J. Linguistics **35** (1999), 273–296. Printed in the United Kingdom
© 1999 Cambridge University Press

Words and feet[1]

ERIK FUDGE

University of Reading

(Received 4 August 1997; revised 2 June 1998)

This article challenges the commonly-held assumption that words, feet and syllables fit into a single phonological hierarchy, and in particular that feet are always sub-units of words. A number of facts of English rhythm are cited which cannot be accounted for unless words and feet are taken to be units of different hierarchies: many of the differences of rhythm between phrases of otherwise similar phonological structure which are noted by Abercrombie (1964) are accounted for most naturally by differences in placement between, on the one hand, word- and stress-group-boundaries, and, on the other, foot-boundaries.

1. Introduction

Many of the theories which aim to deal with phonological elements larger than the segment make an assumption that gives rise to a number of problems: that is the assumption that feet occupy a place in the hierarchy between words and syllables (see, for example, Selkirk 1980; Nespor & Vogel 1986; Harris 1995: 872; Kager 1995: 369; McCarthy & Prince 1995: 320). This article deals with two specific problems arising from this assumption. On the one hand, it is impossible to maintain the hierarchy consistently (sections 3–4); on the other hand, it fails to yield a convincing account of certain facts of rhythm in English (sections 4–6).

Both kinds of difficulty are shown to be removed by the postulation of two distinct hierarchies of elements, as proposed in the model of section 2. In section 2.2, special attention is paid to the nature of feet. Section 3 deals with the relationship between words and feet, and section 4 with that between

[1] Some of the ideas of this article have emerged from the lectures I have given on this and related topics at the University of Reading. An earlier version of the article was written during a sabbatical term spent in the University of Groningen, Netherlands. I should like to express my thanks to students and staff at both institutions, but especially to Dicky Gilbers and Klarien van der Linde, who were my hosts at Groningen and helped to make my stay there so happy and profitable. I also thank Heather my wife, for agreeing that I should work away from home during this time. The present version of the article owes a good deal to *JL*'s anonymous referees.

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

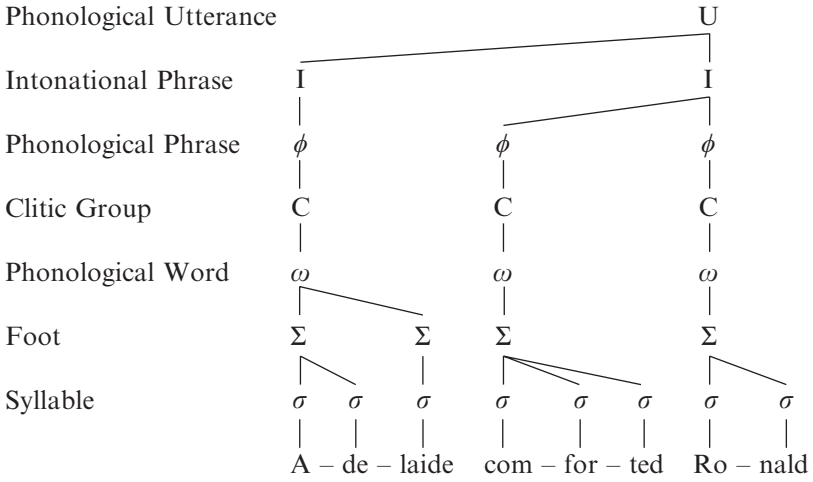
clitic groups and feet. Section 5 sets the scene for the treatment of the examples in section 6.

2. The model used in this study

2.1 *How many 'levels' or 'strata' are required?*

Nespor & Vogel (1986) (henceforth N&V) propose a model of the phonological hierarchy (1) in which there are seven levels, each of which has to be treated as separate and distinct from the others. They make the rather odd claim (pp. 11–12) that the theoretical description of every language is required to include all seven of these units, even if there are some of the seven which it does not appear to use actively. This suggests that a model with fewer distinct levels might be more appropriate.

(1) *The seven phonological units of Nespor & Vogel's prosodic hierarchy, with an example*



2.1.1 *Reduction in the number of levels*

The criteria for setting up of different strata in a hierarchy are well characterised by Pike: 'Crucial to [this] is the analytical requirement that unit types on one level must in some sense (1) control the occurrence and relative (fixed or free) order of included constituents, and (2) be structurally organised in a manner which in some sense is sharply in contrast with the layer next higher or lower in the hierarchy' (1967: 437).

The fact that most scholars support (even if not explicitly) the requirement

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

of Pike's condition (2) is, I believe, one reason why other approaches to the prosodic hierarchy have generally assumed a smaller number of levels than N&V's seven.

For some of these seven levels, it appears that level L is much more like its neighbour level A than its other neighbour B – if this is so, then Pike's criterion (2) suggests that L belongs on a separate stratum from B, but the same stratum as A. Typically, the definitions appeal to related types of phenomena, they intervene in comparable phonological processes, etc. For instance N&V's *I* is much more like their *U* than it is like ϕ or any other of their smaller units; on this basis we claim that N&V's *U* and *I* can be regarded as (different-sized) units on the same stratum (section 2.1.2). Similarly Clitic Group *C* is more like Phonological Word ω than it is like Phonological Phrase ϕ, and ω is more like *C* than it is like foot Σ, suggesting that *C* and ω are also different-sized units on the same stratum (section 2.1.3). This suggests grouping the levels together in a smaller number of strata (possibly three or four rather than N&V's seven: see section 2.3 below). There is then a much greater likelihood that all of these strata will have some part to play in all languages.

Between the 'levels' within one stratum a relation of SUPERUNIT to UNIT is argued for: *U* is the superunit on the same stratum as *I*, *C* is the superunit of ω. It would not necessarily be the case for any particular language that each stratum would require both its unit and its superunit. On each stratum, the presence vs. absence of superunit in addition to the unit would constitute a relevant parameter; this would avoid making N&V's claim that having seven levels was a phonological universal.

2.1.2 *Affinities of* I *and* U *in the work of Nespor & Vogel*

Perhaps the most obvious case of units forming a closely related pair is N&V's pair Intonational Phrase (*I*) (1986: 187–220) and Phonological Utterance (*U*) (1986: 221–247). I suggest that on the following grounds they may be grouped together on one stratum.

(a) The criteria invoked in N&V's definitions of the two units are rather similar to one another in their appeal to syntactic structure (cf. (2) and (3)); the other units do not appeal to syntax in the same way.

(2) *Intonational Phrase Formation*
'I. I *domain*
An *I* domain may consist of
a. all the ϕs in a string that is not structurally attached to the sentence tree at the level of s-structure, or
b. any remaining sequence of adjacent ϕs in a root sentence.' (p. 189)

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

(3) *Phonological Utterance Formation*
'I. U *domain*
The domain of *U* consists of all the *I*s corresponding to X^n in the syntactic tree.' (p. 222)

(b) Their reliance on logico-semantic factors in the determination of which of their constituents is the strong one.

For the Intonational Phrase: 'Rather than being determined structurally, the strong node under *I* is determined on the basis of semantic factors such as focus and given *vs.* new information.' (p. 190)

For the Phonological Utterance: '*U*-level restructuring…depends not only on phonological and syntactic factors, but also on factors of a logico-semantic nature' (p. 221). '… *U* consists of those *I*s that are dominated by the same node in the syntactic tree' (p. 222). 'Except in sentences with special emphasis patterns…, it is generally not possible to say that one *I* is "more prominent" than another' (p. 222) though, on grounds of 'sentence-final intonation', and lengthening of elements (p. 223), N&V conclude that 'the rightmost node dominated by *U* is the strong one' (this assumes coincidence between *U* and the sentence, a view which N&V in fact reject later (pp. 234–244)).

Determination of strong nodes within *I* and *U* is thus carried out (if necessary) on the basis of logico-semantic information. Furthermore, there is a difference of kind between *I* and lower elements in the hierarchy, in which normally the position of the strong node can be determined structurally.

(c) Consideration of the references to 'pause' in N&V's subject index (p. 320) makes it clear that the relevance of pause phenomena is great for *I* and *U*, but has virtually no part to play in the definition of lower units.

(d) 'Restructuring' of a constituent is possible for both *I* (pp. 193–205) and *U* (pp. 237–244). Some (though not all) of the factors causing restructuring are shared by both, e.g.

(i) Length of a constituent (pp. 194, 240);
(ii) Speed of utterance (pp. 195, 240);
(iii) Style of speech (pp. 195, 244).

On the other hand, restructuring of *I* takes the form of a dividing-up of a long *I* into several shorter ones, whereas restructuring of *U* takes the form of building fairly short *U*s into one longer one.

On these grounds I would group these two units together on a stratum to which the name Tone Group could be given. If Intonational Phrase constitutes the basic 'ToneGroup' unit, then Phonological Utterance will be a 'SuperToneGroup' on the same stratum. The two units then have much in common with the 'Minor' and 'Major Tone Groups' of Trim (1959).

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

2.1.3 *Affinities of* ω *and* C *in the work of Nespor & Vogel*

The following seem good grounds for grouping N&V's Phonological Word *ω* and Clitic Group *C* on one stratum:

(a) A Word plus an Affix makes a word in most languages: a Word plus a Word makes a word in some languages (N&V, pp. 112–113). If the Clitic is a unit which is sometimes almost affix-like and sometimes almost word-like (pp. 146–149), then Word plus Clitic might be expected to form a word-like unit in at least some languages. It is therefore not surprising that N&V define the Clitic Group as consisting of a Word plus Clitics (or, in certain cases, as a Clitic 'such that there is no possible host with which it shares more category memberships' (p. 154)).

(b) In some languages word-stress reapplies to the Clitic Group, for example Latin (pp. 146, 160–161) and Modern Greek (pp. 146–147, 151–155); this leads us to infer that the Clitic Group is a Word-like unit, at least in some languages.

2.2 *What is a foot?*

2.2.1 *Has 'foot' been used in several radically different ways?*

The foot might well be defined by different criteria, and with different aims in view. In particular, it is quite clear that the term is not used identically in the fields of speech rhythm and metre in verse. In verse prosody the strong element of a foot may precede the weak (as with trochees and dactyls) or follow it (as with iambs). On the other hand, in studying the speech rhythm of English and similar languages, the unit called the foot must begin with its strong element. In this respect the speech-rhythmic foot is more like the bar in music, where the first beat is always strong, than it is like the verse foot.

Again, what of the claim made by Selkirk (1984: 38–42) that there are really two types of 'speech rhythm' foot which are quite different from one another, and must not be confused with each other?

Selkirk introduces the foot as follows: a foot consists of a stressed syllable together with zero, one, or more stressless syllables, so that 'The description of the distribution of stressed and unstressed syllables in words is no longer a matter of rules assigning the feature [±stress], but of rules that indicate (among other things) what constitutes a well-formed foot in the language, often in terms of the nature of the component syllables. On this view, the foot is a unit of phonotactic description, much like the syllable' (Selkirk 1984: 15).

At this point in the text, Selkirk appends footnote 22, which reads: 'This use of the notion "foot" is thus not the same as Abercrombie's. Abercrombie 1964 defines feet in terms of stress; stress patterns therefore have some prior, independent existence. In the theory of Selkirk …, Halle and Vergnaud …,

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

and Hayes…, for example, "stress" has no existence independent of feet. Patterns of stress are simply patterns of feet, the strongest syllable of the foot being the one that might be called the "stressed syllable"' (1984: 419).

But surely this case is quite different from the comparison drawn above between 'speech feet' and 'verse feet'. In this instance it is clear that, however different their starting points, Abercrombie and the metrical phonologists are actually using the concepts chosen (termed 'foot' in both cases as it happens) to account for the same set of phenomena, and eventually to develop the same kind of unit.

This situation is no different from that of the syllable – one scholar can, like Selkirk, begin to define the syllable from a consideration of phonotactics, another scholar from a consideration of suprasegmentals, and a third from a consideration of articulatory performance. The essence of the syllable, though, is that it lies at the intersection of at least these three types of criteria, and the fact that some cases arise in which the three fail to converge totally does not vitiate the whole enterprise, or imply that the three types of syllable are fundamentally distinct from each other.

2.2.2 *The relation of the foot to other elements*

To place the foot between the syllable and the phonological word (whether as a superunit of the syllable or as a unit with phonological word as its superunit) appears to lead to undesirable complications. While it is uncontroversial to say that the foot consists of one or more syllables, the trouble arises from the assumption that the word consists of one or more feet. To make this assumption entails ignoring how certain words relate to feet (section 3).

We therefore argue that there are two hierarchies to be reckoned with. In the one, the phonological word is directly superordinate to the syllable; in the other the foot is directly superordinate to the syllable. The relation between the word and the foot is thus no longer a hierarchical matter.

The same is true of the stress-group and the foot. As shown in section 4, to subordinate the foot to the stress-group is also to lose the opportunity of accounting for certain rhythmic differences in English.

2.3 *The model used*

What we are in fact arguing for is that there are two distinct hierarchies linking tone-groups and syllables, as shown in Figure 1. These two hierarchies relate to different tasks: among other things, the first (which I shall refer to as HIERARCHY I in what follows) assigns accents (potential stresses) to certain parts of an utterance, while the second (HIERARCHY II) determines which potential stresses are made into actual stresses when it comes to realising the utterance in speech. This is consistent with the distinction drawn by Bolinger

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

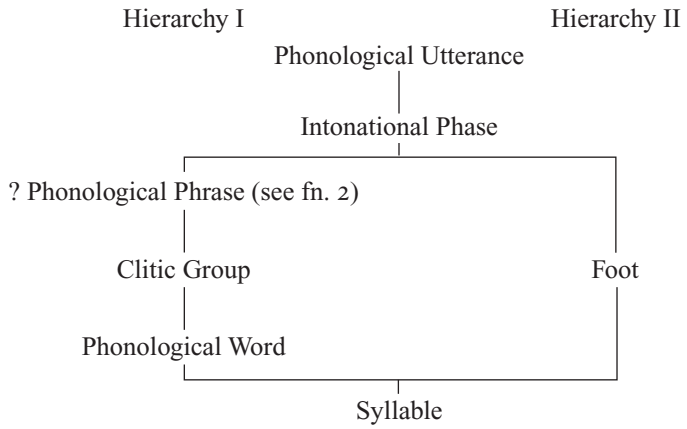


Figure 1
Two distinct hierarchies linking Phonological Utterance and Syllable

(1981: 26–28; cited by van der Hulst 1984: 177–178) between 'syllabic rhythm' (our Hierarchy II) and 'accentual rhythm' (our Hierarchy I).

Hierarchy I has no direct rhythmic implications. The words of an utterance are put together in a string whose linear order reflects lexical and syntactic constraints. The number of syllables and placement of accents within lexical items are both randomly variable over the possibilities permitted by the language; we cannot therefore guarantee that every accent will get a stress, or that the 'Perfect Grid' rule (Goldsmith 1990: 194) will apply straightforwardly to every string of words it encounters.

The interplay of these two hierarchies appears likely to be important in a diachronic context. Might this be what Luick was referring to in his claim (1898: 357) that the loss of certain syllables in English was triggered off by '[das] schwanken zwischen logischer gliederung [Hierarchy I] und mechanischer rhythmisierung [Hierarchy II] des wortmaterials'? On syntactic and semantic ('logisch') grounds, the phrase from the Lord's Prayer *yworþe þi wille* 'thy will be done' would be split into N&V's clitic groups as in (4a), while on rhythmic grounds it would be split into feet as in (4b).

(4) (a) yworþe | þi wille
be-done thy will
(b) (y-) | -worþe þi | wille

The prosodic model that will be assumed in this paper is therefore as follows:[2]

[2] The place of Nespor & Vogel's Phonological Phrase ϕ in Hierarchy I is problematic: there are factors which might indicate that it is a 'SuperSuperWord' and there are others which

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

Hierarchy I has three strata:

(i) TONE-GROUPS: SUPERTONEGROUP (roughly = N&V's *U*, and Trim's major tone-groups)
TONEGROUP (roughly = N&V's *I*, and Trim's minor tone-groups)

(ii) WORDS: SUPERWORD (roughly = N&V's *C*, and Chomsky & Halle's 'phonological word')
WORD (roughly = N&V's ω)

(iii) SYLLABLES

Hierarchy II also has three strata:

(i) TONE-GROUPS: As for Hierarchy I
(ii) FEET
(iii) SYLLABLES: As for Hierarchy I

Sections 5 and 6 show how this hypothesised framework operates for English.

2.4 *Is the foot a phonological universal?*

It is generally agreed that the foot must have a particularly important part to play in accounting for what is normally termed 'isochronous stress'. In languages which exhibit this feature, feet, rather than syllables, or other units, are approximately equal in duration. To put this in a more precise manner which avoids many of the objections raised against 'isochronous stress': feet begin as suprasegmental timing slots which are equal in duration, but during the derivational process may be lengthened or shortened slightly by such factors as number of syllables, and the amount of segmental material contained in each.

It is clear that, in any language which needs the foot, a Phonological Utterance (Major Tone-group – see section 2.1.2) must consist of an integral number of feet; there is, of course, the possibility that the first foot of the utterance is stressless, much as the first bar of a piece of music may lack an accented beat. Intonational Phrases (Minor Tone-groups) also probably consist of an integral number of feet, with the same proviso.

If the raison d'être of the foot is to account for rhythm in languages where the isochronous stress principle operates, then is the foot needed at all for other languages? It could well be argued that syllable-timed languages like

might indicate that it could be regarded as a 'SubTonegroup'. I shall maintain an open mind on whether it is a necessary unit for prosodic purposes. For the range of data considered in section 6, it appears to play no distinct role, and I shall omit further reference to it in this article.

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

French need no such unit; I would certainly want to argue that to equate the foot with any roughly similar unit used for other purposes in such languages is arbitrary. If that is so, then to have the foot in a hierarchy distinct from that of the other units is an advantage: French would on the one hand have all the elements of Hierarchy I, and on the other call for a decision about Hierarchy II – we would have to conclude either (a) that Hierarchy II has no place in the description of French, or (b) that Hierarchy II does play a part in French, but that the rhythmically relevant unit coincides with the syllable rather than being a special kind of grouping of syllables. Alternative (a) would have the consequence that French would have no rhythmically relevant units, which appears unlikely: all languages have a rhythmic principle, and hence need a rhythmically relevant unit. This would suggest that (b) is the correct solution.

This leads to the following tentative conclusions:

(a) All languages have all the units of Hierarchy I.
(b) All languages have some rhythmically relevant element(s) on Hierarchy II:
  (i) the syllable itself;
  (ii) the foot (stressed syllable plus all following syllables up to but excluding the next stressed syllable); or possibly
  (iii) some mixed solution involving the interplay of (i) and (ii).
(c) All languages which have the foot will show rhythmic differences dependent on the way feet and stress groups relate to each other (see section 4).

3. The difficulties of relating feet to phonological words

These difficulties arise from N&V's insistence that '… as required by the Strict Layer Hypothesis …, all the feet of a given string must be grouped into phonological words, and no other category may be so grouped. Each foot is thus exhaustively included in a ω; that is, it is never the case that the syllables of a single foot belong to different phonological words' (1986: 109). In this section we shall try to show that, for English at least, this cannot be the case; the notation used will be a simplified form of N&V's, in order to show the problems arising from having a single hierarchy.

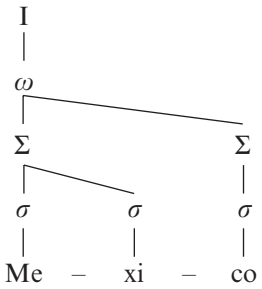
3.1 *Variability in relation between feet and words*

A complication in stating the relation between feet and phonological words is that many words may be 'footed' in several different ways depending on such factors as emphasis or speed of utterance. It would therefore be helpful if individual words were 'footed' after being put into longer stretches, rather than bringing all their own footing with them from the lexicon.

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

For example, as a word in isolation, *Mexico* may well be said to consist of two feet, as is shown in (5). We must, however, consider what exactly it means to talk of 'a word in isolation'.

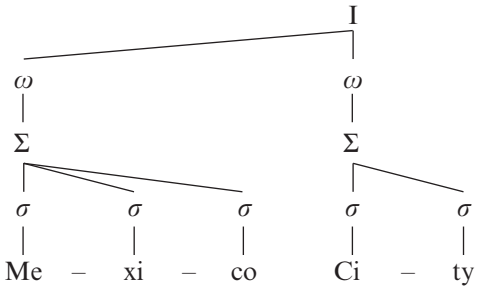
(5) *The word* Mexico *spoken in isolation*



To pronounce a word 'in isolation' means we are making a whole Phonological Utterance out of it, giving it an intonation nucleus. The number of stresses, hence the number of feet, may vary: the more careful the speech (and pronunciation of a word in isolation is normally a matter of rather careful speech), the more stresses are likely to be actualised, and hence the more feet are likely to be contained in the phonological utterance. Just as we find where the main stress (potentially prominent syllable) of an English word is by making it into a phonological utterance, we also by so doing find what the potential foot structure of the word might be.

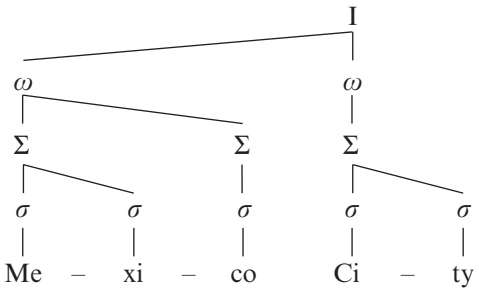
When the word is part of a longer phrase, not all stresses are always actualised: in the phrase *Mexico City*, the word *Mexico* is much more likely to have one foot (6) than two (7). In this phrase it normally takes no longer to say *Mexico* than it takes to say *City*; to utter (7) is to leave what is perceived as too long a pause between the two words.

(6) Mexico City *with* Mexico *as one foot*



https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

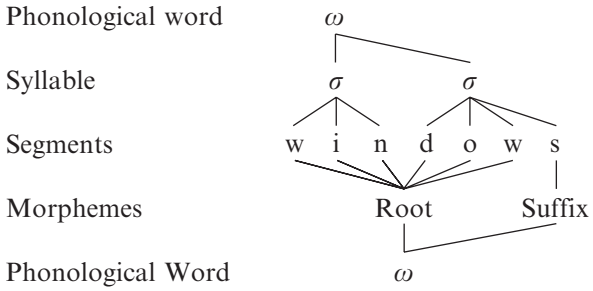
(7) Mexico City *with* Mexico *as two feet*



3.2 '*Perfect Grid*' *is not always possible*

The attempt to relate phonological words to feet in fact exhibits difficulties comparable to those arising from the attempt to relate syllables and morphemes. Words consist of syllables, and words consist of morphemes (see (8)), but we cannot generalise about morphemes consisting of syllables (*door* consists of one syllable, *window* of two, etc., but suffixes like plural /z/ and past tense /d/ are normally less than one syllable) – nor can we generalise about syllables consisting of morphemes (/dɔː/ consists of one morpheme, /gəʊz/ consists of two, but the /wɪn/of *window* is less than a morpheme).

(8) *Lack of correspondence between syllables and morphemes*



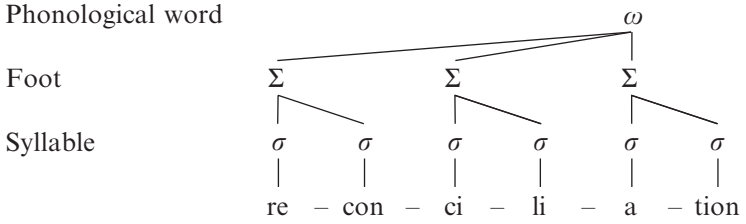
Likewise the words given in much of the literature as examples of foot structure are for the most part 'convenient' examples which conform to the pattern imposed by 'Perfect Grid' – one of alternating strongs and weaks, or stressed and unstressed: *reconciliation* is typical (see (9)). The more 'inconvenient' examples exhibit immediately adjacent strongly-accented syllables (e.g. *advantageous*, see (10)), or too many immediately adjacent unaccented syllables (e.g. *generatively*, see (11)).

Although it is possible for the long string of unaccented syllables in this last example to be 'rationalised' by having a further accent placed on the

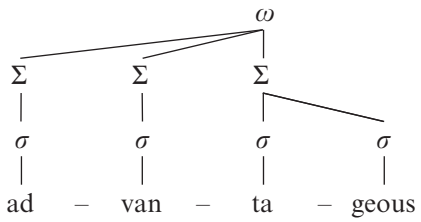
https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

syllable *-tive* (giving two feet as in (12)), it is equally possible for the string to be maintained as in (11).

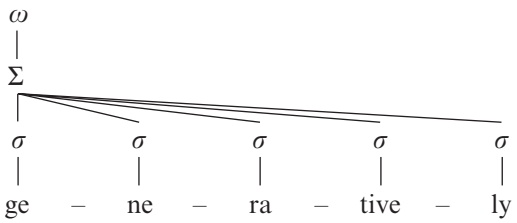
(9) *Prosodic structure of reconcilation*



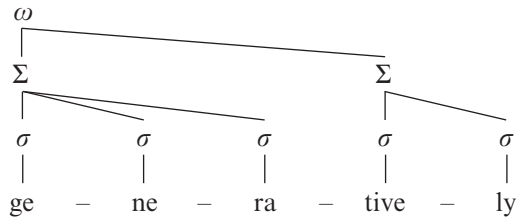
(10) *Adjacent accented syllables*



(11) *A large number of unaccented syllables together*



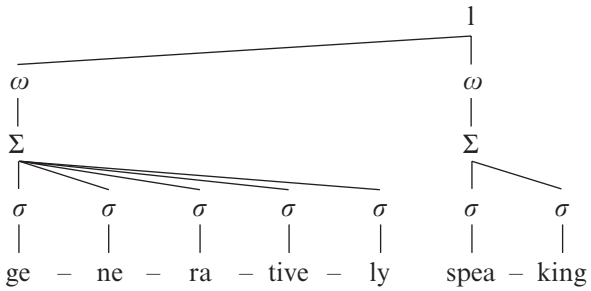
(12) *'Perfect Grid' restored*



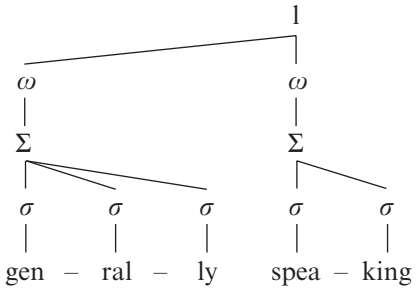
The phrase *generatively speaking* may easily be said with two feet (13), taking no longer to say than the almost 'Perfect Grid' *generally speaking* (14). Perfect Grid thus by no means imposes itself on every utterance.

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

(13) *A non-Perfect Grid*



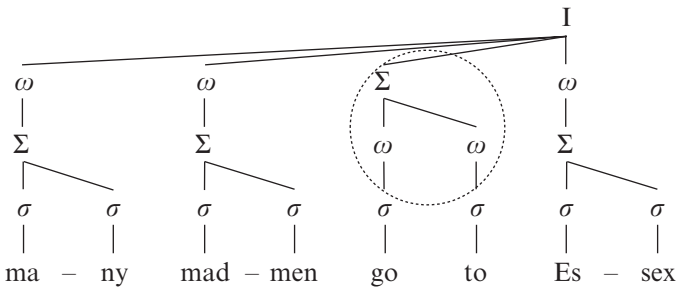
(14) *A near-Perfect Grid*



3.3 *Feet may cross word boundaries*

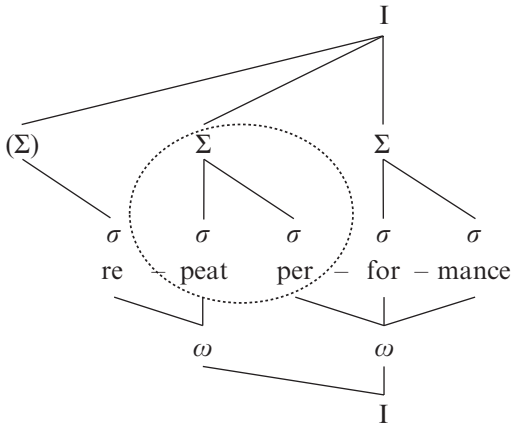
Sometimes a foot clearly consists of two words: for instance the two words *go* and *to* form a single foot in the well-known example *Many madmen go to Essex* (15). Again, when the first of a pair of adjacent words ends in a stressed syllable and the second begins with an unstressed one, a foot may contain a piece of each word, as in the phrase *repeat performance* (see (16)). For a more detailed discussion of these and related problems, see Giegerich (1992: 258–272).

(15) *Feet may consist of more than one word*



https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

(16) *Foot structure may cut across word boundaries*



Morphemes often resyllabify when they join together to form words (*dri.ving*), but we do not try to use this fact as grounds for pushing morpheme and syllable into the same hierarchy. In the same way, words often 're-foot' when they join together to form phrases; once again, we should not use this fact to try to force word and foot into the same hierarchy. Giegerich's attempt to do this (1992: 260–268) fails in that it either leaves such syllables relating directly to words and hence not part of feet (pp. 262–264), or produces nonsensical words (pp. 264–267) (like *-peat per-* in our example (16)).

3.4 *Nespor & Vogel's definitions do not relate word and foot*

If we consider (17), the definition of phonological word in N&V (1986: 141), we should note that the foot is nowhere referred to.

(17) *Phonological Word Construction*
ω *domain*

A. The domain of ω is Q (the terminal node of the syntactic tree); or

B.I The domain of ω consists of
   a. a stem,
   b. any element identified by specific phonological and/or morphological criteria,
   c. any element marked with the diacritic [+W] (for example, phonologically independent suffixes in Dutch).

This omission makes it difficult to see that a word can coherently be said to 'consist of feet' in the same basic sense that, for N&V, a clitic group

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

consists of phonological words (cf. Definition (18) below), or a phonological phrase consists of clitic groups (cf. Definition (19)). 'Specific phonological criteria' (B.Ib in (17) above) might conceivably include foot structure, but hardly refers to it in any genuinely defining way.

(18) '*Clitic Group Formation*
'I. C *domain*
The domain of *C* consists of a ω containing an independent (i.e. nonclitic) word plus any adjacent ωs containing
a. a D(irectional) CL(itic), or
b. a CL(itic) such that there is no possible host with which it shares more category memberships.' (N&V, p. 154)

(19) '*Phonological Phrase Formation*
'I. ϕ *domain*
The domain of ϕ consists of a *C* which contains a lexical head (X) and all *C*s on its nonrecursive side up to the *C* that contains another head outside of the maximal projection of X.' (N&V, p. 168)

This suggests that phonological words are not defined in terms of their constituent feet, and, further, that words do not necessarily have feet as constituents.

4. The difficulties of relating feet to clitic groups

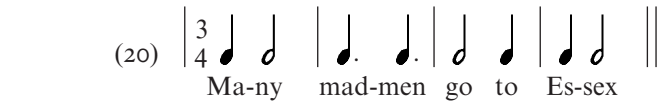
Feet do not relate in a simple fashion to clitic groups either. Like words, clitic groups (and come to that phonological phrases) must contain an accent, but do not necessarily begin with an accent: *to Essex* is a clitic group in example (15). A foot, however, normally has to begin with an accent; if it does not, it cannot contain an accent, and has to begin on an 'up-beat'. In this example, *to* belongs with *Essex* in the clitic group structure, and with *go* in the foot structure. If this is so, why try to fit one type of unit within the other in the same hierarchy?

4.1 *Feet, clitic groups and rhythm in English*

There is, however, important interplay between clitic groups and feet. Differences between clitic group structure and foot structure may be utilised to account for some complex rhythmic differences. It should also be noted that, if feet were invariably constituents of clitic groups (so that clitic group boundaries were invariably also foot boundaries), these rhythmic differences could not be accounted for.

As suggested in section 1 above, feet are rather analogous to bars in music; in fact, musical notation has been found quite helpful by a number of scholars. Following Abercrombie's well-known proposals (1964: 218–219), the rhythm of our example can be approximately represented as in (20).

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

(20) 

The normal rhythm of a two-syllable foot is ♩♩ (shorter followed by longer) if the first syllable of the Foot is light ('short' in Abercrombie's terms) (as in *many*, *Essex*), and ♩♩ (about equal length) if the first syllable is heavy ('long') (as in *madmen*). However, the rhythm of the foot *go to* does not follow this principle, because the *to* belongs with *Essex* in clitic group structure, and therefore is pulled nearer to it in the rhythm: the result is the rhythm ♩♩ (longer followed by shorter) of Abercrombie's 'Type C' foot.

Other examples given by Abercrombie (1964: 219) of Type C feet (shown here as (21)) are also divided by clitic group boundaries internally: foot boundaries are shown by |, and clitic group boundaries (my addition) by [ ].

(21) (a) | [*tea*] [*for*] | [two] |
(b) | [I'll | see you] [a | *gain*] [*to* | morrow]

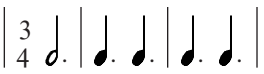
I find it difficult to agree with his inclusion of the further example (22) under this heading: I have ♩♩ rhythm in this instance. This is probably a matter of idiolectal difference between us: Abercrombie gives the pronoun *I* clitic status, as in (22a), whereas I give the pronoun full word status as in (22b).

(22) (a) [per | *haps*] [*I* | did] |
(b) [per | haps] [I] | [did] |

This suggests that the crucial fact leading to a Type C foot is that the syllable(s) undergoing the delay are immediately preceded by a clitic group boundary, but no foot boundary, and are immediately followed by a foot boundary, but no clitic group boundary (as in (22a)). When the following foot boundary has a clitic group boundary coinciding with it (as in (22b)), the delay does not take place.

This approach can be used to explain the difference between the well-known pair of expressions (23) and (24) (Abercrombie 1964: 218).

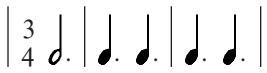
(23) [Take] [Grey] [to London] 

(24) [Take] [Greater] [London] 

Giegerich (1992: 270–272) treats the difference here in terms of the presence of a word-boundary between *Grey* and *to* in (23): a word-final

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

stressed vowel is lengthened before a word-boundary, and the following syllable is then shortened in order to keep feet constant in length. This solution does not work, because the lengthening does not operate consistently before all word boundaries – only before those followed by cliticised words: compare (22a) with (22b), and (23) with (25).

(25) [Take] [Grey] [through] [London] 

Differences between the examples of trisyllabic feet are also accountable for in similar terms. Strict musical notation begins to be insufficient for the task – in particular, I abandon the notion of '1 minim = 2 crochets' in favour of merely 'minims take longer than crochets'; this, of course, entails that time-signatures can no longer properly be used. Examples (26)–(30) below are taken from Abercrombie (1964: 220), though the musical notations are my own.

(26) | [*one*] [*for the* | road] | 

(27) | [*anything*] | [more] | 

(28) | [*seven*] [*o'* | clock] | 

(29) | [*after*] [*the* | war] | 

(30) | [*nobody*] | [knows] | 

5. The notation used in the examples

Probably the most effective way of explaining the way the hierarchies operate is to work through some examples. Examples 1, 2 and 3 (Figures 2, 3 and 4 respectively) show how the same string, with the same syntactic structure, may be intonated, stressed and footed in several quite different ways. Example 4 (Figures 5, 6) indicates how supertonegroups with more than one tonegroup are handled.

Notation in Figures 2–6 is as follows:

(i) N&V (1st column): Five of N&V's seven hierarchical levels. 'Foot' has been removed because it is now dealt with in Hierarchy II; 'phonological phrase' has been omitted (see footnote 2).
(ii) EF (2nd column): The three hierarchical levels that I would propose for Hierarchy I, with 'superunits' on two of them (see section 2.3).

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

(iii) Hierarchy I is shown in the box structure diagram at the top of each figure: I prefer to use boxes rather than branching tree-diagrams, since (a) this gives an unambiguous notation for showing the domains and foci of suprasegmental elements, and (b) I follow N&V (1986: 7–10, 84–87) in adopting n-ary branching between levels, avoiding the complication of unwanted structure. The disadvantage of the notation is that 'headedness' needs to be shown separately (e.g. the oblique stroke notation for the focus of word-stress).
(iv) Dotted arrows: the route by which the intonation nucleus 'percolates down' through the levels on Hierarchy I.
(v) Square brackets: these enclose commentary on the processes that work to ensure the proper percolation of the final patterns of pitch and stress.
(vi) 'r': indicates reduced words ('weak forms') within a stress-group. For words which alternate between strong and weak forms (*was*, for example), the weak form is the unmarked one, so that the 'r' notation is strictly redundant, but included only for its mnemonic value. The marked situation (strong form of the word) is brought about by either (a) situational factors such as contrastive stress, or (b) a limited range of syntactic factors, e.g. in Standard English a stranded preposition is usually strong.
(vii) Oblique stroke: indicates the stressed syllable of the word (computable by word-stress rules or, for a few words, marked in lexical storage).
(viii) Double oblique stroke: indicates the location of contrastive stress (applied under the influence of features of the situation in which the utterance is produced).
(ix) The working of Hierarchy II is shown in the bottom half of each figure. The number of actual stresses determines the number of feet needed. Feet are equal in duration to begin with, but may need adjustment depending principally on the number of unstressed syllables they contain.
(x) Stressed syllables, notated as 'o' together with an indication (where appropriate) of level and/or direction of pitch movement, are placed as 'accented beats' in each foot.
(xi) The interaction of feet and superwords (stress groups) in determining the actual rhythm is stated. Unstressed syllables are indicated by '.', and their pitch is determined by the pitch movement associated with the stressed syllables.

6. The examples

6.1 *Example 1* (Figure 2)

This consists of the sentence *Bill was at a conference*, pronounced in what might be described as a neutral manner: no emphasis or contrastive stress, no situation with particular aspects which will make special intonation likely.

We will assume that the syllable- and word-structure of this utterance are

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

N&V	EF							
U	SuperTG	**[Intonation: unmarked]**						
I	TG	**[Final in SuperTG, so *Falling nucleus*] [Sentence-stress unmarked, so nucleus on *last SuperW*]**						
C	SuperW		**[Nucleus on *unreduced W of Super W*]**					
ω	W		**r**	**r**	**r**	**[Nucleus on *word-stressed S of W*]**		
S	S	/ **Bill**	**was**	**at**	**a**	/ **con -**	**fer -**	**ence**

Two stresses are needed, therefore TWO FEET are called for.

F

Unstressed syllables: 3 in the first foot, 2 in the second.
So STRETCH THE FIRST A LITTLE and SHORTEN THE SECOND.

Insert the stressed syllables as first 'beat' of each foot.

F

Bill con-

Feet/*SuperWord interaction*: The second foot contains no SuperW boundaries, so the rythm FOLLOWS THE SYLLABLE WEIGHTS.

The first foot contains a SuperW boundary before *was*, so the rhythm of the unstressed syllables FOLLOWS THE 'DELAYED' PATTERN.

F

Bill was at a con- fe- rence
[wəzətə]

Figure 2
Example 1

uncontroversial and unchanging, for Examples 2 and 3 as well as this one; stress-group and tone-group structure, however, are the things that vary to give the different readings of the sentence. In this example the structures on these levels are maximally unmarked: there is no need to indicate special

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

intonation or contrastive stress. The placement of the intonation nucleus and other sentence-stresses are spelt out by the downward route which the dotted arrows follow through the structural diagram.

N&V	EF							
U	SuperTG	**[Echo-question, so Intonation: HighRise]**						
I	TG	**[Final in SuperTG, so *bears nucleus*] [Contrastive stress on *was*, so nucleus on *that SuperW*]**						
C	SuperW		//	**[Nucleus on the *only W of Super W*]**				
ω	W			**r**	**r**	**[Nucleus on the *only S of W*]**		
S	S	/ **Bill**	// **was**	**at**	**a**	/ **con -**	**fer -**	**ence**

Three stresses are needed, therefore THREE FEET are called for.

F | | | |

Unstressed syllables: none in the first foot, 2 in the second, 2 in the third. So SHORTEN THE FIRST A LITTLE and STRETCH THE OTHERS.

Insert the stressed syllables as first 'beat' of each foot.

F | o
Bill | o
was | o
con- |

Feet/SuperWord interaction: The third foot contains no SuperW boundaries, so the rhythm FOLLOWS THE SYLLABLE WEIGHTS.

The second foot contains a SuperW boundary before *at*, so the rhythm of the unstressed syllables FOLLOWS THE 'DELAYED' PATTERN.

F | o
Bill | o . .
was at a
[wɒz] [ət] | o . .
con fer ence |

Figure 3
Example 2

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

N&V	EF					
U	SuperTG	**[Correction, so Intonation: Fall-Rise]**				
I	TG	**[Final in SuperTG, so *bears nucleus*] [Contrastive stress on *was* & *a*, so nucleus on *last of these* SuperWs]**				
C	SuperW		//		//	**[Nucleus on the *unreduced W of SuperW*]**
ω	W			**r**		**[Nucleus on the *only S of W*]**
S	S	/ **Bill**	// **was**	**at**	// **a**	/ **con - fer - ence**

Four stresses are needed, therefore FOUR FEET are called for.

F | | | | |

Unstressed syllables: none in the first foot or the third, 1 in the second, 2 in the fourth.

So SHORTEN THE 1ST AND 3RD A LITTLE, LEAVE THE 2ND AS IT IS, and STRETCH THE 4TH A LITTLE.

Insert the stressed syllables as first 'beat' of each foot.

F | o Bill | o was | o a | o con- |

Feet/*SuperWord interaction*: The fourth foot contains no SuperW boundaries, so the rhythm FOLLOWS THE SYLLABLE WEIGHTS.

The second foot contains a SuperW boundary before *at*, so the rhythm of the unstressed syllables FOLLOWS THE 'DELAYED' PATTERN.

F | o Bill | o was . at | o a | o . . con-fer-ence |

[wɒz] [ət] [eɪ]

Figure 4
Example 3

6.2 *Example 2* (Figure 3)

The same sentence syntactically as in Example 1. This time it represents a hearer's echo-question reaction on being given the news that Bill had been at a conference, when he had believed that Bill never attended conferences. The

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

N&V	EF						
U	SuperTG	**[Intonation: unmarked]**					
I	TG	**[Non-Final in SuperTG, so *Fall-Rise nucleus*] [Sentence-stress unmarked, so nucleus on *last SuperW*]**					
C	SuperW	**[Nucleus on *only W of SuperW*]**					
ω	W	**r**	**r**	**[Nucleus on *word-stressed S of W*]**			
S	S	/ **When**	**it**	/ **stops**	/ **rain -**	**ing**	I'll

Two stresses are needed, therefore TWO FEET are called for.



Unstressed syllables: none in the first foot, 1 in the second.
So SHORTEN THE FIRST A LITTLE and LEAVE THE SECOND AS IT IS.

Insert the stressed syllables as first 'beat' of each foot.



Feet/*SuperWord interaction*: The second foot contains no SuperW boundaries, so the rhythm FOLLOWS THE SYLLABLE WEIGHTS.

There are two unstressed syllables belonging in a SuperW with *stops*, so the rythm of the unstressed syllables FOLLOWS THE 'DELAYED' PATTERN, and they are timed as closely as possible to *stops*.

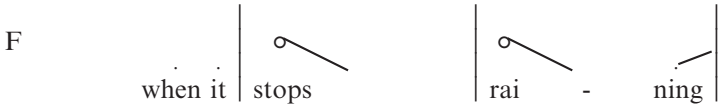


Figure 5
Example 4: first TG

typical echo-question intonation is a High Rise, and it focuses on the non-negativeness of the verb, so the nucleus is attracted on to the verb *was*. The weak form [wəz] is therefore not possible, and must be replaced by the strong form [wɒz].

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

N&V	EF					
U	SuperTG	**[Intonation: unmarked]**				
I	TG		**[Final in SuperTG, so *Falling nucleus*] [Sentence-stress unmarked, so nucleus on *last SuperW*]** ⋯▼			
C	SuperW		**[Nucleus on *unreduced W of SuperW*]** ▼			
ω	W		**r**		**r** **[Nucleus on *only S of W*]** ⋯▼	
S	S	-ing,	**I'll**	/ **go**	**to**	/ **town**

Two stresses are needed, therefore TWO FEET are called for.



Unstressed syllables: one in the first foot, none in the second.
So LEAVE THE FIRST and SHORTEN THE SECOND A LITTLE.

Insert the stressed syllables as first 'beat' of each foot.



Feet/SuperWord interaction: The first foot contains a SuperW boundary before *to*, so the rhythm of the unstressed syllables FOLLOWS THE THE 'DELAYED' PATTERN.

There is one unstressed syllable belonging in a SuperW with *go*, so the rhythm of this unstressed syllable FOLLOWS THE 'DELAYED' PATTERN, and it is timed as closely as possible to *go*.



Figure 6
Example 4: second TG

6.3 *Example 3* (Figure 4)

Again the same sentence syntactically as in Example 1. It is now uttered in response to a statement that Bill was at, say, a syntax conference, by a hearer who knows that Bill was in fact at a phonology conference. The typical

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press

'correction' intonation is a Fall–Rise, and the focus is the determiner of *conference*, so the nucleus is attracted on to *a*, which must therefore have its strong form [eɪ]. The verb *was* also receives a certain amount of emphasis, and therefore has the strong form [wɒz].

6.4 *Example 4* (Figures 5, 6)

A sentence forming a supertonegroup, with two clauses, each of which constitutes a tone group. Each tone group is treated in a separate diagram; the single supertonegroup is thus 'torn' between the two diagrams. The intonation and sentence-stress are neutral: this means the non-final tone group will have a Fall-Rise intonation, and the final one a Fall.

REFERENCES

Abercrombie, D. (1964). Syllable quantity and enclitics in English. In Abercrombie, D., Fry, D. B., MacCarthy, P. A. D., Scott, N. C. & Trim, J. L. M. (eds.), *In honour of Daniel Jones*: *papers contributed on the occasion of his eightieth birthday*. London: Longmans. 216–222.

Bolinger, D. L. (1981). Two kinds of vowels, two kinds of rhythm. Bloomington, IN: Indiana University Linguistics Club.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Giegerich, H. J. (1992). *English phonology*. Cambridge: Cambridge University Press.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.

Goldsmith, J. A. (ed.) (1995). *The handbook of phonological theory*. Oxford: Blackwell.

Harris, J. W. (1995). Projection and edge-marking in the computation of stress in Spanish. In Goldsmith (ed.), 868–887.

Hulst, H. van der (1984). *Syllable structure and stress in Dutch* (Linguistic Models 8). Dordrecht: Foris.

Kager, R. (1995). The metrical theory of word stress. In Goldsmith (ed.), 367–402.

Luick, K. (1898). Beiträge zur englischen Grammatik, III: Die quantitätsveränderungen im laufe der englischen sprachentwicklung. *Anglia* **20**. 335–362.

McCarthy, J. J. & Prince, A. S. (1995). Prosodic morphology. In Goldsmith (ed.), 307–366.

Nespor, M. & Vogel, I. (1986). *Prosodic phonology* (Studies in Generative Grammar 28). Dordrecht: Foris.

Pike, K. L. (1967). *Language in relation to a unified theory of human behavior* (Janua Linguarum, series maior 24). The Hague: Mouton.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry* **11.** 563–605.

Selkirk, E. O. (1984). *Phonology and syntax*: *the relation between sound and structure*. (Current Studies in Linguistics 10.) Cambridge, MA: MIT Press.

Trim, J. L. M. (1959). Major and minor tone-groups in English. *Le Maître Phonétique* **112.** 26–29. [Reprinted in Jones, W. E. & Laver, J. (eds.) (1973). *Phonetics in linguistics*: *a book of readings*. London: Longman. 320–323.]

Author's address: *Department of Linguistic Science,*
University of Reading,
P. O. Box 218,
Whiteknights,
Reading RG6 6AA,
U.K.
E-mail: *ErikFudge@aol.com*

https://doi.org/10.1017/S0022226799007641 Published online by Cambridge University Press