Do grammars change when they leak?

Citation for published version (APA):

Published in:
New perspectives on English historical linguistics

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Do grammars change when they leak?¹

0. Introduction

The organisation of this programmatic paper is as follows. Limitations of space mean that the treatment is necessarily both selective and allusive. Some general remarks about language change (§1) are followed by criticism of the conventional view of synchronic grammar (§2) and a consideration of reanalysis in diachrony (§3). I then discuss ‘leakiness’ of the grammar of the English NP in synchronic terms (§4) and suggest graduated change as a plausible alternative to reanalysis for some kinds of diachronic change (§5). The appearance and loss of intermediate forms are considered in §§6-7, and my example of reanalysis reappears in the conclusion (§8).

1. A typology of language change²

All living languages are subject to change. How do they change? Most linguists now distinguish between actuation or innovation on the one hand and propagation or diffusion on the other. Diffusion is the spread of a variant from the point where it has become an option for a number of speakers. Diffusion of change requires the prior existence of variants — alternative ways of saying the same thing. Some variation appears to be stable
over long periods (Raumolin-Brunberg 2002): variation is a necessary but not a sufficient condition for change to occur.

And why does language change? The answer to this question is, of course, dependent on the theoretical position of the analyst. If the most salient property of language is its grammar, an internalised set of rules unconsciously built up and used by an individual speaker-hearer, then language change may be seen as a consequence of new generations inducing a slightly different grammar from that internalised by their parents’ generation, perhaps because of a slightly changed preponderance of some usage, the cause of which is not really grammatical in origin but some contingent ‘performance’ factor. The new grammar in turn leads to a further change in the output of its speakers, and so things move on. The process of language acquisition in childhood will be critical, and the favoured form of analysis will (usually) be formal and structural. Alternatively, if language is something which crucially belongs to and exists in a speech community, then speaker interaction and relative social status may be the fundamental engines of language change. Speakers may adjust their usage to (or against) community norms throughout adolescence and perhaps beyond, and change is not confined to the acquisition process. The requisite analysis will be sociolinguistic and statistical. Then again, if speakers and hearers are regarded as autonomous individuals, anxious above all to maximise their communicative efficiency, yet other considerations may be identified, typically involving speaker intentions.

We can group three main types of force for change under the headings structural, social and functional. Each has its own champions in the literature, though the dividing lines are not always clear-cut. Beyond these three broad categories there are
extralinguistic factors to consider too. Here I will mainly be looking at change from the structural and functional points of view. For a fuller discussion see Croft (2000).

2. Synchrony without leakage

Many approaches to synchronic language description start from the working hypothesis that grammars don’t leak. That is, there are a certain number of parts of speech or word classes (the number may depend on the theory or the analyst), and every word in a particular utterance of a sentence belongs to just one of them. Constituent structure analysis gives a unique bracketing of these words. (In a theory with movement there may be a series of structures in the derivational history.) Utterances which appear not to fit are either

- ignored,
- or arbitrarily made to fit,
- or assumed to be in need of a better analysis which remains to be discovered,
- or relegated to ‘mere’ performance (or E-language),
- or — in extremis — accepted as anomalous but regarded as minor and peripheral parts of the language system.

This is unfair to many linguists, but as far as mainstream formal syntax is concerned, not entirely unjustified.

What utterances don’t fit? Here is a simple example. If we consider the sentences

(1) Jim bought a watch in the market.
(2) Jim put his bicycle in the shed.
(3) Jim gave money to charity.

(4) Jim took pleasure in mathematics.

(5) Jim paid attention to his teacher.

(6) Jim set fire to the car.

we find that each has different properties despite its superficially similar NP$_1$ V NP$_2$ PP structure. For example, the PP in (1) is an adjunct, whereas the PPs in (2) and (3) are generally regarded as complements; in (3), though, but not (2), the PP commutes with an indirect object. The PP of (4) is syntactically somewhat equivocal — adjunct or complement? — while at the lexical level one might recognise take pleasure in as a transitive group-verb equivalent to enjoy. Similar remarks could be made about (5), though the syntax is not identical: (5) passivises much more readily. So does (6), but now the verbal idiom is closer to being frozen and NP$_2$ is not open to modification; indeed we could ask whether it makes sense to assign fire to the word class Noun at all.

We could go on adding sentences with shades of variation and subtly different behaviour until a structuralist model ran out of ways of distinguishing items. The working hypothesis cannot be upheld.

3. Reanalysis

Even a watertight system of synchronic analysis still gives many possibilities of change by creation of variants and selective preference among them. One such route is by reanalysis. We need an example. There was (and still is) a prepositional verb run over:

(7) The rabbit ran over the meadow.
(8) My car ran over a bottle (lying in the road).

(9) \[ \text{VP ran_{intr} } [\text{PP over [NP a bottle]}] \]

*Over* is a preposition. However, the same linear string, *ran over a bottle*, is in principle open to reanalysis in English so that *over* becomes an adverbial particle in a transitive phrasal verb:

(10) \[ \text{VP ran_{trans} } [\text{part over [NP a bottle]}] \]

(As it happens, no one has come up with a wholly satisfactory formal analysis of either prepositional or phrasal verbs, so the reader’s pet analyses won’t necessarily be the same as mine in (9) and/or (10). All that is needed is agreement that analyses of the two should be different from each other.)

Why should this reanalysis take place? It is actually unlikely in the context of (8) but far more likely if the pragmatic situation is:

(11) My car ran over a bottle (standing in the road).

This time the bottle is definitely affected by the action: it will be knocked over and in all probability broken. Hence the re-interpretation of *over*: it need no longer refer to the trajectory of the car across and above the obstacle; now it can refer to the trajectory of the bottle away from the upright position and is resultative. Thus *run over* is no longer a prepositional verb like *run into*, but a phrasal verb like *knock over*. It has been reanalysed, in this instance both semantically and syntactically. Of course, sadly, we take more interest in such things when the victim is human:

(12) A drunk driver ran over two pedestrians.

We see the change semantically in that the car need not literally pass on top of its victims. And we see it syntactically in that we will now also begin to attest sentences like
A drunk driver ran two pedestrians over. not previously possible.

I chose this rather grisly example because the reanalysis is so intuitively clear. Nevertheless it would be unfair to let this example represent all syntactic reanalyses, (a) because this reanalysis is clearly not purely a matter of syntax, and (b) because even without the change in run over (which is historically attested, incidentally), such everyday patterns as go for and work out are rather embarrassing for a syntax without constructions. On the other hand, whether reanalysis ever is purely syntactic is questionable (Bill Croft, personal communication, 2 Dec. 2002). I cannot enter into this question here.

What is the mechanism of reanalysis? I don’t mean the technical mechanism, which will depend on the particular syntactic model in use; I mean the social mechanism. In diachrony I take it to be the attribution by a younger generation of choices from the watertight grammar for the analysis of some pattern which differ from the analysis of the older generation, whether different word class assignment or different structure or both. Remember: change implies variation. The variants in this case are implicit ones, alternative structures for the same explicit form, and on the individual level the change occurs during acquisition. Statistically the change will enter the language as the new analysis gradually prevails over time. This is essentially the mechanism of Lightfoot (1979) and many succeeding works.

Another mechanism allows for more than one grammar in a population at the same time. The idea of competing grammars is found in Pintzuk (1991) and other works. It was devised in order to account for explicit variation in synchronic grammar, but I
guess it could be used for the kind of implicit, underlying variation in the string analysed above. One of my problems with that approach is the prospect of a snowballing multiplicity of grammars to account for all the variation encountered in practice, and allowing it to deal with lexical variation like this would exacerbate that problem enormously.

Yet another approach has been tried for a specific historical problem by Wim van der Wurff (1992), where he imagines a scenario of contact between dialects (or social networks), in which an innovation in one group has to be interpreted in a particular way by speakers of another group with a different grammar who encounter it, leading to reanalysis by them and starting a chain of developments as one group’s output is reacted to by another. Something similar was recently suggested by Anthony Warner for the history of *do* (2002). Notice here that reanalysis and change is not confined to the acquisition phase. I am very sympathetic in general to accounts which allow for significant changes in an individual’s language during adolescence at least and perhaps right through adulthood. But the scenarios demand enormous ingenuity and are necessarily highly speculative.

Although there are difficulties with all these models of reanalysis, I certainly don’t wish to reject reanalysis outright. But here I want to follow a different tack. Reanalysis of an individual form moves instantaneously from one watertight analysis to another (even if the consequences take time to play out and/or spread). But synchronic grammar just isn’t watertight. The prior assumption is wrong.
4. *Synchrony leaks*

Edward Sapir’s (1921:38) dictum that “All grammars leak” has long been practically a motto of mine, though I (and probably others too) have unintentionally misappropriated it: Sapir was specifically writing about the mapping between form and meaning. It isn’t entirely clear that he is endorsing the interpretation which I wish to attribute to the saying, namely that the categories and structures imposed by grammarians on a language are never watertight. Nevertheless, that is the sense in which I intend my title to be read.

Here I will concentrate on simple NPs. The basic structure used by the *Cambridge Grammar* (2002:331) is

\[(14) \quad \text{NP} \]

\[
\begin{array}{c}
\text{Determiner:} \\
D \\
\text{Modifier} \\
A \\
\text{Head:} \\
\text{Nom} \\
\text{Head} \\
N \\
\end{array}
\]

\[
\begin{array}{c}
\text{the} \\
\text{old} \\
\text{man} \\
\end{array}
\]

They give reasons for not using the Abney-style DP analysis. And note that they use ‘Determiner’ for the *function*, as in (14), whereas I will stick to old habits and use it for the *category*. Something like the three-way distinction of D, A, N is very widely accepted cross-linguistically and in English grammar, in the majority of frameworks.

What does it buy us?

Determiners are function-words, of general semantics (definiteness, quantification, etc.), and mutually exclusive. Adjectives are lexical, content words, and
can be iterated. ‘Lexical’ is an imprecise term which tends to involve specificity of meaning and openness of class. A summary of some key properties is given in Table 1.

I have added Prn, since pronouns can act as head of NP, like N, but differ in many other ways, including general incompatibility with D and A. Almost every cell needs a bit of qualification, as we tell our students:

- Prn can co-occur with A in certain circumstances: *lucky me, poor you*
- N can iterate, though not as head of NP: *pit-head strike ballot*
- not all A show comparison: *potential, mere*
- not all A can occur in predicative position: *potential, mere*
- some A always occur in predicative position and do not function attributively in NP at all: *asleep*

So far these subtleties do not vitiate the basic D-A-N analysis of the standard NP. I should emphasise that I think that D, A and N are useful descriptive generalisations, probably with genuine psychological reality. However, in categorial terms it might be interesting to arrange the four categories mentioned so far in the following diagram:

```
(15)
   _________
  |        |
  |  D     |
  |_______|
   |        |
  |  A     |
  |_______|
   |        |
  |  N     |
  |_______|
   |        |
  |  Prn   |
  |_______|
```

The reason is that every boundary between adjacent categories is leaky in PDE, whereas Prn and A are entirely distinct and so too (I think) D and N. Let me briefly exemplify the shading of one category into another. For the A-N boundary I refer to Denison (2001),
for example the word *key*, found in PDE at various points along the continuum from clear N to almost wholly adjectival uses like:

(16) There are several reasons. The *most key one* for many victims is …

(2001 *Women’s Hour*, BBC Radio 4 (15 Nov.))

(17) There are *key fundamental flaws* in this method.

(2002 Damian Byrne, abstract of article in online *Journal of Language and Linguistics* 1.1)

For Prn~N, we have the analytical possibility of treating Prn as a subcategory of N, since they are the two categories which can normally be head of NP. Historically there are shifts in the usage of OE *mann* to *man*, and arguably of *self, body*, and so on from the domain of N to that of Prn.

The relation between D and Prn is interesting. On one approach, the core of the D category is the articles, since almost their sole function is to determine a noun, and they cannot occur in any pronominal (that is, pro-NP) function. The remainder of the central determiners are then defined, at least in part, by contrastive distribution: any NP item which cannot co-occur with the articles is itself a determiner too. However, as is well known, apart from the articles, almost every determiner *can* occur in pronominal function.

(18) What do you think of *this*?

(19) *Some* are born great.

There are good reasons to analyse such items in two different ways: as an NP whose head (and sole constituent) is Prn, bringing out the analogy with

(20) What do you think of *him*?
(21) They are born great.

Or as headless NPs with a missing or ellipted N, bringing out the analogy with

(22) What do you think of this idea?

(23) Some people are born great.

If we treat the possessive determiner my and the pronoun mine as variants of the same form, then there is even more overlap between D and Prn. (Historically, of course, that is exactly what they were.) Apparently rather similar is no and none.

The status of pronominal determiners is tricky, therefore, reflected in a rather confusing analysis in Quirk et al. (1985:870-873). Huddleston and Pullum — normally so decisive — actually treat pronominal determiners as a kind of blend, what they call the fused-head construction. I reproduce their diagram of an explicitly partitive fused-head (Huddleston and Pullum 2002:412):

(24)

```
NP
  Head: Nom
    Det-Head: D
    Comp: PP
      few
      of her friends
```

This is meant to capture the idea that certain kinds of pronominal determiner (though not personal pronouns, as in (25)-(26)), are simultaneously the grammatical part of the NP, acting as modifier to the lexical part, and the head.
Coming at the gradient from the other side, note that even the core of Prn, the personal pronouns, have some determiner-like uses, though only the plural ones (and not 3 PL in standard):

(25) *We/#us scholars* must stick together.

(26) *You people* are all wrong.

(27) *You idiot(s)!*

(28) *#Them ideas* are really stupid.

Only the vocative (27) allows singular as well as plural. The *Cambridge Grammar* regards these, or at least (25)-(27), as determinatives in category, pointing out that the predeterminer *all* can precede the plural items as normal (Huddleston and Pullum 2002:374, 353 n.14). Dick Hudson has an analysis in which Det and Prn are unified as a single category on the same lines as analyses which treat adverbial particles as intransitive prepositions (Hudson 2000a); and of course there is the DP Hypothesis (Abney 1987).

As for the last pairing, D~Adj, here we have a veritable minefield. Numerals have long been difficult to classify, but there are subtler difficulties too. Several of us have pointed out the difficulties of classifying *such* both as pre-determiner and post-determiner (Denison 1998:117-18, Mackenzie 1997),

(29) such a nuisance

(30) no such luck

and the recent change from

(31) *such another → another such*
Now there are at least two analyses (Huddleston and Pullum 2002:435, Spinillo 2003) which make a strong case for analysing *such* not as any kind of determiner in PDE but as A.

5. *Diachrony leaks*

Let us extend this discussion of the Det~A boundary into diachrony. The *Cambridge Grammar* (Huddleston and Pullum 2002:392-3) regards *various* and *certain* as marginal determiners in PDE because of their occurrence in a partitive ‘fused-head’ construction (*certain of the delegates*), general semantics, non-generic semantics, and — for *certain* only — use with *a* in

(32) This gave her *a certain authority*.

It is interesting to compare this predominantly syntactic account with *OED*’s comments on the lexical history:

I. 1. a. Determined, fixed, settled; not variable or fluctuating; unfailing. To avoid ambiguity from confusion with sense 7, the adj. is sometimes put after its n., *as a certain day, a day certain*.

[...]

II. 7. a. Used to define things which the mind definitely individualizes or particularizes from the general mass, but which may be left without further indentification [sic] in description; thus often used to indicate that the speaker
does not choose further to identify or specify them: in *sing.* = a particular, in *pl.* = some particular, some definite.

Different as this seems to be from sense 1, it is hardly separable from it in a large number of examples: thus, in the first which follows, the *hour* was quite ‘certain’ or ‘fixed’, but it is not communicated to the reader; to him it remains, so far as his knowledge is concerned, quite indefinite; it may have been, *as far as he knows*, at any hour; though, *as a fact*, it was at a particular hour. (The absolute uses are in B 4-6.)

a1300 *Cursor M.* 8933 Ilk dai a certain hore| þar lighted dun of heuen ture Angels. [etc.]

In other words, there are two very different lexical senses of *certain* which can be argued to be categorially different as well, yet there are many early examples which are equivocal. Given that the gradual — or at least *graduated* — nature of semantic change is reasonably widely accepted, why not allow that syntactic change may proceed by small steps too? I have alluded to many synchronic examples of gradience within the NP, where a word or a construction is neither one thing nor the other. Many will actually have started off as ‘one thing’ and moved during the history of English towards ‘the other’. Note that this is not necessarily an argument for *slowness* of change, merely for graduatedness of at least some changes.


6. How intermediate forms arise

How and why should this happen? One mechanism would be via what Quirk called serial relationship (1965, 1968). This is a plausible idea, though not part of a fully worked out model of language, in which new intermediate forms are made possible precisely because they share characteristics with, and therefore largely overlap with, already-existing grammatical forms on either side. One of his matrices (1968:Table 5, p.172) is reproduced here as Table 2.

The columns 1-7 represent variant sequences in which the particular verbs shown at the left either are (+), or are not (−), attested, for example:

4. He would \( X \) to come every day

5. He \( X \) that

6. He \( X \) us to come every day

Quirk writes (1968:172): “it is because of the gradience, in fact, that we [...] may find in speech ‘mistaken’ constructions of the form \( He \ doesn’t \ want \ that \ anyone \ should \ ... \ ”. Later he explains

(33) He was said to be foolish

in a similar way, as a passive that is fully acceptable despite the non-existence of any corresponding active (1968:176).

Another mechanism with some promise of forming part of a coherent linguistic theory would be dual inheritance, in a Construction Grammar framework where the
properties of a given form may be inherited from more than one more general source construction (Goldberg 1995, Hudson 2000b), but *contra* (Bresnan 1997).

Either way, intermediate forms would arise when they serve some communicative purpose and because speakers — as opposed to linguists — are not confined to producing structures and grammars which are wholly self-consistent and maximally elegant and economical. If sufficiently useful, the innovations would diffuse into the language and might lead to permanent changes in the categories or structures of its grammar. But they might well not. Genuinely intermediate forms are often unstable historically.

7. Why intermediate forms disappear

If we assume that humans categorise the world — things in general, that is, and therefore also linguistic ‘things’ — roughly according to the principles identified by Rosch (1978, 1988), then one consequence is that we unconsciously adjust our categories and our categorisations to accommodate in the most satisfying way what we encounter. This was the assumption underlying Anthony Warner’s (1990) account of the development of a modal category in English. Perhaps it is not too much of a stretch to assume that one solution to categorial awkwardness is for language users to find ways of avoiding the troublesome forms. (Again, this was applied by Warner to the loss of old preterite-presents like *witan* and certain lexical meanings of *shall, can*, etc. which fitted least comfortably into the emerging modal category.) The psycholinguistic experiment with pictures of cups and bowls is well known (Labov 1973). There the language user
typically does make a choice, if sometimes an arbitrary one, as to where to draw the dividing line.

We have, then, the outline of a mechanism which disfavours intermediate forms. I have no explanation for why some intermediate forms seem nevertheless to show long-term stability; the best example I know is the P~A form near, though most other former P~A words have moved to one side or other of the boundary.\(^5\)

With two opposing principles in tension we thus allow for both stability and change; it is worth noting, though, that such a diachronic model is rather at risk of post-hoc-ery.

8. Concluding remarks

When a lexical item develops a new sense, there are often equivocal ‘bridge’ examples which users do not need to allocate to older or newer sense, since either works fine in the context. It might be compared to a superposition of quantum states whose wave function only collapses when observed. Linguists, however, tend to insist on observing, whereas ordinary speakers and hearers sometimes don’t need to look analytically but only need to have the right chunk of language with the right overall interpretation. And I’m suggesting that this can happen not just in lexical semantics but in morphosyntax too. It might even happen with the example I used earlier to illustrate reanalysis:

(34) The car ran over a hedgehog.
Perhaps neither speaker nor hearer will need to decide whether this is the prepositional or
the phrasal verb. (And if that doesn’t work, keep substituting different animals — *frog*,
*rabbit*, *sheep*, *elk* — until you find the right size for this to be true.)

Much remains to be done at the diachronic level, including the tracing and
verification of specific examples, and following enough of them to allow safer
generalisations about the relative importance in language change of intermediate forms,
and about their markedness and (in)stability. I don’t think this is merely a notational
variant of reanalysis.

What would such an enterprise buy us? It seems to me that the case for leakiness
of grammar in synchrony is difficult to deny. The history of a language is the passage
through successive synchronic states — or conversely, the simultaneous existence at any
given moment of many historical states in a layered manner. (Notice that this latter re-
formulation sounds very much like a description of a grammaticalisation gradient, but I
intend it to apply to a wider range of changes than grammaticalisations or lexicalisations.)
Either way, a picture of change which both reflects and makes use of a more plausible
picture of grammar, is in my opinion a desirable aim.
Note

1 I am grateful to Bill Croft and to the audience at 12ICEHL for comments, and to my
daughter Rosie’s timely contribution on (screen-)saving face.

2 This section is adapted from draft material prepared for Hogg and Denison (in prep.).

3 Later he discusses the arbitrariness of parts of speech and concludes “how they not
merely grade into each other but are to an astonishing degree actually convertible into
each other” (Sapir 1921:118). Once again, though, this is really to do with cross-
linguistic comparison rather than assignment of an individual word in one language to a
word class.

4 I have shortened pronoun as Prn to distinguish explicit pronouns from the empty
category Pro of GB Theory.

5 See now Aarts, Denison, Keizer and Popova (in press, 2004:Introduction) for a review
of work on mixed and gradient categories.
References


**Tables**

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>A</th>
<th>N</th>
<th>Prn</th>
</tr>
</thead>
<tbody>
<tr>
<td>lexical</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>iterate</td>
<td>–</td>
<td>+</td>
<td>(–)</td>
<td>–</td>
</tr>
<tr>
<td>number marking</td>
<td>(–)</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>comparison</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>case marking</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>can act as predicate</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>?</td>
</tr>
</tbody>
</table>

*Table 1: Distinctive features of some NP categories*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>intends</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>wants</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>seems</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>–</td>
</tr>
<tr>
<td>has</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>used</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>is</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>may</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*Table 2: An example of serial relationship*