such as **Am going to the cinema*.

This analysis also explains why contrastive subjects cannot be dropped (example 6). Contrastive pronouns receive phonological stress in English, whereas non-

contrastive pronouns do not (Wells 2007/2006: 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 only in cases where the verb is stressed (examples 12a, b); CLITICISE is not applicable 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; see Zwicky & Pullum 1983). As *I* would only receive stress if contrastive (Wells 2007/2006: 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

Prosody provides further evidence to suggest that the process of subject pronoun drop in English is sensitive to the phonological domain, not the syntactic one. Despite the claim that has been maintained above, it is not in fact strictly accurate to claim that SPD is solely an utterance-initial phenomenon. Consider (27) below. (I have marked intonation using the system of Wells (2007/2006), with ` signifying a fall and ˇ signifying a fall-rise, and underlining signifying nuclear stress.)

(27) Are you going to the party?

ˇYeah, e thought I `would.

Compare the ungrammatical⁹ example (28):

(28) But you said he was a doctor!

?* \sim Yeah, *e* said he was a doctor of phi`losophy, not of \sim 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 (27), the intonation is ‘reset’ after the fall on \sim Yeah, while in (28) it is not; the high tone created by the rise on \sim Yeah continues until the fall on phi`losophy. We can interpret this in terms of intonational phrase (IP) boundaries; intonation ‘resets’ after an IP boundary, as below:

(29) (a) || \sim Yeah, || *e* thought I \sim would. ||

(b) *|| \sim Yeah, | *e* said he was a doctor of phi`losophy, || not of \sim 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. We can therefore recast WEAK- σ_1 -DEL as follows:

(30) *Initial weak syllable deletion in informal spoken English*

If the first syllable in an English *intonational* phrase does not bear phonological stress, it can be deleted in informal registers of spoken English.

This rule, along with CLITICISE, appears to account for all of the relevant data.

3.3 *Against a syntactic analysis*

Gerken (1991)'s investigation of subject pronoun drop in infant language provides evidence to believe that subject pronoun drop is a phonological phenomenon, but not a watertight proof that it 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 fact, as we will see later, we will have to adopt a syntactic analysis for written English, so the possibility certainly cannot be dismissed. There are, however, various grounds on which a syntactic explanation for spoken English should be dispreferred. I set out these grounds below.

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.

The 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 SPD in spoken and written English to be too great to support a unified analysis.

3.3.2 *One phonological rule versus several syntactic rules*

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

- (31) (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 are of different categories from each other. They all also inhabit different locations in the syntactic tree; indeed, *I'm* in (31b) is generally analysed as being split across two different positions (without tying oneself down to a particular analysis, we can say that *I* is generally considered to be in a specifier position and *'m* in a head position). 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 spelled out as silence’. If we do this we are both being stipulative and clearly missing a generalisation. WEAK- σ_1 -DEL captures the generalisation neatly in only one rule, a rule strongly motivated by the empirical evidence.

3.3.3 *Salience*

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 salience of a subject. Contrastive subjects cannot be dropped, as shown in example (6). A metrical analysis allows us to account for (6) 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 2007/2006: 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 weak syllables at the left edge of an intonational phrase’, which we have seen generally holds true in English.

4 WRITTEN ENGLISH

I have argued above that the phenomena of subject pronoun drop in spoken and written English merit separate analyses. This section will present evidence that SPD in written

English is a syntactic phenomenon, not linked to the process of syllable drop discussed above.

4.1 *Written SPD is not 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 SPD 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 SPD 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 SPD as a phonological phenomenon? Such an analysis may be possible, but it would require considerable research into what the nature of the 'PF' of written language is. While not completely discounting this possibility, I will argue that the distribution of subject pronoun drop lends itself better to a non-phonological explanation.

4.2 Neither is written SPD due to pragmatics

It is possible that subject pronoun drop in written English could be explained due to purely pragmatic factors. That is, actual constraints in the world, whether lack of time or space, can allow 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 SPD, 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, both in spoken and written English (see Nariyama 2004 for discussion of the spoken case). 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 following text message, received by the author:

(32) 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 (3GPP 1998). 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), obviously 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 usually 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), where pronouns are dropped essentially without constraint. But in general there is no such resemblance. For example, Janda provides the below example (1985:443):

(33) 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:

- (34) (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 can be no purely pragmatic explanation why (34a) is acceptable in the diary register, while (34b) is not. There is clearly a pragmatic, or discourse-related, component to subject pronoun drop in written English, in so far 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.

4.3 *SPD is syntactic: positive evidence*

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 SPD in written English is a syntactic phenomenon. Recall that Haegeman (1997) shows that there is an asymmetry in the nature of possible preposed material before subject pronoun drop; that non-arguments can be preposed but arguments cannot, as in (18a, b). Such an asymmetry suggests that SPD in written English is sensitive to syntactic structure; we are therefore best placed looking for a syntactic explanation. More generally, the pattern of forbidden configurations that we saw in section 2.2.3 clearly has a common syntactic domain. Sensitivity to yes/no and *wh*-question formation, and

to subject-verb inversion more generally, is indicative of a syntactic phenomenon.

Many possible syntactic explanations have been proposed in the literature for ‘diary drop’, including an analysis where SPD is analogous with Germanic topic drop (Haegeman 1990), or an account reliant on a ‘truncation’ of the higher parts of the CP layer (Haegeman 1997, based on the cartographic approach of Rizzi 1994, 1997; Horsey 1998 recasts the argument in a Minimalist framework). However, these analyses cannot account for the existence of subject drop in embedded clauses (Haegeman & Ihsane 1999, 2001), for which the field is awaiting an analysis. It would be neat and parsimonious if an account could be found to unify root and embedded SPD, but as yet it appears that none has been; and indeed, given the distinction between ‘conservative’ diary dialects which forbid embedded SPD and ‘liberal’ dialects which permit it, it may be that embedded and root SPD instantiate two quite different phenomena.

5 FURTHER WORK

The above findings allow us some deeper insights for our theory of language. As was briefly touched upon in section 4.1, 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,¹² 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. However, this seems unsatisfying, for various reasons. In general, written English does not differ hugely from spoken English. If the present work 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 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? I have not considered these matters in great detail, but they would provide an interesting avenue for further research.

6 CONCLUSION

We have seen 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. Ramifications resulting from the separation of spoken and written grammars in this respect have been noted above. We do not yet have a complete account of this phenomenon, which awaits further research.

Author's address:

Research Department of Linguistics

University College London

Chandler House

2 Wakefield Street

London

WC1N 1PF

UK

E-mail: a.weir@ucl.ac.uk

REFERENCES

- 3GPP. 1998. Technical realization of the Short Message Service (SMS). *3GPP Technical Specification* 03.40. <http://www.3gpp.org/ftp/Specs/html-info/0340.htm> (21 February 2009).
- Gerken, L. 1991. The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30(4), 431–51.
- Haegeman, Liliane. 1990. Understood subjects in English diaries. *Multilingua* 9(2), 157–99.
- Haegeman, Liliane. 1997. Register variation, truncation, and subject omission in English and in French. *English Language and Linguistics* 1(2), 233–70.
- Haegeman, Liliane, & Tabea Ihsane. 1999. Subject ellipsis in embedded clauses in English. *English Language and Linguistics* 3(1), 117–45.
- Haegeman, Liliane, & Tabea Ihsane. 2001. Adult null subjects in the non-pro-drop languages: two diary dialects. *Language Acquisition* 9(4), 329–46.
- Horse, Richard. 1998. *Null arguments in English registers: A Minimalist account*. BA thesis, La Trobe University.
- Hyams, Nina M. 1986. *Language acquisition and the theory of parameters*. Dordrecht: D. Reidel.
- Janda, Richard D. 1985. Note-taking English as a simplified register. *Discourse Processes* 8(4), 437–54.
- Nariyama, Shigeko. 2004. Subject ellipsis in English. *Journal of Pragmatics* 36(2), 237–64.

- Rizzi, Luigi. 1994. Early null subjects and root null subjects. In Teun Hoekstra and Bonnie Schwartz (eds.), *Language acquisition studies in generative grammar* (Language Acquisition and Language Disorders 8), 151–76. Amsterdam: John Benjamins.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in generative syntax*, 281–337. Dordrecht: Kluwer.
- Wells, John C. 2007. *English intonation: An introduction*. Cambridge: Cambridge University Press. (Originally published 2006).
- Zwicky, Arnold M., & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English N'T. *Language* 59(3), 502–13.

FOOTNOTES

- 1 I am indebted to Peter Ackema, Liliane Haegeman, and Caroline Heycock for their helpful comments and criticism during the writing of this article.
- 2 Although subjects can be dropped from many forms of written communication, throughout the present work 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.
- 3 This example is due to Liliane Haegeman (personal communication).
- 4 I am indebted to Neil Bennet for pointing out this possibility. This case perhaps also accounts for certain fixed exclamations such as *Can do!* and *Will do!*, which do not require subjects; possibly there is an implicit contrast here with these expressions' negative counterparts.
- 5 They acknowledge that 'given the low numbers of sentences with *wh*-fronting in our extracts, the absence of the null subject may be a sampling accident' (1999: 134).
- 6 By 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.
- 7 It 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.
- 8 If a contrast exists between these two examples, we can explain it by noting that phonotactic constraints in English rule out a nasal + plosive cluster at the beginning of a

phonological word, as in *'m gonna*, but permit a cluster /s/ + plosive, as in *'s gonna*.

9 Ungrammatical in my judgement, that is. I admit that intuitionistic data start to get hazy at this point. One avenue of further research would be to investigate corpora of informal spoken English to confirm or refute the analysis I present here.

10 Interestingly, *fictional* diary entries often *are* restricted in length. Haegeman & Ihsane (2001: 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 SPD 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 SPD is forbidden in e.g. root *wh*-questions.

11 Although having said that, this clearly isn't the full story; there is no semantic need for the *wh*-cleft (*What e did...*) in (33). 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?

12 There may be objections to this, 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.