

Reflections on English personal pronouns

J. Lambek, McGill University

When can an English personal pronoun represent a noun phrase in the same sentence? Assuming agreement in person, gender and number, we attempt to answer this question with the help of a purely syntactic criterion, depending on a superficial grammatical analysis of the sentence, which recognizes a few key constituents only, some subject to minor editing.

1. Personal pronouns and their features.

Personal pronouns exhibit the features person, number, gender and case, the first three of which are essentially determined by semantics, though subject to some inconsistencies. The English pronoun *we* is supposed to refer to a number of people, which must include the speaker and may or may not include the hearer. Yet, it can refer to the speaker alone when uttered by the queen, or to the writer alone in a scientific article. The pronoun *we* can even refer to the person addressed, when uttered by a doctor or nurse, as in

How are we this morning?

The pronoun *you* originally referred to a group of people including the hearer. In this capacity it may now be replaced by *you all* or *youse* in certain dialects. Today it replaces the obsolete *thou* (surviving only in the Quaker *thee*) in referring to a single hearer, although it still requires the plural inflection of the verb.

There is a tension between grammatical and semantical number, as illustrated by the following:

A number of people know Esperanto; but their number is small, even if it is not insignificant.

The gender in English is strictly determined by the sex of the person referred to, as opposed to the practice in some other European languages. For example, in both French and German, the word translating the English word *person* is feminine, even when referring to a male. In both these languages, the gender of the possessive pronoun agrees with the noun it modifies; in German it also agrees with the case. In English it only agrees with the sex of the possessor.

There are fewer distinguishable grammatical cases in English than in many other languages. For example, *he/she* refer to the subject (Latin nominative), *him/her* to the direct or indirect object (Latin accusative or dative), and *his/her* are in the possessive case (Latin genitive). Compare this with four cases in German and six in Latin.

Not only pronouns, but other noun phrases as well can denote persons. Clearly, *the old man* refers to a male, *the pretty girl* to a female, and *the members of the senate* to a plurality of persons of unspecified sex. Morphology tells us that *girls* and *members* are plural; in other languages, such as Latin, the morphology may also provide a clue to the sex of the person, which is left to semantics in English.

Leaving aside such questions as to whether Santa Claus, animals or ships are persons, we are still faced with a problem in connection with quantifiers. Which person or persons is referred

to by *no old man*, *a pretty girl* or *many members of the senate*? Concerning the first of these, one is reminded of Homer's cyclops, who believed that *Nobody* was the name of Odysseus, and of the medieval *Saint Nemo*, who was asserted to have ascended to heaven.

In much of what follows, we will make the blanket assumption that the noun phrases under consideration refer to persons of known gender. We will return to the question of quantifiers in Section 11 below.

2. Pronouns representing other noun phrases.

A personal pronoun occurring in a sentence may refer to the person or persons referred to by another noun phrase in the same sentence. I will say that the pronoun *represents* the noun phrase. The noun phrase may occur earlier or later, in which two cases one traditionally speaks of an *anaphoric* or *cataphoric* relation respectively. It may also enclose the pronoun, as in

a girl who lost her ring.

In any case, the pronoun should agree with the noun phrase it represents in number and gender in principle, although there may be a tension between syntax and semantics, as discussed in Section 1.

Aside from such agreement, when can a pronoun represent another noun phrase in the same sentence (or even in the same text)? This is something never taught in language courses, yet it is a question which has attracted much lively discussion among linguists in the second half of the twentieth century. McCawley [1988] cites contributions to this question by Bach [1986], Carden [1986], Chomsky [1981], Hankamer and Sag [1976], Jackendoff [1972], Lakoff [1968], Postal [1971], Reinhard [1976, 1983], and Ross [1967, 1986]. I will rely on McCawley's summary and evaluation of the various claims and counter-claims and refer the reader to his bibliography.

The general consensus seems to be that the answer to this question on the syntax-semantics boundary need not be taught and is part of the speaker's innate "universal grammar". There also seems to be a consensus that, aside from feature agreement, the answer can be formulated in syntax alone. It is usually expressed in terms of the prevailing *X-bar* theory, ultimately in terms of the rather technical notion of "*C-command*".

Although arriving on the battle-field much after the battle, I decided to take another look at this question. It seems to me that some of these technicalities are not really needed here, nor is the widely accepted geometric representation of sentence structure by planar trees with labeled nodes. I admit that these trees offer a convenient tool to the grammarian, but I do not believe that they have any psychological reality. Of course, they are equivalent to a labeled bracketing of the sentence; but I do not think that a complete constituent analysis of the sentence is necessary for the present purpose and I claim that the recognition of certain key constituents suffices.

Although the kind of grammar I now prefer is a "pregroup grammar" (see e.g. Lambek [2004]), this will play no rôle in the present discussion, except for a brief digression in Section 13.

To simplify my exposition, I will confine attention to third person pronouns. After all, noun phrases in the first and second person are quite rare. Moreover, I will defer the often impersonal pronoun *it*, which poses special problems unrelated to the present investigation, to Section 12.

It will be convenient to treat reflexive pronouns as noun phrases, e.g.

himself = his + self.

(I am indebted to Ed Keenan for pointing out that this widely accepted analysis may in fact be historically justified.) I will assume that the reader can recognize certain key constituents of a sentence without carrying out a complete constituent analysis.

3. Certain key constituents.

The criterion for representability of noun phrases by pronouns I wish to propose will depend on recognizing certain key constituents of the sentence:

(0) Direct sentences, both declarative and interrogative, such as the following.

*John likes her,
did Jane kiss him?
Whom did she kiss?*

(1) Noun phrases, e.g.

*the pretty girl,
his girlfriend,
the girl [whom] he loved.*

(2) Indirect sentences, e.g.

*that she kissed him,
whether she arrived,
whom she will see.*

These may of course be viewed as special kinds of abstract noun phrases.

(3) Indirect quasi-sentences, e.g.

*whom to see,
[for her] to arrive,
his/him arriving.*

These may also be viewed as abstract noun phrases.

(4) Subordinate clauses, e.g.

*although she saw him,
while kissing him.*

(5) Spatio-temporal prepositional phrases, such as introduced by the prepositions

above, on, after, with,

but not by

for, to, about, by,

The preposition *with* may be treated as referring to spatio-temporal proximity in some contexts, but not in others, when it denotes instrumentality.

As it will turn out, the distinction between spatio-temporal and other prepositional phrases is required for our proposed criterion, but it ought to come up for other reasons as well. For example, look at the unorthodox word order:

Above his head passed a plane.
On her forehead was painted a caste-mark.
After the deluge arose a new generation.
With John came his girlfriend.

All these sentences allow the optional inclusion of the word *there* before the verb. On the other hand, the following are not normally admitted:

- * For Jane bought he a book.
- * To Jane promised he a ring.
- * About Jane spread he a rumour.
- * By John was written a book.

The first three of these would not even allow *there* to be inserted.

4. Representability criterion.

In most of what follows, we will assume that noun phrases under consideration represent persons whose sex is known. Here is the proposed representability criterion:

CRITERION. A personal pronoun can represent a (personal) noun phrase in the same sentence if and only if it agrees with the noun phrase in person, number and gender, and either

(A) the pronoun is properly contained in a key constituent of type (1) to (5) and the noun phrase is not,

or

(B) the noun phrase is properly contained in a key constituent of type (0) to (4) and the pronoun occurs *after* this constituent.

There may be some ambiguity about what is meant by “the same sentence” in the above criterion, since one sentence may occur as a constituent of another. So let us declare that the criterion already applies to the smallest sentential constituent containing both the pronoun and the noun phrase in question. As a consequence of conditions (A) and (B) we state:

(C) If the pronoun and the noun phrase are both properly contained in the same minimal key constituent then the former cannot represent the latter.

For example, *he/him* cannot represent *John* in the following:

John likes him,
John talks to him,
He likes John,
He talks to John.

Our representability criterion is subject to some provisos that will be discussed in later sections. Some of the key constituents may require slight editing before the criterion is applied

(see e.g. sections 7 to 9). Condition (A) may require a proviso that the noun phrase, when occurring after the pronoun, refers to a definite person (see Section 11).

Our criterion can also be applied to texts longer than sentences. Moreover, it can be extended to anaphoric expressions other than personal pronouns, such as *the old man*. However, we will refrain from discussing these generalizations here, in order to keep the present exposition within reasonable bounds.

5. Illustrating Condition (A).

In the following sample sentences and noun phrases we follow the usual procedure of labeling constituent noun phrases and pronouns by the same numerical subscript if and only if they are intended to denote the same person. We write (A1) to refer to constituent (1) when illustrating Condition (A), etc. We will leave (A3) and (A5) to Sections 7 and 9 respectively, as they require a substantial amount of editing.

(A1) Noun phrases.

John₁ likes (his₁ girlfriend).
(His₁ girlfriend) kissed John₁.
John₁ married (the girl he₁ loved).
(The girl he₁ loved) married John₁.

In the following sentences I have put the subscript 2 on the relative pronoun *who* and the optional relative pronoun *whom*:

John met (a girl who₂ had lost (her₂ ring))₂,
John met (the woman [whom₂] (her₂ daughter) loves)₂.

Indeed, these relative pronouns could be generated with the help of an “inflector” Rel as follows:

Rel *she₂* → *who₂*,
 Rel *her₂* → *whom₂*.

This analysis makes it clear why (A1) does not license

**(a girl who₂ loves her₂)₂,*
 **(the girl [whom₂] she₂ loved)₂,*

both of which are ruled out by Condition (C).

The same device will be helpful to explain

(the man whose₁ daughter loves him₁)₁.

Here the relative clause is restrictive: it modifies the noun *man* and not the noun phrase *the man*, which is therefore not a constituent (hence (B0) does not apply). However, if we analyze

Rel *his₁* → *whose₁*,

the implicit noun phrase *his₁ daughter* becomes a proper constitutional and (A1) can be applied to justify *his₁*. Thereafter, (A1) can be invoked again to justify *him₁*.

(A2) Indirect sentences.

Here are two examples:

*John₁ wished ([that] he₁ had met Jane sooner),
(That he₁ met her) was John'₁s luck.*

Recall that the complementizer *that* is optional in the object position, but obligatory in subject position. Note that the second sentence is more commonly rephrased thus:

It was John's luck that he met her.

We will return to this formulation in Section 12.

(A4) Subordinate clauses.

Here are three examples:

*John₁ snores (when he₁ sleeps);
(Although he₁ snored), John₁ slept quite well;
Jane₂ slept beside John₁, (while she₂ heard him₁ snore).*

In the first of these, we need not decide whether the subordinate clause modifies the sentence or only the verb phrase.

6. Illustrating Condition (B).

Again, we postpone (B3) and (B5) until later.

(B0) Direct sentences.

*(John₁ met Jane₂) and she₂ kissed him₁
..... [he₁] kissed her₂,
..... *[he₁] kissed him₁.*

Here *[he₁] *kissed him₁* violates (C).

*(Did John₁ kiss Jane₂) or [did he₁] ignore her?
..... did she₂ ignore him?
..... *[did he₁] kiss him₁?*

Evidently the material in square brackets should be inserted editorially before citing (B) or (C).

(B1) Noun phrases.

*(John'₁s girlfriend) kissed him₁;
(The girl [whom] John₁ liked) kissed him₁;
(Jane, whom John₁ liked) kissed him₁;
(The girl who liked John₁) kissed him₁.*

(B2) Indirect sentences.

*(That Jane₂ kissed John₁) pleased her₂, but not him₁;
(Whom Jane₂ wanted to kiss) was up to her₂.*

(B4) Subordinate clauses.

(When Jane kissed John₁) he₁ objected.

7. Editing indirect quasi-sentences.

Some serious editing is required to handle the following sample sentences:

*X let (Y kiss Z),
X saw (Y kiss Z),*

where X, Y and Z are noun phrases. Here a pronoun in place of Z can represent X but not Y , as we would expect if we think of Y as the subject of the quasi-sentence in parentheses, even though it is in the accusative case. However, a pronoun in place of Y also cannot represent X , so it would help to think of Y also as an object complement of *let*. We resolve this dilemma by writing

*X₁ let Y₂ (#₂ kiss Z_i),
X₁ saw Y₂ (#₂ kiss Z_i),*

where #₂ is an invisible copy of Y_2 . Now we can have $i = 1$ but not $i = 2$, which would violate Condition (C).

The same construction applies to verbs of causation

let, make, help, have

and perception

see, hear, feel,

as in

*John₁ helped Jane₂ (#₂ cook for him₁),
John₁ heard Jane₂ (#₂ call him₁).*

What do verbs of causation and perception have in common? There seems to be a folk-philosophy which considers perception as a kind of causation. Some early philosophers, e.g. Empedocles, believed that light-rays pass from the eye to the object seen. Even more recently, Bishop Berkeley proclaimed that existence is caused by perception.

Similar editing will help out where the infinitive requires *to*:

*John₁ hoped for Jane₂ (#₂ to kiss him₁);
John₁ wanted [for] Jane₂ (#₂ to kiss him₁),*

where the complementizer *for* is optional. However, when the complementizer is present, a different analysis is possible, one in which *for Jane to kiss him* is viewed as an abstract noun phrase:

John₁ wanted (for Jane₂ (#₂ to kiss him₁)).

If so, the following would also be admissible:

John₁ wanted ([for him₁] (#₁ to kiss Jane₂));

but here *for him* is usually omitted to yield

John₁ wanted (#₁ to kiss Jane₂).

Here the invisible noun phrase $\#_i$ plays the rôle of Chomsky's PRO. It would seem that the subscript i must agree with that of the last noun phrase preceding $\#_i$, at least as long as the abstract noun phrase appears in object position.

Here are some examples when (A) is applied several times:

*Jane₂ wondered (whether she₂ should let John₁ (#₁ kiss her₂)),
John₁ wondered (whether Jane₂ wanted him₁ (#₁ to kiss her₂)),
John₁ expected Jane₂ (#₂ to let him₁ (#₁ kiss her₂)),
Jane₂ persuaded John₁ (#₁ to ask her₂ (#₂ to kiss him₂)).*

When the indirect quasi-sentence appears in subject position, the complementizer is obligatory. Thus we have:

*(For Jane₂ (#₂ to kiss him₁)) pleased John₁,
(For Jane₂ (#₂ to kiss John₁)) pleased him₁,*

by (A3) and (B3) respectively. We even have:

(For Jane₂ (#₂ to kiss him₁)) pleased her₂,

where both (A3) and (B3) are invoked.

(A3) alone is invoked in

- (a) *(For her₂ (#₂ to kiss him₁)) pleased John₁,*
- (b) *([For her₂] (#₂ to kiss him₁)) pleased Jane₂.*

However, *For her₂* may not be deleted in (a):

- (c)* *(#₂ to kiss him₁) pleased John₁.*

It would seem that, when the indirect quasi-sentence appears in subject position, the subscript of $\#_i$ must agree with that of the first noun phrase following the indirect quasi-sentence. Of course

**(#₁ to kiss him₁) pleased John₁*

is ruled out by (C), as long as $\#_1$ is viewed as a noun phrase.

8. Quasi-sentences from participles.

Quasi-sentences built from participles in place of infinitives may be handled similarly, provided we introduce inflectors:

Poss = possessive (genitive) case,
Part₁ = present participle

as in the rewriting rules

Poss *John* → *John's*,
Poss *him* → *his*,
Part₁ *kiss* → *kissing*.

It is convenient to regard *his* as the possessive of *him* rather than of *he*, in order to make the analogy between participles and infinitives work better. I suggest that the inflector Poss replaces the complementizer *for* and the inflector Part₁ replaces the word *to* which completes the infinitive, thus allowing a comparison between

for Jane₂ (#₂ to kiss him)

and

Poss *Jane₂ (#₂ Part₁ kiss him) → Jane'₂s (#₂ kissing him).*

We can now analyze the sentences

*John₁ denied Jane'₂s (#₂ having kissed him₁),
Jane₂ denied [her₂] (#₂ having kissed John),*

where *her₂* would normally be omitted.

Sometimes the possessive inflector is not admitted, as in

John saw Jane₂ (#₂ kissing Bill),

so we ought not say

**John saw [his] (#₁ kissing Jane),*

and should say instead

John₁ saw (him₁self) (#₂ kissing Jane).

The complementizer Poss is obligatory when the indirect quasi-sentence appears in subject position:

- (a) *(John'₁s #₁ kissing Jane₂) displeased her₂/him₁,*
- (b) *(His₁ #₁ kissing Jane₂) displeased John₁/her₂,*
- (c) *(#₁ kissing Jane₂) displeased John₁/*her₂.*

Here (a) is justified by (B1), (b) by (A1)/(B1) and (c) is obtained from (b) by deleting *his₁*. Recall that, according to the remark at the end of Section 7, the subscript of #₁ in (c) must agree with that of the first noun phrase following the indirect quasi-sentence, hence the object *her₂* is not admissible in (c).

It should be pointed out that the indirect quasi-sentences built from a participle ought not be confused with the noun phrase such as

the kissing of girls

where *kissing* is a genuine noun.

9. Spatio-temporal prepositional phrases.

The story of prepositional phrases is not too clear cut. They may modify nouns, verbs, verb phrases or sentences, as in the following examples:

*John looked at the (girl with a telescope),
John (looked with a telescope) at the girl,
John (looked at the girl with a telescope),
With a telescope, John looked at the girl.*

In all but the first of these sentences, the preposition *with* expresses instrumentality. In the first example, it may express spatio-temporal proximity; but this is more clearly so in

John₁ looked at (the girl with him₁),

which may already be justified by (A1), since *him* occurs properly inside a noun phrase. However, I would invoke (A5) to justify

John₁ took the girl (with him₁),

where the presumably spatio-temporal prepositional phrase modifies the verb, though discontinuously.

Clearer applications of (A5) are found in

*John₁ saw an eagle (above him₁),
(After him₁), John₁ expected a deluge.*

(A5) does not apply to

*John₁ bought a book (for *him₁),*

since *for* is not spatio-temporal.

But why does (A5) not apply to the following?

*John₁ was (beside *him₁).*

We attempt an explanation by editing the spatio-temporal prepositional phrases in a manner similar to what was done above (in Section 7) for indirect quasi-sentences. We will insert an invisible noun phrase $\#_i$ before the preposition, as in

*John₁ saw (an eagle)₂ ($\#_2$ above him₁);
(John'₁s girlfriend)₂ slept ($\#_2$ beside him₁);
($\#_2$ beside him₁), John₁ saw a spider₂;
($\#_2$ above him₁), (John'_{s1} kite)₂ flew by.*

We will stipulate that the subscript of $\#_i$ must agree with the object of a transitive verb, but with the subject of an intransitive verb, the verb being the principal verb of the sentence to the right or left of which the prepositional phrase occurs.

We list some further examples, to be justified by (A5):

*John₁ took (the girl)₂ ($\#_2$ with him₁);
($\#_2$ with him₁), John₁ took (enough provisions)₂;
(John'₁s girlfriend)₂ did not sleep ($\#_2$ beside him₁);
(his₁ girlfriend)₂ did not sleep ($\#_2$ beside John₁).*

The last two of these examples can also be justified by (B1) and (A1) respectively.

Since $\#$ is viewed as a noun phrase, Condition (C) rules out

**John took Jane₂ ($\#_2$ with her₂),
($\#_1$ without John₁) ($\#_1$ he could not sleep).

Here are some more elaborate examples:

Jane let John₁ (#₁ take (the girl)₂ (#₂ with him₁)),
Jane₂ let John₁ (#₁ take the girl from her₂),
Jane let John₁ (#₁ take the girl for him₁self).

In the last two of these, the prepositions *from* and *for* are not spatio-temporal. Neither is the preposition *of* in the following:

(A picture of John₁) was painted by him₁,
(A picture of him₁) was painted by John₁,

which may be justified by (B1) and (A1) respectively, although we might prefer *him₁self* in place of *him₁* in both examples.

10. Examples invoking two rules.

We look at a few examples where more than one instance of our criterion is invoked.

Although John₁ liked her₂, he₁ did not kiss Jane₂.

To justify *her₂* we use (A4), for *he₁* we use (B4).

His₁ fondness for Jane₂ led John₁ to kiss her₂.

Here (A1) justifies *his₁* and (B1) justifies *her₂*.

When he₁ kissed Jane₂, John₁ fell in love with her₂.

(A4) justifies *he₁* and (B4) justifies *her₂*.

When John₁ met her₂, he₁ kissed Jane₂.

(A4) justifies *her₂* and (B4) justifies *he₁*.

(The man she₂ loved)₁ married (a girl who₂ did not love him₁)₂.

Here (A1) justifies both *she₂* and *him₁*.

(His₁ girlfriend)₂ let John₁ (#₁ kiss her₂).

Here (A1) justifies *his₁* and (A3) justifies *her₂*.

(His₁ girlfriend)₂ asked John₁ (#₁ to kiss her₂).

This is quite similar to the above.

Sometimes applying two rules can lead to ambiguities. While gender agreement allows us to unambiguously interpret

John₁ told Jane₂ that he₁ could see her₂

and

John₁ told Jane₂ that she₂ could see him₁,

there are no clues for distinguishing between

John₁ told Bill₂ that he₁ could see him₂

and

John₁ told Bill₂ that he₂ could see him₁.

However, if we replace *could* by *would* or *should*, people seem to prefer

John₁ told Bill₂ that he₁ would see him₂.

and

John₁ told Bill₂ that he₂ should see him₁.

I guess the reason is that people subconsciously paraphrase these two sentences as follows:

John₁ told Bill₂ : I will see you,

and

John₁ told Bill₂ : you shall see me,

translating the indirect quote into a direct one. These sentences express John's wish or expectation respectively.

The situation is reversed if we replace the indirect affirmative sentence by an indirect interrogative one. Now people seem to prefer

John₁ asked Bill₂ whether he₁ should see him₂

and

John₁ asked Bill₂ whether he₂ would see him₁.

These are presumably paraphrased as follows:

John₁ asked Bill₂ : shall I see you?

John₁ told Bill₂ : will you see me?

These sentences express John's request for information about Bill's wish or expectation respectively.

I wonder whether people suffering from Asperger's syndrome, who are alleged to have difficulty visualizing other people's mental states, exhibit the same preferences.

There are some ambiguities where I don't see a preferred interpretation:

(John₁'s brother)₂ met (his_i friend),

where $i = 1$ or 2 , each by (B1) or (A1).

(His_i brother) met (John₁'s friend)₂,

where $i = 1$ or 2 , each by (A1).

(The man John₁ knew)₂ met (a girl he_i liked),

where $i = 1$ by (A1) or (B1), or $i = 2$ by (A1).

(Jane₂ met (the woman who₃ had lost (her_i ring))₃,

where $i = 2$ or 3 , each by (A1).

11. Problems with quantifiers.

Surprisingly, the presence of quantifiers usually makes little difference to the representability of noun phrases by pronouns. In

(The girl)₂ liked (her₂ boyfriend)

we can easily replace the definite article *the* by the indefinite *a* or by the quantifiers *some* or *no*, even though we find it hard to explain whom

a/ some/ no girl

refers to. However, I find it difficult to carry out the same replacement on

(Her₂ boyfriend) liked (the girl)₂

and to accept the subscripts in

(a) **(Her₂ boyfriend) liked (some girl)₂*,

and similarly when *some* is replaced by *a* or *no*.

At first sight, these examples suggest a proviso to Condition (A) in our representability criterion:

provided the noun phrase, when occurring *after* the pronoun, is definite (meaning that it refers to a definite person).

On the other hand, the following seem to me to be acceptable:

(b1) *(For John to kiss her₂) is expected by (some girl)₂*;

(b2) *(That he₁ likes girls) is (no man's)₁ fault*;

(b3) *(Even when he₁ snores), (a man)₁ can sleep well*.

Presumably, the indefinite article in the last example can be interpreted as a universal quantifier, and the same is true in

If he₁ likes her₂, (a man)₁ may kiss (a woman)₂.

Also the following suggests an interpretation with two universal quantifiers:

(Unless he₁ likes her₂), (no man)₁ should kiss (a woman)₂.

I must confess that, at this stage, I don't know how to restrict Condition (A) so that (a) is disallowed, but (b1) to (b3) are admitted. I suspect that this can be done with the help of Discourse Analysis (see Kamp and Reyle [1993] and Preller [to appear]).

12. Some remarks about the pronoun *it*.

The pronoun *it* deserves special consideration. Technically, a third person pronoun, it has no gender, though some people would say it is of neutral gender, and it does not usually denote a person, unless perhaps a small child. More commonly, it will denote an inanimate object or an abstract one. Often it may denote nothing at all, as in *it is raining*. As long as

it denotes something, the above consideration of representability will also apply, provided we replace “person” by “thing”.

I am here interested in another function of the pronoun *it*. There is some reluctance to use indirect sentences or quasi-sentences as subjects of sentences. Consider for example

- (1) *That people will die is certain;*
When they will die is uncertain;
For people to have to die is sad;
Whether to commit suicide is an option.

Though acceptable, these sentences may be replaced by

- (2) *It is certain that people will die;*
It is uncertain when they will die;
It is sad for people to have to die;
It is an option whether to commit suicide.

One reason for the reluctance to employ (1) is that such constructions, when iterated, quickly exceed Miller’s [1956] limit on the number of chunks of information in the short-term memory. Consider, for example, the sentences with increasing complexity:

- (3) *War sucks;*
That war sucks sucks;
That that war sucks sucks sucks;
That that that war sucks sucks sucks sucks.

If the reader does not like the iterated appearance of the verb *sucks*, she can replace it by *is bad*, *is true*, *is known* etc successively.

If the reader is interested, she will find in Section 13 an argument to show that uttering the last of the above requires a temporary storage of nine chunks of information, in the terminology of Miller’s [1956], who claims that people can hold at most seven (plus two?) chunks of information in their short-term memory.

On the other hand, we can replace this hard to process sentence by

- (4) *It sucks that it sucks that it sucks that war sucks,*

where the short-term memory need only hold six chunks.

13. A mathematical detour.

It may be of interest to see how (3) and (4) of Section 12 are analyzed in the kind of grammar proposed by Harris [1977], a rudimentary form of the pregroup grammar I now prefer [2004]. To make things as simple as possible, let us assume that we are given two *basic* types N (for noun phrase) and S (for statement), from which other *simple* types N^ℓ, N^r, S^ℓ, S^r are formed. These generate a partially ordered monoid (semigroup with unity element 1) of *types*, whose partial order is denoted by an arrow, subject to the contraction rules

$$X^\ell X \rightarrow 1, \quad XX^r \rightarrow 1,$$

where $X = N$ or S .

Now assign the following types (elements of the monoid) to English words:

$$\textit{war} : N, \quad \textit{sucks} : N^r S, \quad \textit{that} : NS^\ell.$$

Then

$$\begin{array}{c} \textit{War sucks} \\ \underline{N (N^r S)} \rightarrow S \end{array}$$

is seen to be a statement by a simple calculation on types. To see that the last sentence of (3) is well-formed, we require the calculation

$$\begin{array}{cccccccc} \textit{That} & \textit{that} & \textit{that} & \textit{war} & \textit{sucks} & \textit{sucks} & \textit{sucks} & \textit{sucks} \\ (NS^\ell) & (NS^\ell) & (NS^\ell) & N & (N^r S) & (N^r S) & (N^r S) & (N^r S) \end{array} \rightarrow S$$

Looking at the first five words, we arrive at the type

$$(NS^\ell) (NS^\ell) (NS^\ell) \underline{N (N^r S)}$$

before any contraction can take place, thus listing nine simple types in temporary storage. Elsewhere [2004], I have suggested that simple types should be identified with Miller's [1956] chunks of information, so we have reached his reluctant upper limit of what the short-term memory can bear.

On the other hand, let us assign to *it* the type

$$it : S\bar{S}^\ell S^\ell N,$$

then we can calculate successively

$$\begin{array}{c} \textit{It sucks} \\ (S\bar{S}^\ell S^\ell \underline{N}) (N^r S) \end{array} \rightarrow S\bar{S}^\ell \underline{S^\ell S} \rightarrow S\bar{S}^\ell$$

$$\begin{array}{c} (\textit{It sucks that}) \\ (S\bar{S}^\ell) (\underline{\bar{S} S^\ell}) \end{array} \rightarrow S\bar{S}^\ell$$

$$\begin{array}{c} (\textit{It sucks that}) \textit{it} \\ (S\bar{S}^\ell) (\underline{S\bar{S}^\ell S^\ell N}) \end{array} \rightarrow S\bar{S}^\ell S^\ell N$$

$$\begin{array}{c} (\textit{It sucks that it}) \textit{sucks} \\ (S\bar{S}^\ell S^\ell N) (\underline{N^r S}) \end{array} \rightarrow S\bar{S}^\ell \underline{S^\ell S} \rightarrow S\bar{S}^\ell$$

etc. The reader will easily check that calculating the type of (4) of Section 12 never requires holding more than six simple types in temporary storage.

14. Some belated remarks about the literature.

It goes without saying that much of the literature on pronouns and other anaphora is based on the revolutionary ideas of Noam Chomsky. I have referred here only to his notion of PRO and have ignored the other empty categories discussed in his paper with Howard Lasnik, reprinted in Chapter 1 of Chomsky [1995]. I guess the possible noun phrase $\#_i$ introduced here for editing indirect quasi-sentences and spatio-temporal prepositional phrases should fall under this heading.

Although the present account does not invoke the ingenious technical device of *C-command* (see McCawley [1998]), it is subject to the same objections raised there.

For example, McCawley mentions

**(What he₁ denied) was that Nixon₁ was a crook,*

which contradicts a prediction made with the help of C-command. It also contradicts (A2), if *what he denied* is viewed as an indirect question, or (A1), if it is preferably taken to be a noun phrase equivalent to *that which he denied*.

I will attempt one explanation for why the present analysis fails. Deleting the words *what* and *was* in the above sentence, we obtain what might be called a “sub-sentence”

**he₁ denied (that Nixon₁ was a crook).*

Here (A) is no longer applicable, because *he* is not properly contained in a key constituent which excludes the indirect sentence, and (B) is not applicable because it precedes the latter.

The above objection does not apply to

(That Nixon₁ was a crook) was (what he₁ denied),

which is admissible by (B2) or (A1). It also does not apply to

(What Nixon₁ denied) was (that he₁ was a crook),

which is acceptable by (B1) or (A2), where even the sub-sentence

Nixon₁ denied (that he₁ was a crook)

is acceptable by (A1).

We may make our attempted explanation explicit by stipulating:

(D) For a pronoun to represent a noun phrase in a sentence it is necessary that it does so in any sub-sentence in which both occur.

Here “sub-sentence” is to be interpreted rather liberally to allow for the deletions of *what* and *was* above, as well as for the “minimal sentential constituent” mentioned in the remark immediately preceding Condition (C) in section 4.

15. Conclusion.

We have attempted to answer the question when a personal pronoun can represent a noun phrase in the same sentence. Aside from the obligatory agreement in person, number and gender, we have postulated two purely syntactic conditions, which involve only a superficial recognition of certain key constituents of the sentence. We have avoided reliance on a more sophisticated grammatical analysis based on the popular geometric representations by planar trees, in particular *X*-bar theory, which is often invoked in this context.

However, some preliminary editing of the key constituents was necessary, in particular, when looking at quasi-sentences, such as those corresponding to Latin “accusative with infinitive”, and spatio-temporal prepositional phrases. We had to introduce some invisible noun phrases, vaguely resembling the empty categories discussed by Chomsky [1981, 1995]. A more complete treatment of prepositional phrases has not been attempted here.

The noun phrases to be represented by personal pronouns were assumed to denote persons, hence to be definite. Surprisingly, the technique we employed usually also applies to indefinite and quantified noun phrases, which cannot be interpreted as referring to a particular person. There are exceptions to this when the noun phrase occurs on the right of the pronoun. A complete understanding of what goes on here still escapes me and presumably should be investigated with the help of Discourse Analysis (see Kamp et al. [1993] and Preller [t.a.]).

To keep this paper within reasonable bounds, we have confined attention to third person pronouns, mostly in the singular, although other noun phrases may behave like pronouns in admitting internal reference to noun phrases occurring in the same sentence, or even in the same text.

REFERENCES

- N. Chomsky, *Lectures on government and binding*, Foris Publications, Dordrecht, 1981/1983.
- N. Chomsky, *The minimalist program*, The MIT Press, Cambridge, Mass., 1995.
- Z. Harris, *Mathematical Structure of Language*, Interscience Publishers, New York, 1977.
- H. Kamp and U. Reyle, *From discourse to logic*, Kluwer Academic Publishers, Dordrecht, 1993.
- J. Lambek, A computational algebraic approach to English grammar, *Syntax* 7(2004), 128-147.
- J.D. McCawley, *The syntactic phenomena of English*, Volume I, The University of Chicago Press, 1988.
- G.A. Miller, The magical number seven plus or minus two: Some limits to our capacity for processing information, *Psychological Review* 63(1956), 81-97.
- A. Preller, Towards discourse representation via pregroup grammars, *Journal of Logic and Information*, to appear.

ACKNOWLEDGEMENTS

I would like to thank Bernie Lambek and Anne Preller for their helpful comments on the manuscript.