



**HAL**  
open science

# The Stress Patterns of English Verbs: Syllable Weight and Morphology

Quentin Dabouis, Jean-Michel Fournier

► **To cite this version:**

Quentin Dabouis, Jean-Michel Fournier. The Stress Patterns of English Verbs: Syllable Weight and Morphology. Nicolas Ballier; Anne Przewozny; Jean-Michel Fournier; Eiji Yamada. New Perspectives on English Word Stress, Edinburgh University Press, pp.154-191, 2023, 9781399519939. halshs-04252455

**HAL Id: halshs-04252455**

**<https://shs.hal.science/halshs-04252455v1>**

Submitted on 21 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## The Stress Patterns of English Verbs: Syllable Weight and Morphology

Quentin Dabouis and Jean-Michel Fournier

### 0 Introduction

Most studies about English verb stress, however interesting from a theoretical point of view, seem to be based on limited sets of examples, leading their authors to assertions about the stress behaviour of English that are not actually vindicated through comprehensive data. Following Lionel Guierre’s teachings, this study sets out to analyse as comprehensive a corpus as possible, based on dictionary data: though limited in some extent by the very nature of the source, it can still be seen as a first step towards a reliable description of the category. The paper is structured as follows: first, we review the literature on English stress and distinguish two generalisations which were proposed in Chomsky and Halle’s *Sound Pattern of English* (section 1). One generalisation refers to the weight of the final syllable of verbs while the other refers to the presence of a semantically opaque prefix. We then set out to evaluate the empirical validity of these two generalisations, as well as to propose a global overview of how English verbs are stressed, all morphological categories included. In section 2, we present how the data was collected, cleaned and classified. In the following section, we detail the stress distributions found in all the morphological categories found in the data (section 3). At the end of the presentation of our results, we measure the empirical validity of the two generalisations in the relevant morphological categories and show that they have similar efficiencies, but different theoretical costs. Crucially, we find that the data supports an analysis including semantically opaque morphological constituents among the possible determiners of stress.

### 1 English Verbs: Stress, Syllable Structure and Morphology

#### 1.1 The Two Generalisations

In Chomsky and Halle (1968; henceforth *SPE*), there are two parameters determining the position of primary stress in verbs that many morphologists would now analyse as monomorphemic. The first parameter is the segmental makeup of the final syllable. In their Main Stress Rule, a verb gets penultimate stress if the final syllable contains a ‘weak cluster’ (VC; for example *astónish*, *édit*, *consíder*, *imágin*e, *intérpret*...) and final stress otherwise (for example *maintáin*, *eráse*, *appéar*, *colláps*e, *exháust*, *tormént*, *usúrp*...). Since Kahn (1976), the terms ‘weak cluster’ or ‘strong cluster’ have been abandoned and replaced by the notion of syllable weight. Heavy syllables are now commonly defined as syllables ‘which have branching anywhere within the rhyme constituent’ (Carr 1999: 76), but the final consonant in verbs is usually assumed not to contribute weight to the final syllable (see section 4.1). The generalisation can thus be formulated as (1).

#### (1) The weight-based generalisation.

- In verbs, assign primary stress to
- a. The penultimate syllable if the ultima is light
  - b. The final syllable otherwise

The second parameter invoked by Chomsky and Halle is morphological. ‘Complex verbs’, that they describe as being ‘morphologically analyzable into one of the prefixes *trans-*, *per-*, *con-*, etc. followed by a stem such as *-fer*, *-mit*, *-cede*, *-curn* or *-pel*’ (*SPE*: 94), have a different stress behaviour. They analyse these complex verbs as containing a = boundary. In their theory, this boundary has to be explicitly mentioned in the formulation of rules, which has the consequence of blocking (1a) and leads to the application of (1b), regardless of the weight of the ultima. For dissyllabic verbs, this means all complex verbs should have final stress, which allows them to

account for final stress in words such as *perμίt*, *concurr*, *compél*, *detér* and *transfér*, whose final syllable is light.

Longer verbs are said to undergo the Alternating Stress Rule (ASR): if they receive final stress because of (1b), that stress is retracted to the antepenultimate syllable (for example *éxercise*, *ánalyze*, *cómplicate*, *clárify*). Once again, complex verbs have a different behaviour: they do not undergo the ASR when a = boundary is found between the penult and the ultima, as in *comprehénd*, *apprehénd*, *intervéne*, *introspéct*, *introduíce*. To account for that behaviour, Chomsky and Halle, formulate the ASR as in (2).

## (2) Alternating Stress Rule

$V \rightarrow [1 \text{ stress}] / \_ \_ C_0 (=) C_0VC_0 [1 \text{ stress}] C_0 ]_{NAV}$

This Rule allows for the presence of a = boundary between the antepenult and the penult (for example *com=plicate*) but not between the penult and the ultima (for example *contra=dict*). This correctly predicts antepenultimate stress in the former case and final stress (no retraction) in the latter case. An alternative way to analyse the behaviour of complex verbs is Fournier's (2007) 'Germanic Law' in (3).

## (3) The Germanic Law

Overlook the prefix(es) and assign stress normally to what remains.

This law means that stress will be assigned to the root of opaque prefixed verbs, regardless of the number of prefixal syllables or the length of the root (for example *contáin*, *devélop*, *intervéne*). Note that, in Fournier's model, the Germanic Law is subordinated to the effects of stress-affecting suffixes such as *-ate*, which may impose primary stress to be placed on the prefix (for example *désignate*, *étricate*, *rénovate*).

Although both are used in *SPE* to account for the stress behaviour of verbs (although the Germanic Law is not formulated as in (3)), the two generalisations have known different fates in the literature.

### 1.2 Two Generalisations, Two Different Fates

Most proposals since *SPE* (1968: 69) claim that the stress pattern of simplex verbs is determined by the weight of their final syllable (see for example Burzio 1994: 43; Hammond 1999: 263; Hayes 1982). This is also the rule that one finds in textbooks such as Cruttenden (2014: 244-245) and Roach (2000: 110). Only the weight-based generalisation is put forward and little is said on *SPE*'s 'complex verbs'. The reason for that state of affairs might have to do with the fate of the = boundary.

Shortly after the publication of *SPE*, the = boundary was done away with by Siegel (1974, 1980), who introduced the now commonplace 'Class I' and 'Class II' classification of English affixes, which corresponds to the other two boundaries found in *SPE*: + and #, respectively. Her new analysis of stress placement in long verbs does not radically differ from that of *SPE*: she claims that the blocking of the ASR has nothing to do with the presence of the = boundary but with the morphological nature of the final syllable: if it is a root, the ASR does not apply (for example *interséct*) and, if it is a suffix, it does (for example *désignâte*).<sup>1</sup> However,

---

\* We would like to thank Ricardo Bermúdez-Otero, Patrick Honeybone, Jacques Durand, Philip Carr, Renate Raffelsiefen, Ingo Plag, Sabine Arndt-Lappe and the audiences of the several conferences in which sections of this paper have been presented for their useful comments and suggestions. All mistakes are ours alone.

<sup>1</sup> Incidentally, Siegel's criticism is based on dubious examples that she describes as 'the ones I have been able to find' (without mentioning how the search had been conducted) such as *bicuspid*, *inhibit*, *prohibit*, *inhabit*,

dissyllabic verbs are not explicitly discussed. This is problematic because doing away with the = boundary to replace it with the + boundary has the consequence that these words should behave like words with no internal structure. As a consequence, words such as *permi*t or *compé*l are incorrectly predicted to receive penultimate stress.

This did not stop the reference to these 'complex verbs' from progressively disappearing from generative analyses of stress placement in English verbs. Such verbs are explicitly mentioned in early works such as Halle and Keyser (1971: 37) and Liberman and Prince (1977) but examples of 'complex verbs' have sometimes been used to illustrate the claim that verbs with a heavy final syllable have final stress (for example *prevent* and *decide* in Burzio (1994: 43)). Hammond (1999: 71) observes that '[w]hile most speakers of English are incapable of assigning any sort of meaning to elements like *ob-*, they are aware that it forms some sort of building-block in the vocabulary of English', which does not prevent him from including them in his analysis of 'monomorphemic' words in English.

*SPE*'s analysis did not use = only for the purposes of stress assignment. It was also used to account for the specific behaviour of the prefixes introduced with the = boundary with regards to vowel reduction (sometimes called 'destressing').<sup>2</sup> It has often been claimed that these prefixes generally see their vowel reduces even though it is in a closed syllable, as opposed to non-prefixed words (*SPE*: 118; Halle and Keyser 1971: 37; Liberman and Prince 1977: 284-285; Guierre 1979: 253; Selkirk, 1980; Hayes 1982; Halle and Vergnaud 1987: 239; Pater 2000; Hammond 2003; Collie 2007: 129, 215, 318-319), which has recently been empirically confirmed (Dabouis and Fournier 2019).

Reference to the 'complex verbs' in *SPE* and their specific stress behaviour did not disappear in all strands of the literature. It is mentioned in Fudge (1984: 165) and has been developed within the framework put forward by French linguist Lionel Guierre (Chapter 2, this volume; Fournier 2010b; Guierre 1979, 1984; Trevian 2003). Many authors have assumed that *SPE*'s = boundary applied to Latinate items only, but this seems to be an overinterpretation as it is never said explicitly in *SPE*. Following Guierre (1979), we assume that the words concerned can be Latinate (as most examples cited so far) or Germanic (for example *begin*, *forget*, *understand*) and that what characterises them is mainly their semantic opacity.

However, there is one issue which we have not yet addressed. 'Complex verbs' have often been treated alongside words with no internal structure in most post-*SPE* analyses of verb stress placement (see the examples above), so one can wonder if this departure from previous analyses was motivated by a different analysis of their structure. In other words, can *SPE*'s 'complex verbs' really be treated as morphologically complex?

### 1.3 Are *SPE*'s 'Complex Verbs' Actually Complex?

In the morphological literature, the status of historically prefixed verbs in contemporary English are a well-known problem (Anderson 1992: 55; Aronoff 1976: 12-15; Bauer et al. 2013: 15-16; Castair-McCarthy 2002: 23-26; Katamba and Stonham 2006: 23; Mudgan 2015; Plag 2003: 24-27) as they are recurring forms with meanings which are not always clearly identifiable and therefore constitute a challenge to the standard definition of the morpheme as the minimal meaningful unit. Ben Hedia and Plag (2017) clearly formulate the way this problem can be addressed:

---

*evanesce*, *immature*, *premature* or *determine*. She claims that *SPE* predicts that these verbs would receive stress on their prefix. This is clearly wrong: *SPE* predicts stress retraction only in the case where the verbs receive final stress. Apart from *immature* and *premature*, all of these verbs have a light ultima and are correctly predicted to receive penultimate stress. Moreover, *bicuspid*, *immature* and *premature* all have semantically transparent prefixes and would have been described more accurately using # than +.

2 They also use it to account for /s/-voicing (*SPE*: 95), which will not be discussed here.

Locative *in-* belongs to a set of Latinate forms in English that are often discussed in the context of morpheme-based approaches to word structure (e.g. Aronoff, 1976, 12; Bauer et al., 2013, 15f; Don, 2014, 15; Lieber, 2010, 41f; Plag, 2003, 24f). In this set we find historically Latin prefixes such as *ad-*, *con-*, *in-*, *re-*, but also bound roots, for example *-ceive*, *-mit* or *-fuse*. These forms may sometimes have a clear meaning, sometimes not. Depending on which type of definition of morpheme one adheres to, such forms are treated as morphemes (if one’s theory allows for morphemes without clear semantic content), or as units below the word level that have no clear semantic content. According to the latter position these strings are called ‘formatives’ and are considered ‘elements contributing to the construction of words’ (Bauer et al., 2013, 16), and thus as some kind of morphologically relevant unit.

We do not want to make any claims on the precise morphemic status of elements such as *ad-*, *con-* or *-mit* within the scope of this paper, though they clearly need to be distinguished from productive units such as *co-* (‘together’), *ex-* (‘former’) or *re-* (‘again’). However, we argue that it is not because they might indeed not be morphemes that they are not some form of morphological unit, which some have called ‘pseudo-morphemes’ and others ‘formatives’. Accordingly, we will retain the idea that, these elements being ‘morphologically relevant’, it is legitimate to assess their potential interaction with phonology.

There are various kinds of evidence which support the idea that words such as *contain* or *submit* are complex. First, there are other phonological phenomena which have been shown to be sensitive to this kind of morphological structure, such as vowel reduction (see previous section) and secondary stress placement (Dabouis 2016, 2017). Second, a number of psycholinguistic studies on lexical access in visual word recognition have showed that ‘morphological decomposition is a process that is applied to all morphologically structured stimuli, irrespective of their lexical, semantic or syntactic characteristics’ (Rastle and Davis 2008, in a review on the issue). This morpho-orthographic decomposition would appear to take place in the early stages of the recognition process, independently of semantics (see Marslen-Wilson et al. (2008) and references therein). Third, there is evidence from reading studies which shows that prefixes are usually not assigned stress in disyllabic words and so have to be identified in computational models of reading (Ktori et al. 2016, 2018; Rastle and Coltheart 2000; Treiman et al. 2020). Fourth, there are phonotactic sequences found in semantically opaque prefixed constructions which do not occur in morphologically simple words (Guierre 1990; Hammond 1999: §3.3.; for example [dh] in *adhere*). Finally, the distributional recurrence of prefixes and roots can be argued to provide enough evidence for morphological segmentation, independently of semantics (Fournier 1996; Pastizzo and Feldman 2004; Taft 1994).

An open question is that of the criteria that should be used to determine whether a given word contains an opaque prefix and a bound root. Indeed, classical morphological analysis crucially uses shared meaning to identify morphemes. Semantically opaque constituents may sometimes be difficult to identify and there is, to the best of our knowledge, no established procedure for doing so. Of course one can refer to etymology but this entails two types of problems (Pastizzo and Feldman 2004). The first problem is that historically morphologically complex words may not be identifiable as such anymore. The evolution of their pronunciation and/or spelling may have altered their shape in such a way that there is no way that they could be parsed by speakers. The second problem is that, if there are synchronic mechanisms used by speakers to parse semantically opaque complex words, these mechanisms could lead to the parsing of words which actually have no historical morphological structure. But for any such

mechanism to be relevant we first need to make sure unaltered historical prefixation does indeed affect stress placement.

Our position will be to assume that most historically prefixed words are still potentially recognisable in contemporary English. One of the goals of this paper is to seek to establish whether historically prefixed verbs do indeed have a stress behaviour that is distinct from that of words with no internal structure, as was claimed in *SPE*. If they do, then it would imply that phonological analyses of English stress cannot do away with what *SPE* called ‘complex verbs’. It would also raise the question of learnability, which has probably led generative linguists to reject the synchronic relevance of this type of morphological structure (as pointed out by Raffelsiefen, personal communication), and open the way to test out the predictions such a model makes: loss of identifiability or assimilatory identification (things that ‘look like’ prefixes).

#### *1.4 Aims of the Paper*

It has been noted before that many studies on English phonology ‘suffer from a scarcity of systematic empirical evidence’ (Domahs et al. 2014) or are based on ‘convenient samples’ (Wenzky 2004: 12). McMahon (2001) also notes that ‘there is undoubtedly a problem in phonology, especially the sort that rather distances itself from phonetics, of reliance on stock examples and introspection’. Consequently, the generalisations discussed in the previous sections have seldom been tested against large datasets. The only study which we are aware of is that of Guierre (1979). His study showed a strong effect of opaque prefixation, but his results did not strongly support the weight-based generalisation, even though he did find a tendency towards final stress in dissyllables with a heavy ultima. Since Guierre’s study, there seems to have been no empirical evaluation of the efficiencies of the two generalisations. Therefore, the first aim of this paper will be to remedy that state of affairs. To do so, we will constitute a large corpus of pronunciation dictionary data. This will allow us to give a detailed empirical account of both the morphology and the stress patterns of English verbs.

We will test the two generalisations presented in section 1.1 and will seek to establish whether both reference to opaque morphology and syllable weight are needed to account for the placement of primary stress in English verbs. If the study were to confirm Guierre’s results that opaque morphology does determine the position of primary stress in these words, the question of the learnability of such structures will have to be addressed. Note that such an endeavour implies, as a first step, to describe the different morphological categories found in verbs since it will allow us to isolate the morphological categories relevant to test the two generalisations by clearly identifying categories which must be excluded on account of their distinct behaviours (most notably suffixed words, either with stress-shifting suffixes or stress-neutral suffixes, whose effects have been shown to override those of opaque prefixes). It will also allow for a detailed description of stress placement in verbs (maybe most interestingly for categories which are not well-described, such as compounds). Our study will thus provide a general view of the morphological and accentual characteristics of English verbs and of how they interact, which is something which, to our knowledge, has not been done before.

## **2 Methodology**

### *2.1 Data Collection*

The data used in this study comes from the Laboratoire Ligérien de Linguistique’s (LLL) Dictionary Database. This is an ongoing project which gathers three reference pronunciation dictionaries (Delbridge et al. 2009; Jones 2006; Wells 2008) and contains additional information such as syntactic categories, word meaning and will contain token frequencies and morphological analyses. The syntactic categories in the database come from the *Macquarie*

*Dictionary*, which is the only dictionary of the three which is not only a pronunciation dictionary.

For this study, we extracted all 5,236 verbs listed in Jones (2006). Only the main pronunciation for British English was extracted. Words which were marked as ‘rare’, ‘obsolete’, as belonging to another dialect of English (American, Australian...) or which had no entry as verbs in the online *Oxford English Dictionary (OED)* were left out. This led to the exclusions of 293 words.

## 1.2 Conversion

One crucial issue with using the syntactic categories in the *Macquarie Dictionary* is that, like many dictionaries of its kind, it tends to attribute several categories to certain words, even when they are predominantly used within a single category. Henceforth, many of the words which are tagged as ‘verbs’ in the dictionary are in fact verbal uses of nouns or, more marginally, adjectives. As pointed out by Bauer et al. (2013: 279) and Bram (2011), it is not always easy to determine which member of a homophonous pair, one a verb and the other a member of another category, should be analysed as the base of the other. More generally, the question of directionality is a hotly debated topic in the literature on conversion in English. Recent work underlines the necessity of using several criteria in order to establish the directionality of conversion. The criteria which have been used in the literature, although none of them is free of limitations (see Bram (2011) for detailed discussion), include: attestation dates, frequency of occurrence, semantic dependency and semantic range. Following these previous works, we based our identification of those verbal uses on these same four criteria, all of which being based on data recorded on the *OED*.

The **dates of first attestation** were taken from the *OED*. The first member of the pair to be recorded was seen as more likely to be the base of the other. In some cases, it was not possible to compare attestation dates because the *OED* refers to undefined dates such as OE (Old English) or eOE (early Old English). As pointed out by Plag (2003: 108), historical information cannot be used alone because ‘complex semantic changes may overwrite the original direction of conversion’.

The **frequency of occurrence** was taken from the *OED* Frequency Bands, which are based on Google Book Ngram data and other corpora. It was used as a rough indication of which member of the pair is used more frequently than the other, in which case the former is seen as more likely to be the base. This is based on the assumption that ‘derived words tend to have a narrower range of meaning’ (Plag 2003: 111) and that, as a consequence, they tend to be used in fewer contexts.

**Semantic range** was established by looking at the number of meanings given for each member of the pair. The member with more meanings was seen as more likely to be the base.

**Semantic dependency** was established using the definitions of both members of the pair. We checked whether the semantic relationship between the two members of the pair fits within one of the thirteen semantic categories listed by Lieber (2004: 92), which are largely based on those listed by Plag (1999: 220), which are themselves based on earlier literature on conversion. Such categories include the locative ‘put (in)to N’ (for example *jail*), the ornative ‘provide with N’ (for example *staff*) or the causative ‘make (more) Adj’ (for example *yellow*).

For all the words in our dataset, we checked whether a noun existed in the *OED* to establish a first list of potential candidates for conversion. We then collected the relevant information for each pair and made our decision based on the four criteria described above. Whenever there were inconsistencies between the four criteria, the decision was based on semantic dependency.

In the case of conversions from other categories, all fairly obvious, these criteria did not seem helpful and were not used.

This led to the identification of 1,395 converted words (see examples in (4)), which were therefore excluded from the dataset as their stress patterns could be attributed to their original syntactic category.<sup>3</sup>

(4) *author, baby, catalogue, engineer, input, invalid, exile, patent, ready, signature, silence, tapestry...*

They are mostly conversions from nouns: only 94 are conversions from other classes (adjectives, adverbs, interjections), such as those in (5).

(5) *busy, even, narrow, oblique, opaque, overnight, farewell, hello...*

It is worth noting that no criterion could have been used on its own. For instance, if we consider the 1,312 verbs classified as denominal, we find that the noun is more frequent than the verb in only 868 (66%) cases and similarly that only 858 (65%) of the nouns entered the language over fifty years before the verb did, according to the *OED*.

The final dataset contains 3,548 words, 1,918 disyllables and 1,630 longer verbs.

### 2.3 Morphological Classification

As discussed in section 1, one of the aims of this paper is to investigate the relationship between obscure morphology and stress placement in English verbs. Therefore, standard analyses for morphological complexity (for example Plag 2003: §5.2.2; Stockwell and Minkova 2001: 58 *ff.*) cannot be used. While words whose elements have identifiable semantics can confidently be analysed as morphologically complex (even though that meaning can sometimes be rather vague; for example *spect* ‘looking, sight’ in *inspect, suspect, spectacle*), it may not be so easy to do so for words with semantically opaque constituents. This is the case for prefixes (for example most *ob-* prefixed verbs) and for roots (for example *-pel, -fer, -mit*).

In the absence of an established procedure for an analysis of morphological structure which is not based on these definitions of the morpheme, it seems preferable that the procedure we adopt should be replicable and therefore based on accessible, verifiable data. One option for prefixed verbs could be to analyse as morphologically complex any word whose constituents occur elsewhere, even though they do not have a clear identifiable meaning. This methodology seems problematic with those cases where the potential prefix has little segmental content and the potential base<sup>4</sup> is not attested elsewhere (for example *attack, esteem, erase*). If one assumes that distributional recurrence of prefixes and roots is a key mechanism allowing speakers to parse semantically opaque prefixed words, then one can wonder how a synchronic analysis could actually treat those words as morphologically complex. However, there might be other properties which could still lead to their parsing as morphologically complex and such a methodology would therefore miss out on some of the data (although probably not a lot of it). Therefore, it was decided to use etymological information, taken from the *OED*, prior to any further, and synchronic, considerations. Doing so has the advantage that other researchers can replicate our findings using the same objective criterion, even though, at this stage, it comes at the cost of the limitations discussed in section 1.4, that is to say the assessment of the conditions of learnability. But precisely, it also opens up the possibility to lead a contrastive analysis amongst these cases, and thus define further criteria of relevance to contemporary speakers,

---

3 The datasets used in the paper are available on <https://lll.cnrs.fr/actualite/the-stress-patterns-of-english-verbs/>.

4 We will use the term ‘base’ rather than ‘root’ since in a number of cases, the right-hand constituent (or the left-hand one in the case of suffixed form) is morphologically complex (for example *abbreviate, embellish, emulsify*), at least in the view adopted here.



entailing a more restrictive classification method. Of course, one can just as much refute the whole idea and group this category with the base (or simplex) category as distinct results will be given for each of our categories and therefore allow for such a stance.

Since an active base in contemporary English may influence the behaviour of a prefixed verb, whatever the status of its prefix, all such cases have been identified. However, even with an active prefix, the relationship between prefix and base can be either compositional (for example *disprove*, *rewrite*, *unlock*) or not (for example *compile*, *depose*, *react*): we call the latter ‘opaque’ constructions. The two types have therefore been further distinguished. A restrictive interpretation of compositionality led to deem ‘opaque’ all cases of verbs which, though their meaning is transparent, are formed from a nominal (or marginally other non-verb classes) base, on account of category change (for example *devoice*, *discourage*, *endear*): as will be seen, whatever their classification it does not affect the position of primary stress. We identified 144 such cases, 128 from a nominal base and 16 from an adjectival base.

Constructions whose first element is one of the following adverbial or prepositional elements: *back-*, *by-*, *down-*, *fore-*, *in-* (locative), *off-*, *out-*, *post-*, *up-*, *with-* (for example *backfire*, *foreclose*, *outdate*) have been listed separately. In construction, these units seem to behave as prefixes in the case of verbs, but as roots in the case of nouns (Abasq 2007). They are mostly compositional in the case of verbs.<sup>5</sup>

In the case of suffixed verbs, we listed all cases which contain final sequences which may be analysed (at least historically) as suffixes, even when their base is not active in contemporary English (for example *happen*, *vibrate*, *polish*, *compliment*, *fraternize*), that is to say the same strategy as that adopted for prefixed words. We then further divided all suffixed entries between so-called stress-preserving and so-called stress-shifting suffixes,<sup>6</sup> the free or bound<sup>7</sup> status of the base being indifferent in the case of stress-shifting suffixes, as is well established (as far as main stress is concerned). Certain suffixes have been treated as being alternatively stress-preserving and stress-shifting. The verbal suffix *-ize* has often been considered as one of these ‘mixed’ (or ‘dual-level’ in Stratal Phonology) suffixes (Bermúdez-Otero 2018; Guierre 1979: 711-714; Selkirk 1982; Szpyra 1989: §2.2.3; Trevian 2007, 2015: §13.2) and will therefore be presented separately.<sup>8</sup>

This distinct treatment of suffixed words is crucial, as stress-preserving suffixes are irrelevant for stress assignment and stress-affecting suffixes directly determine the position of primary stress, regardless of other factors (and so may neutralise the effects of other types of morphological structures).

The ‘Compounds’ category contains all cases formed from the association of two free bases, the most widely accepted definition (for example *blindfold*, *skyjack*, *cross-examine*). These

---

5 These adverbial constructions are grouped with prefixed constructions within the ‘compositional’ and ‘opaque prefixed’ categories in Table 5.1 below. See also section 3.1.

6 This terminology, used by Fudge (1984) distinguishes classes of suffixes based on their stress behaviour (although other works have shown that differences in stress behaviour are correlated with other characteristics). It is largely co-referent with other terminologies such as the + vs. # boundary distinction in *SPE*, ‘Class-I’ vs. ‘Class-II’ suffixes (Siegel 1974), ‘Level-I’ vs. ‘Level-II’ suffixes (Kiparsky 1982) or ‘stem-level’ vs. ‘word-level’ suffixes (Bermúdez-Otero 2018): ‘Stress-preserving’ suffixes generally have stress on the same syllable as that which bears stress in the base (for example *similar* → *similarly*) while ‘stress-shifting’ suffixes are associated to a fixed position of primary stress, which can be either on themselves (for example *kitchen* → *kitchenette*) or on a syllable which precedes them (for example *égal* → *égalité*).

7 Following Bauer et al. (2013: 14), we assume that a free morph is ‘a morph which has the ability to be a word-form on its own’ and that a bound morph is one ‘that lacks this ability’.

8 Giegerich (1999) claims that all suffixes have a dual affiliation and that, in a stratal model, stratal affiliation is base-driven, not affix-driven.

words need to be treated separately as the previous literature has shown that compound words have a stress behaviour that is distinct from that of non-compound words.

All other constructions which did not fit into any of the previous categories, and whose numbers were too low for an independent analysis, were grouped into the ‘Other’ category: truncations, back-formations, neoclassical formations, bound bases compounds and some other, obscure cases (for example *liaise*, *animadvert*, *choreograph*, *genuflect*, *flabbergast*).<sup>9</sup>

Finally, the ‘simplex’ (or root) category (for example *argue*, *harangue*, *listen*) consists of words for which no morphological structure can be identified.

The distribution of the data among the different categories we have just presented is shown in Table 5.1.<sup>10</sup>

Table 5.1. Morphological distribution in the dataset

	Disyllables		Long verbs		TOTAL	
Compositional left-edge constructions	227	12%	362	22%	589	17%
Opaque prefixed constructions - free base	364	19%	110	7%	473	13%
Opaque prefixed constructions - bound base	665	35%	105	6%	771	22%
Simplex	278	14%	15	1%	293	8%
Bound base + suffix	112	6%	18	1%	130	3%
Stress-shifting suffixes	40	2%	617	38%	657	19%
Stress-preserving suffixes - free base	114	6%	6	<1%	120	4%
-ize - free base	2	<1%	334	20%	336	9%
Compounds	88	4%	24	2%	112	3%
Other	28	1%	39	2%	67	2%
<b>TOTAL</b>	<b>1,918</b>		<b>1,630</b>		<b>3,548</b>	

The difference in the morphological distribution of the two inventories is striking:

- Disyllables are massively ‘left-edge constructions’ (they are constructed on the left of the root: compositional + opaque prefixed: 66%), while the rest is roughly equally divided between simplex and suffixed forms (about 15% each). This is clearly different from nouns; in his dataset, Guierre (1979) similarly found that 77% of verbs are prefixed while only 9% of nouns are.
- Longer verbs are mostly suffixed (60%), though left-edge constructions still represent a significant part of the inventory (35%), and long simplex verbs are extremely rare. This account is however somewhat misleading: suffixation is mostly due to only 3 suffixes, *-ate*, *-ify* and *-ize*, and no less than 335 (that is to say an additional 20% of all long verbs) out of the 617 verbs with a stress-shifting suffix are also prefixed.<sup>11</sup> Stress determination apart, English verbs of all length thus display a remarkable proportion of left-edge constructions.

9 It is clearly a kind of ‘leftovers’ group, but it seems more convincing to admit them as such within an effort at a comprehensive description of the English verb category.

10 Relevant examples are given in the text above.

11 The verbs with stress-preserving suffixes are not taken into consideration since these suffixes do not affect the pronunciation of the base. In other words, the structure before the suffix is that of the base, not that of the verb itself.

### 3 The Stress Patterns of English Verbs in British English: A Comprehensive Description

Results will be presented according to phonological behaviour types: first compositional constructions, stress-shifting suffixes whatever the status of their base (active/inactive) and its morphology, stress-preserving suffixes and the mixed suffix *-ize* when they are attached to a free base, whatever its morphology; then the remaining cases, divided into five categories: compounds, other, prefixed words, non-prefixed suffixed (with stress-preserving suffixes) words with a bound base, and simplex words. Given the limits of this article, only samples will be displayed but the full data is available on demand.

#### 3.1 Compositional Constructions (n=589)

Semantically compositional constructions whose leftmost formative is a prefix (411 words) or an adverbial particle (178 words) tend to behave as two distinct phonological domains, as evidenced by stress clashes (for example *dèclássify*, *rèwríte*, *òutbáalance*) or morphological geminates (for example *di[ss]atisfy*, *mi[ss]pell*).<sup>12</sup>

Primary stress is on the base in the vast majority of words (581/589, 99%). The only words in which that is not the case are shown in (6).

(6) *fórecast*, *ínfix*, *óutcry*, *óutsource*, *cóuntercharge*, *cóuntersign*, *cóuntersink*, *óversew*

Interestingly enough, all constructions with an adverbial particle (or ‘adverbial constructions’) are formed from an active base, with the exception of locative *in-* (which is not surprising given its history;<sup>13</sup> for example *infer*, *inject*, *inspect*),<sup>14</sup> and are mostly semantically compositional, again except for *in-* constructions (for example *inflame*, *inform*, *intend*): despite our relatively restrictive definition of compositionality, only 25 cases (for example *withdraw*, *outwit*, *understand*) were analysed as opaque, that is < 14% of the class (*in-* constructions excluded). These cases are included within the opaque prefixation categories presented below.

#### 3.2 Stress-Shifting Suffixes (n=657)

Some suffixes determine the position of primary stress in vast parts of the data. The most common one is *-ate*, with 536 words. Verbs in *-ate* usually have final stress in 33/36 disyllables (for example *dictáte*, *negáte*, *gradáte*... but *cúrate*, *fúrcate*, *striáte*)<sup>15</sup> and antepenultimate stress in 494/500 longer verbs (for example *démonstrate*, *originate*, *maturate*... but *destrúncate*, *èquilibráte*, *inspíssate*, *óxygenate*, *péregriúate*, *térgiversáte*)<sup>16</sup>. It seems interesting to associate to this class the 3 disyllables and 13 longer verbs in *-ite/-ute* (for example *uníte*, *salúte*, *èxpédite*,

---

12 See Dabouis (2016), Raffelsiefen (1993)). In Prosodic Phonology, this is commonly analysed as evidence that the prefixes have their own prosodic words (Booij and Rubach 1984; Raffelsiefen 1993, 1999, 2007; Szpyra 1989; Wennerstrom 1993).

13 As stated in the *OED*: ‘Since *in-* prefix1 [from Germanic] and *in-* prefix2 [from Latin] are identical in form, and to a great extent in sense, they come in later use to be felt as one and the same prefix’, both also adverbs in their original languages though the modern adverb comes from Germanic. This explains why the corpus contains mostly forms inherited from Latin (through French in a number of cases) with opaque bases like the examples above, despite the active adverb.

14 And *upbraid*, its relationship with *braid* being nowadays farfetched at best.

15 Note that in American English, there is a significant difference ( $\chi^2 = 14.45$ ,  $df = 1$ ,  $p < .001$ ) between prefixed verbs ending in <-ate> (for example *ablate*, *collate*, *debate*, *inflate*, *relate*, *translate*) and suffixed verbs ending in *-ate* (for example *castrate*, *dictate*, *graduate*, *migrate*, *rotate*, *vibrate*), as the former have final stress as their main pronunciation in 13/14 words, while the latter have final stress in 9/35 words.

16 Words containing a hiatus (for example *álienate*, *améliorate*, *detériorate*, *orientate*) can be analysed as either being stressed /(-)1000/ or /(-)100/. We are treating them here as being stressed /(-)100/.

*cónstitute, pérsecute*), with similar stressing except the *-tribute* verbs (*atríbute, contribúte, dístríbute*<sup>17</sup>), which seems to allow for a *-Vte* (or */-V̄t/*) constraint.

The second largest class concerns the 89 verbs ending in *-ify/-efy*, which all have antepenultimate stress (for example *divérsify, glórify, rárefy, solémnify, vílify*).<sup>18</sup>

The dataset also contains 9 words ending with the suffix *-esce* and which all have final stress (for example *àcquiesce, èfflorésce, lùminésce*). We have associated the verb *rèminísce* to this category because, although it does not formally contain the suffix *-esce*, it seems to owe its final stress to an (orthographic?) analogy with this category.

The 6 remaining verbs contain various Class-I suffixes and follow the regular patterns associated with these suffixes: *cashíer, búccanéer, elèctionéer, appórtion, dísilútion, envíision*.

Let us finish this section by pointing out that words with opaque prefixes behave in the same way as words without such prefixes, even when that entails stressing on the prefix (for example *ácclimate, cómpensate, dédicatè, émigrate, ímplicate, óbfuscate, prómulgate, súbjugate* cp. *áternate, círculate, dóminate, émulate, ímitate, óscillate, póstulate, strángulate*): the rules or constraints which regulate stress placement in words with stress-shifting suffixes clearly dominate the potential ‘stress-repelling’ effects of prefixes, just as much as they interfere with the potential effect of the weight of the final syllable.

### 3.3 Stress-Preserving Suffixes – Free Base (n=120)

Contrary to stress-shifting suffixes whose associated stress-pattern is systematic whatever the status of the base, stress-preserving suffixes preserve the pronunciation of free bases. Apart from a few exceptions, no stress shifts are observed.

The most commonly found stress-preserving suffix is *-en* with 59 verbs transparently derived from a free base (for example *awáke → awáken, fát → fátten, shórt → shórten*).

The second most common stress-preserving suffix is *-le*. It is found only in 42 disyllables. Stress, obviously enough, never shifts to the second syllable (for example *cráckle, hándle, síckle*). Although the frequentative (and sometimes diminutive) meaning of this historical verbal suffix is often still apparent, the suffix is clearly no longer active in contemporary English: these cases could alternatively be analysed as non-derived.<sup>19</sup> Additionally, 10 of these words display important alterations of the root, which further questions their derived status: *boom → bumble, daze → dazzle, game → gamble, joust → jostle, nose → nuzzle, prate → prattle, stride → straddle, throat → throttle, wade → waddle*.

Finally, we find 19 other words containing stress-preserving suffixes. Among those, the 17 shown in (7) fully preserve the stress pattern of their base.

(7) *banish, blandish, bumper, burnish, distinguish, flutter, jigger, moulder, mullock, multiply, notice, pester, prophesy, putter, skitter, upholster, whimper*

The other two show an exceptional stress shift: *bómb → bombárd* and *ímage → imáginè*. However, considering the semantic drift between the two words, *imagine* might be better analysed as a deradical formation. Similarly, *bombard*, originally converted from the noun referring to the military engine, has now lost that reference in favour of a direct but somewhat

---

17 Both *contribute* and *distribute* have variants with antepenultimate stress.

18 Including *satisfy*, with the very same historical suffix *-fy*, but the only one in contemporary English not preceded by either *e* or *i* (the latter being either part of the stem in loanwords or connective in English formations).

19 Though such is not the position we have adopted here, it must be noted that the same observation also applies to frequentative verbal *-er*, and verbal *-ish* inherited from French infinitive *-ir*; see examples in (7) below.

loose relationship with *bomb*, with an intensive value that has been lost in all the other derivatives in *-ard*. Additionally, these derivations are both one of a kind.

### 3.4 *-ize* – Free Base (n=336)

The dataset contains 318 *-ize* derivatives which may be analysed as stress-preserving constructions considering that the position of stress is mostly identical to that found in the base (for example *cánon* → *cánonize*, *chárácter* → *chárácterize*, *pròpagánda* → *pròpagándize*, *stýle* → *stýlize*). The remaining 18 words do show stress shifts, as shown in (8).

- |  |   |
|--|---|
| (8) <i>advért</i> → <i>ádvertize</i>           | <i>épilogue</i> → <i>epílogize</i>                |
| <i>àlumínium</i> → <i>alúminize</i>            | <i>épigram</i> → <i>èpigrámmatize</i>             |
| <i>ànaesthésia</i> → <i>anáesthesize</i>       | <i>Galváni</i> → <i>gálvanize</i>                 |
| <i>ánthropomorph</i> → <i>ánthropomórphize</i> | <i>gélatin</i> → <i>gelátinize</i>                |
| <i>apòtheósis</i> → <i>apótheosize</i>         | <i>hypnósis</i> → <i>hýpnótize</i>                |
| <i>áttitude</i> → <i>àttitúdinize</i>          | <i>immúne</i> → <i>ímmunize</i>                   |
| <i>canál</i> → <i>cánalize</i>                 | <i>pànegýric</i> → <i>panégyrize</i>              |
| <i>cátholic</i> → <i>cathólicize</i>           | <i>vólatile</i> → <i>volátitalize</i>             |
| <i>éblem</i> → <i>emblématize</i>              | + parasynthetic <i>demóralize</i> ← <i>morale</i> |

However, for a number of these cases the derivation can be questioned in contemporary English and some might prefer to classify these among bound base constructions.<sup>20</sup>

### 3.5 Compounds (n=112)

Main stress tends to be on the first constituent in disyllables (75% of 89 words):

- (9) *blíndfold*, *hándcraft*, *kídnáp*, *skýjack*... vs. *dèep-frý*, *hànd-píck*, *vòuchsáfe*...

but is equally divided between first and second constituent in words of three syllables or more (25 words):

- (10) *bábysit*, *wíndow-shop*... vs. *cròss-exámine*, *stàge-mánage*...

Let us examine the data more closely to see which factors could explain the position of stress in these compounds.

Previous research has shown that verbal compounds are quite uncommon and that they are often formed through conversion or backformation (Castairs-McCarty 2002: 60-61; Marchand 1969: 100 ff.; Plag 2003: 154-155). The 271 converted compounds present in the initial dataset have already been excluded (see section 2.2). Following Bauer et al. (2013: 280-281), we identified 26 cases of backformation based on the dates of attestation (taken from the *OED*) of verbal compound and its base and the semantic dependency between them.<sup>21</sup> These cases are shown in (11).

- (11) *air-condition*, *back-pedal*, *brainwash*, *bulldoze*, *chain-smoke*, *crash-land*, *cross-refer*, *dive-bomb*, *freeload*, *freeze-dry*, *gatecrash*, *ghostwrite*, *headhunt*, *jaywalk*, *jerry-build*,

<sup>20</sup> The derivation is also not simply *-ize* in the cases of *attitud/inize*, *emblem/atize* and *epigram(m)/atize*.

<sup>21</sup> We treated as semantically dependent compounds whose definition in the *OED* or Dictionary.com mentions the base (for example *bulldoze* ‘to clear away by or as if by using a *bulldozer*’).

*lip-read, mass-produce, plea-bargain, shoplift, shrink-wrap, spring-clean, stage-manage, steamroll, touch-type, typewrite, windsurf*

We then compared the stress patterns of these compounds to those of their bases and found that the position of primary stress is systematically identical. Therefore, we cannot exclude the possibility that the stress patterns of these verbal compounds actually reproduce those of their bases, and so they were excluded from further analyses.

Most studies on compound stress have focused on noun + noun constructions and there is no previous extensive work on verbal compounds that we know of. Previous work on nominal compounds has shown the importance of lexicalisation, the semantic relationship between the two elements, length, informativity and analogy (see, among others, Bell and Plag 2013; Giegerich 2009; Plag 2006). The data was therefore coded as follows:

- Lexicalisation was approached using two variables:
  - SPELLING: more lexicalised compounds tend to be spelled as one word while less lexicalised compounds tend to be spelled hyphenated or as two words (Plag et al. 2007, 2008);<sup>22</sup>
  - FREQUENCY: the frequency of the whole construction was extracted from SUBTLEX-UK (Van Heuven et al., 2014) and log-transformed, as is standard procedure.
- Less informative elements have been found to be less likely to be stressed (Bell and Plag 2013). Following Bell and Plag (2013), these are assumed to be more frequent and have a greater number of senses. Therefore, we used two measures of informativity:
  - FREQUENCY of each of the two constituents, following the same procedure as that used for the whole construction.
  - SYNSETS for the two constituents. Synsets correspond to sets of words with similar meanings. We used the WordNet lexical database (<https://wordnet.princeton.edu/citing-wordnet> - Accessed 01/05/2021) to manually extract the number of synsets for each constituent.
- We coded for the number of syllables of the first and second constituents, of the whole construction, and of syllables following the last strong syllable of the first constituent.
- The syntactic categories of Constituent 1 and Constituent 2 were coded for.

Given the limited size of our dataset and the diversity of the semantic relationships found in it, we chose not to study this variable, for it would seem difficult to draw any conclusions with such small samples. For the same reasons, it is difficult to study analogy as few constructions share a constituent, but this parameter is discussed below.

These factors were tested in a binary logistic regression with the position of main stress in the compound as the dependent variable.<sup>23</sup> Two variables were found to be significant predictors of the position of stress: the syntactic category of Constituent 1 and spelling. The results are shown in Table 5.2 and the distribution of the data is shown in Figure 5.1.

Table 5.2. Binary logistic regression for the position of main stress in verbal compounds

---

22 See also Guierre (1979) on the observation that compounds spelled as one word usually have more initial stress than compounds which are hyphenated or spelled as two words.

23 Some variables were correlated (Synset count and frequency, spelling and frequency) and so they were tested separately.

	95% C.I.			p-value
	Lower	OR	Higher	
CATC1-NOUN	0.010	0.053	0.224	0.000163
CATC1-ADJ	0.034	0.272	1.809	0.183954
CATC1-VERB	0.0725	0.372	1.680	0.211397
SPELLING-ONEWORD	0.053	0.194	0.627	0.008040

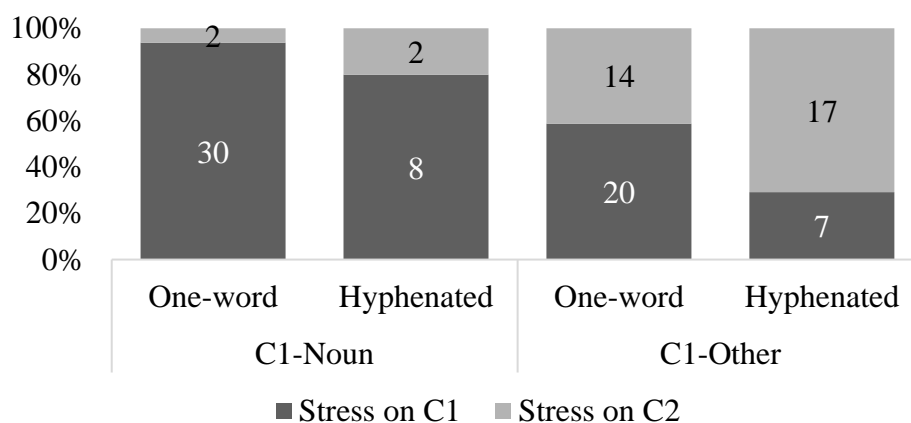


Figure 5.1. Position of main stress in verbal compounds depending on the syntactic category of Constituent 1, spelling and the length of Constituent 2

As can be seen in Figure 5.1, stress on Constituent 2 is more likely if Constituent 1 is not a noun and if the compound is hyphenated. The absence of effects from the other factors such as informativity could be due to the restricted size of the dataset.

Let us finish with the question of analogy. It has long been observed that, in noun-noun compounds, compounds which share a constituent tend to be stressed in the same way. For example, compounds which end in *street* are stressed on their first constituent while those which end in *avenue* are stressed on their second element (Arndt-Lappe 2011). As mentioned above, our dataset being quite small, there are few instances of compounds sharing a constituent. However, we have grouped those which do in Table 5.3.

Table 5.3. Position of main stress in compounds which share a constituent

	Shared constituent	Stress on Constituent 1	Stress on Constituent 2
Constituent 1	<i>about</i>	∅	<i>about-face, about-turn</i>
	<i>air</i>	<i>air-cool, air-dry</i>	∅
	<i>back</i>	<i>backcomb, backfill, backslide, backtrack, backwash</i>	<i>backdate, backfire</i>
	<i>cross</i>	<i>crossbreed, crosscut, crosshatch</i>	<i>crosscheck, cross-examine, cross-fertilize, cross-question</i>
	<i>hand</i>	<i>handcraft</i>	<i>hand-pick</i>
	<i>side</i>	<i>sideslip, sidestep</i>	∅
Constituent 2	<i>cast</i>	<i>broadcast, typecast, webcast</i>	∅
	<i>dry</i>	<i>air-dry, drip-dry</i>	<i>spin-dry</i>
	<i>fill</i>	<i>backfill</i>	<i>fulfil</i>
	<i>fry</i>	<i>pan-fry, stir-fry</i>	<i>deep-fry</i>
	<i>jack</i>	<i>highjack, skyjack</i>	∅
	<i>start</i>	<i>jump-start, kick-start</i>	∅
	<i>surf</i>	<i>bodysurf, windsurf</i>	∅

As can be seen from Table 5.3, no clear picture emerges from this variable alone. Although this is consistent with previous research, which has shown that analogy is not the only factor at play, additional research using more extensive data might allow for a clearer picture on the importance of analogy in verbal compounds and how it might interact with the other factors seen previously.

In sum, we have found that the position of main stress in verbal compound can be related to lexicalisation, as compounds which are spelled as one word are more likely to get stress on their first constituent than those spelled with a hyphen. We have also found effects of syntactic categories, with compounds whose first constituent is a noun being more likely to get stress on this element than those whose first constituent has another category.

### 3.6 ‘Others’ (n=67)

As stated above, this inventory includes truncations, back-formations, neoclassical formations, bound bases compounds and some other, obscure cases:

- (12) *cónfab, liáise, vîviséct, ànimadvért, quádruple, chóreograph, génufléct, mànu fácture, flábbérgast, cáterwaul...*

Given the heterogeneity of this inventory, the stress distribution (41 on the first element, 26 on the second) is most probably of no significance.

So far, none of the categories we described is sensitive to opaque prefixation, either by nature in the case of compositional constructions or compounding, or because the parameter is superseded by stress-shifting or stress-preserving suffixes. From here on, both the weight parameter and the opaque prefixation parameter can come into play: let us take a detailed look at each sub-category.

### 3.7 Opaque Prefixation (n=1244)

We treated as ‘opaque’ any construction whose meaning is not compositional. For example, *exact* does not transparently inherit its meaning from the prefix *ex-* and the free base *act*, nor does *proclaim* transparently inherit its meaning from the prefix *pro-* and the free base *claim*.



Dabouis, Q., Fournier, J.-M. (2023) "The Stress Patterns of English Verbs: Syllable Weight and Morphology", *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.

As stated in section 2.3, this includes semantically transparent cases but whose base is not a verb, on account of category change such as the cases shown in (13).

(13) *debag, de-ice, derail, enchain, enlist, ensnare, dismast...*

As will be seen through the examples, a number of these cases are also suffixed with a stress-preserving suffix (34 as part of a free base, and 20 as part of a bound base): they do not seem to play any part into the placement of primary stress either on the prefix or on the base, which explains why they were included here.

### 3.7.1 Free Base (n=474)

The stress-pattern of the base is not affected; primary stress is on the base, as shown with the examples in (14). This includes 2 parasynthetic formations: *aggrandize, embolden*.

(14) *accústom, accóunt, becóme, besprínkle, defáult, eláipse, emblázon, expórt, preóccupy, rècolléct, redréss, ùnderstánd*

Stress is not base-initial in six words. In five of them, primary stress falls on the prefix (*cómpass, párboil, trésspass, cóunterpoise, rétrograde*), and in the last one, primary stress falls on the second syllable of the base (*àscertáin*).

### 3.7.2 Bound Base (n=770)

In 733/770 (95%) of cases, stress is on the base.

(15) *abhór, commít, distráct, ejéct, obtáin, persíst, refér, survíve*  
*cònttradíct, ìntercépt, ìntermít, rètrogréss*  
*amórtize, contínuue, demólish, elícit, inhábit, remember*

This rate is higher in disyllables, with base stress found in 643/665 (97%) of cases, than in longer verbs, with base stress found in 91/105 (87%) of cases. The words in which primary stress is not on the base are shown in (16).

(16) Disyllables: *ábseil, cómbat, cómfort, cónjure, cónquer, cóver, déstine, díffer, édit, énter, ìnjure, óffer, pérjure, próffer, próssper, rállý, rélay, rével, séver, sójourn, súffer, súmmon*  
Long verbs: *círcumcise, círcumscribe, cómpliment, cóuntenance, díscómfit, éxorcize, ìmplement, ìmportúne, ìmprovise, ìntérpret, óccupy, récognize, réconcile, súpervise*

### 3.8 Bound Base + Suffix (n=130)

In disyllables, stress is initial in 102/112 words (91%), such as those shown in (17).

(17) *chérish, chúckle, glísten, glítter, héarken, mánage, pótter, slíther, sprínkle, vánish...*

The ten words with final stress are *augmént, chastíse, dívîne, fermént, fomént, fréquént, lamént, levánt, rampáge* and *tormént*.

In longer verbs, 17/18 (94%) have antepenultimate stress. These includes 16 words in *-ize* (for example *fráternize, órganize, sánitize*) and *mónitor*. The only word with penultimate stress is *etérnize*.

Following on the observations about the *-le, -er* and *-ish* verbal suffixes in section 3.3, this class contains a number of cases whose historical suffixes have been inactive for quite a long

Dabouis, Q., Fournier, J.-M. (2023) “The Stress Patterns of English Verbs: Syllable Weight and Morphology”, *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.

time, notably *-it* (*merit, posit, visit, vomit*) and nominal *-le* (*nettle, ripple, spindle, tackle, trundle, whittle*): semantically completely opaque, and, contrary to those in section 3.3, never suffixed to a free base, they would probably be best considered as extinct and the verbs analysed as simple roots. We only kept them here for the sake of consistency at this point. This decision, however, does not affect the contrastive analysis between theories that follows since this class of verbs is grouped with the simplex class anyhow.

### 3.9 Simplex Verbs (n=293)

In disyllables, stress is initial in 262/278 words (94%), as in the words listed in (18).

(18) *árgue, cáncel, cópy, fóllow, góvern, hállow, lísten, rénder, scúrry, swállow, vómit...*

The 16 words with final stress are all listed in (19).

(19) *blasphéme, cajóle, caréen, caréss, caróuse, cavórt, equip, maráud, molést, opíne, ordáin, patról, piáffe, posséss, shampóo, usúrp*

In longer verbs, we have:

- 8 words with penultimate stress: *bamboozle, canoodle, finagle, malinge, manoeuvre, sequester, skedaddle, solicit;*
- 7 with antepenultimate stress: *damascene, gallivant, manacle, manifest, massacre, minister, orient.*

### 3.10 Interim Summary

At the end of this detailed presentation of stress and the morphological distribution of British English verbs, and before confronting both treatments of stress placement in the relevant categories, we believe it important to repeat the observation prompted by the figures in Table 5.1 above: English verbs are quite obviously morphologically characterised as heavily built ‘on the left’. The first two categories, obvious to the speaker because of semantic compositionality or the presence of a free base, represent no less than 30% of the whole class and, as such, constitute quite a strong incentive to identify the same left-edge morphology in the 22% (up to 35% in the case of disyllables) of the opaque prefixed constructions with a bound base. In such a context, hypothesising a role to opaque morphology appears rather consistent with the data.

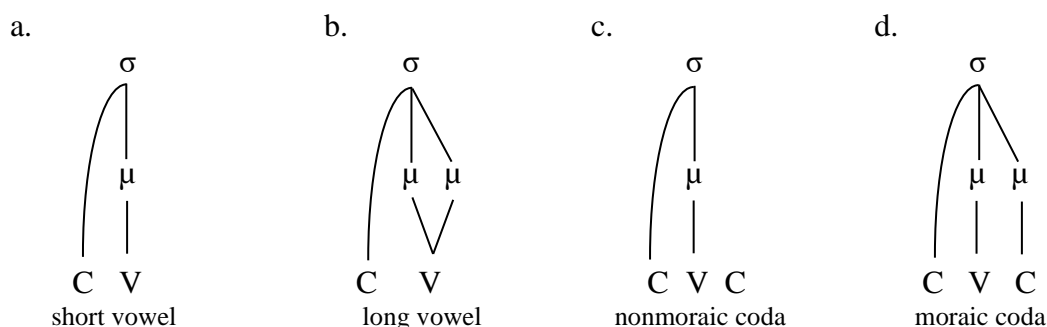
## 4 Syllable Weight and Opaque Prefixation

The aim of this section is to evaluate the scope and efficiency of the weight-based generalisation and of the Germanic Law and to determine whether both are needed to account for stress placement in English verbs. Before this evaluation, it is necessary to go through the proposals which have been made in the literature on what counts as a heavy syllable and, more specifically, what counts as a heavy syllable at the end of verbs.

### 4.1 On Syllable Weight

Nowadays, most practitioners use Moraic Theory (Hayes 1989; Hyman 1985; McCarthy and Prince 1996) to deal with weight-related phenomena. Moraic Theory claims that rhyme constituents possess something which onset constituents do not: they can project morae, that is to say weight-bearing units. The possible weight configurations in Moraic Theory are represented in (20).

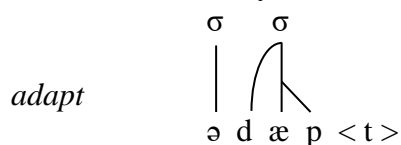
(20)



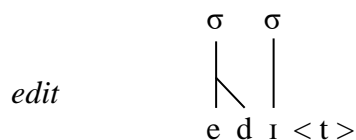
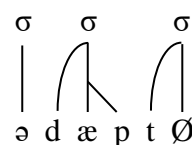
The distinction between light and heavy syllables is represented by the number of morae which syllables carry. A syllable is said to be heavy when it carries at least two morae (it is said to be bimoraic) and light when it carries only one (it is monomoraic). Short vowels project only one mora (20a), long vowels project two morae (20b) and coda consonants may (20d) or may not (20c) project a mora, depending on the language (or, as we will see below, on the nature of the coda).

In English, final syllables have been claimed to have specific properties, as is illustrated mostly by stress and by the distribution of segments (see Hammond 1999: §4). Most analyses consider the final consonant of verbs to be invisible to stress-assignment rules or constraints. This can be achieved through final consonant extrametricality (Hayes 1982), which means that that consonant is invisible to the stress-assigning mechanism. Another option is catalexis, which is the introduction of a final null vowel (Burzio 1994; Hammond 1999), which makes the final consonant the onset of a syllable with an empty nucleus.<sup>24</sup> These two options are represented in (21) with the examples of *adapt* and *edit*.

(21) a. extrametricality



b. catalexis



In both cases, the final consonant does not contribute weight to the preceding syllable. For English verbs, the claim is therefore that final syllables which contain a short vowel may be heavy only if they are followed by at least two consonants.

Moreover, there have been further proposals so as to the number of morae that a given segment may or may not project. Hammond (1999: 147) proposes the mora counts in (22), based on distributional regularities in English and on the assumption that stressed syllables should contain at least two morae and three morae at the most.<sup>25</sup>

24 See also Harris and Gussmann (1998), who argue that all final consonants should be analysed as onsets.

25 His proposal concerns American English. We are here using the British equivalents. Hammond also gives additional constraints which may render coronals heavy depending on the moraicity of the neighbouring consonants which are not relevant here.

(22)	<b>Consonants</b>		<b>Vowels</b> <sup>26</sup>	
	Coronals	(μ) <sup>27</sup>	Lax	μ
	Noncoronals	μ	Tense	μμ
	[ʒ, ɲ]	μμ	[aʊ, ɔɪ]	μμμ
	[ð]	∅		

Additionally, it has been proposed that coda sonorants have variable weight (Burzio 1994: §3.4; Giegerich 1999: 371; Halle and Vergnaud 1987: 257; Selkirk 1984: 127), which we will assume to be the variable projection of a mora, as Hammond proposes for coronals. Therefore, these more recent proposals suggest that stress placement in verbs is not simply determined by the number of consonants following the last vowel but also by the nature of these consonants.

Finally, certain vowels have been argued to be lighter than can be observed on the surface. For instance, [ə] has been argued to be less heavy than full vowels (Burzio 2007) or even nonmoraic (Hammond 1999: 206-207).<sup>28</sup> Another case is that of unstressed final [əʊ],<sup>29</sup> which has been argued to be underlyingly lax (*SPE*: 75; Hammond 1999) and, therefore, which can be interpreted as being monomoraic. Hammond (1999) also extends that claim to final [i].<sup>30</sup>

Consequently, a lot of potential drawbacks for the weight-based analysis can be excluded on the grounds that the surface weight is not the actual weight. In order to test the weight-based generalisation in its most favourable conditions, the following assumptions were adopted to code the weight of the final syllables in our dataset:

- [ə] is moraless
- Final unstressed [i] and [əʊ] are monomoraic (but final stressed [əʊ] is treated along the other long/tense vowels)
- Short/lax vowels are monomoraic
- Long/tense vowels are bimoraic, except [aʊ, ɔɪ], which are trimoraic
- The last consonant is moraless
- Preconsonantly, sonorants and [s] are moraless if unstressed and have one mora if stressed, to the exception of [ɲ], which is bimoraic
- Preconsonantly, other consonants are treated as monomoraic

This set of assumptions gives us the weights in Table 5.4. Syllables which are at least bimoraic were treated as heavy and are shown in grey. An empty cell means that the configuration is not attested in the dataset.

Table 5.4. Weight coding adopted

26 In British English, the lax vowels are [æ, e, ɪ, ʊ, ɒ, ʌ] and the reduced vowels [ə, i, u], the tense vowels are [aɪ, eɪ, i:, (j)u:, əʊ, ɔ:, aʊ]. In later sections, we will assume the r-coloured vowels [ɑ:, ɜ:] to be bimoraic.

27 There have been proposals to account for the particular case of /s/ in /sC/ clusters in terms of variable syllabification (tautosyllabic or heterosyllabic) because it can function as an onset word-initially (Kager 1989: 117-118). Burzio (1994: 61-62) describes the /s/ in /sC/ clusters as having variable weight.

28 Note that this creates circularity in the analysis, which is perfectly identified and claimed by Burzio (2007), who argues that it is a strength of parallel evaluation models, such as Optimality Theory (Prince and Smolensky 1993), over rule-based models: ‘stress and vowel reduction stand in a mutual dependency relation: stress is a determinant of vowel reduction (if stress, no reduction), but at the same time reduction controls syllable weight, and hence the position of stress. Such mutual dependencies are ordering paradoxes for ordered rules, and thus provide a direct argument for the parallel interaction of surface constraints.’

29 A vowel that our American colleagues are used to represent with [o] in American English.

30 A possible explanation is proposed in *SPE*: [i] would be an underlying glide /j/, which would explain the fact that words ending in -C<sub>2</sub>y have antepenultimate stress and prepenultimate stress in words such as *orthodoxy*, *téstimony* or *áuditory*.

Final consonant(s) Vowel	__#	__C#	__C[+son]C# __sC#	__[ŋ]C#	__C[-son]C#
[ə]	0μ	0μ	0μ	-	1μ
[i]	1μ	-	-	-	-
[əʊ]	1μ	-	-	-	-
[aʊ, ɔɪ]	3μ	3μ	3μ	-	-
∨̇	-	1μ	2μ <sup>31</sup>	3μ	2μ
∨̄	2μ	2μ	3μ	-	3μ

#### 4.2 Relevant Data and Results

In order to evaluate the role of syllable weight and that of opaque prefixation, we kept only the relevant data: simplex and deradical verbs. We distinguished two categories:

- Prefixed: opaque prefixed constructions with bound roots (see section 3.7.2)
- Non-prefixed: bound base + suffix (see section 3.8) and simplex verbs (see section 3.9)

These categories were selected because they are those which many practitioners would treat as ‘monomorphemic’ and so potentially free of morphological influences. The dataset thus contains 1,193 items, 1,055 disyllables and 138 longer verbs.

##### 4.2.1 Disyllables

Most of the relevant data concerns disyllabic verbs, in which the study of the position of stress with regards to the two generalisations is relatively easy: the weight-based generalisation predicts final stress in verbs with a heavy ultima while the Germanic Law predicts final stress in the presence of an opaque prefix. Note that they do not necessarily make contradictory predictions: they are actually convergent in 90% of cases, as will be seen below. The two parameters were tested in a binary logistic regression with the position of stress as the dependent variable. The results are shown in Table 5.5. The stress distribution of the data with regards to prefixation and the weight of the final syllable is shown in Figure 5.2.

Table 5.5. Binary logistic regression with morphology (prefixed vs. non-prefixed) and syllable weight (heavy vs. light) as predictor variables and the position of primary stress (initial vs. final) as the dependent variable

	95% C.I.			p-value
	Lower	OR	Higher	
MORPHOLOGY-PREFIXED	0.002	0.006	0.014	< 2e-16
FINALWEIGHT-LIGHT	74.007	180.867	558.917	< 2e-16

31 The only attested cases have final stress.

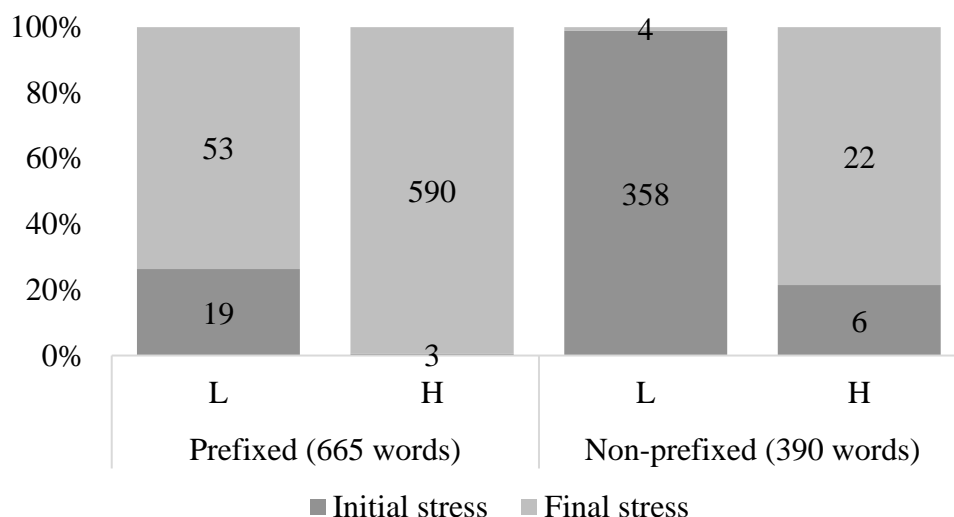


Figure 5.2. Stress distribution of disyllabic verbs depending on the weight of the final syllable (H = heavy and L = light) and of the presence or absence of an opaque prefix.

This figure shows that syllable weight is mostly inoperant in prefixed verbs, which almost systematically have stress on their base regardless of the weight of the final syllable. Crucially, the weight-based generalisation is contradicted by 74% of prefixed verbs with a light final syllable which have final stress, such as those shown in (23).

(23) *abásh, asséss, condémn, detách, emít, propél, obséss, succúmb...*

However, it should be noted that there is a statistically significant difference between prefixed verbs with a light ultima and those with a heavy ultima (Fisher’s exact test,  $p < 0.01$ ). The full inventory of prefixed words with initial stress is shown in (24).

(24) *abseil, combat, comfort, conjure, conquer, cover, destine, differ, edit, enter, injure, offer, perjure, proffer, prosper, rally, relay, revel, sever, sojourn, suffer, summon*

It should be noted that these include words for which historical prefixation is probably highly difficult to perceive because of the lack of paradigmatic support of the base, of the absence of phonotactic cues pointing to the presence of a morphological boundary or of historical change which have altered the segmental content of the prefix, making it unrecognisable.

The effect of syllable weight is clearer in non-prefixed verbs. There too, the difference between verbs with a light ultima and those with a heavy ultima is statistically significant (Fisher’s exact test,  $p < 0.01$ ).

Strikingly, the two generalisations make identical predictions in the vast majority of cases (955/1,055, 90%), that is the second and third columns in Figure 5.2, as most prefixed verbs have a heavy final syllable and most non-prefixed verbs have a light final syllable. If we evaluate the efficiency of each generalisation independently, both have an efficiency of 97%. Therefore, both can be said to be efficient descriptive generalisations of the disyllabic verbs data.

Finally, as noted above, 30% of the verb category are compositional constructions or opaque constructions with a free base: given this characteristic proportion it seems plausible to surmise

that speakers of English should be more sensitive to the presence of prefixes even in constructions with bound bases.

#### 4.2.2 Longer Verbs

As we have seen in section 3.7.2, only 14 words (13%) do not have base-initial primary stress: *círcumcise*, *círcumscribe*, *cómpliment*, *cóuntenance*, *discómfit*, *éxorcize*, *íimplement*, *íimportúne*, *ímprovise*, *intérpret*, *óccupy*, *récognize*, *réconcile*, *súpervise*. In non-prefixed words, Fournier’s (2007) Normal Stress Rule predicts that non-prefixed verbs should have antepenultimate stress. As was seen in section 3.9, this is only confirmed in seven simplex words out of 15 and in 17/18 verbs made up of a bound root and a suffix. Overall, the Germanic Law and the Normal Stress Rule correctly predict the position of primary stress in 115/138 (83%) of long verbs.

The weight-based generalisation also fails to account for the prefixed words listed above, except for *discómfit*, *íimportúne* and *intérpret*. Moreover, it fails to account for an additional six words: *amórtize*, *contínue*, *impóverish*, *íntermít*, *íntromít*, *prètermít* and *rètrogréss*. In simplex verbs, all eight words with penultimate stress are accounted for as they all have a light final syllable (*bamboozle*, *canoodle*, *finagle*, *malinge*, *manoeuvre*, *sequester*, *skedaddle*, *solicit*). However, the seven words with antepenultimate stress are unaccounted for (*damascene*, *gallivant*, *manacle*, *manifest*, *massacre*, *minister*, *orient*). In verbs made up of a bound root and a suffix, all are unaccounted for as 17 contain the heavy suffix *-ize* /-aɪz/ which never receives final stress, and the remaining word, *monitor*, has a light ultima but antepenultimate stress. Let us note that there is no attested case of non-prefixed word with a heavy ultima and final stress. Overall, the weight-based generalisation makes correct predictions in 95/138 (69%) of long verbs.<sup>32</sup>

However, it has been claimed that the suffix *-ize* is a strong retractor when it attaches to roots (see (Bermúdez-Otero, 2018))<sup>33</sup> and so one might want to exclude it from the counts. If we do so, the first approach makes correct predictions in 97/118 (82%) of cases while the weight-based generalisation does in 95/118 (80%) of cases. Therefore, as in disyllabic verbs, both generalisations have similar efficiencies.

#### 4.3 Discussion

The results presented in the previous section have shown that both generalisations have high efficiencies, especially in disyllabic verbs. Therefore, both can be said to constitute accurate descriptive generalisations of the dataset. If one were to choose one of the two generalisations, two elements should be taken into consideration: the theoretical cost and the empirical scope of both generalisations. Let us briefly go through these elements.

Positing the Germanic Law only requires one to assume that opaque prefixes are somehow recognised by speakers. Some may argue that this is already too great a cost but, as discussed in section 1.3, there is independent evidence that this is indeed the case. On the other hand, positing the weight-based generalisation in a way that allows it to reach levels of efficiency close to those of the Germanic Law requires one to accept the assumptions in 0. Some of these assumptions may be independently motivated (for example by phonotactics; see Hammond 1999) while others are more problematic (for example moraless schwa).

---

32 Let us note that the Alternating Stress Rule (see section 1.2) is no longer used in recent analyses. Most of the words to which it was applied contain stress-shifting suffixes and so their stress patterns are accounted by referring to the suffixes.

33 Although doubts can be shed on this claim. In our dataset, five root-derived words have a heavy penultimate syllable, and they are almost equally divided among penultimate and antepenultimate stress: *amórtize*, *etérnize* vs. *éxorcize*, *fráternize*, *récognize*.

Both generalisations have empirical scopes which go beyond the current dataset. The Germanic Law applies to all opaque prefixed words that are not nouns (Fournier 2010; Guierre 1979) and, more generally, allowing the phonology to see opaque prefixes is necessary for other phenomena evoked in section 1.3. The weight-based generalisation extends to other categories, notably nouns, under final syllable extrametricality (Hayes 1982), although that claim has been challenged by Fournier (2007), who claims that weight-sensitivity is only a property of certain subsets of the vocabulary: borrowings and learned words. English secondary stress assignment has also been claimed to be weight-driven (Burzio 1994; Hammond 1999; Pater 2000), but the number of relevant words is extremely small and the claim can only be supported by examples with closed syllables, not long vowels, as shown by (Dabouis et al. 2017) in an exhaustive dictionary-based study.

Therefore, one wishing to choose between the two generalisations will have to take these elements into consideration. An alternative could be to integrate both generalisations, considering that they are not necessarily contradictory. As we have seen in the previous section, they often make identical predictions. The Germanic Law should simply be stated to override the weight-based generalisation so that verbs with light final syllables get final stress. Stress assignment in verbs could then be summed up as in (25).

**(25) Stress assignment in simplex or deradical verbs**

Overlook the prefix(es) and assign stress to what remains as follows: assign primary stress to the final syllable if it is heavy and to the penultimate otherwise.

Interestingly, if one assumes a constraint-based probabilistic model such as Max-Ent-OT (Goldwater and Johnson 2003) in which both generalisations are expressed through independent constraints, such a model would predict that the two generalisations may conspire and yield higher regularity when they make the same predictions. This is precisely what was found in the data, as shown in Figure 5.2.

Even though we have seen that the two generalisations are useful in describing the distribution of stress in verbs, we have also seen that they overlap considerably. This might lead speakers to use only one of the two generalisations and overgeneralise it. There are several ways in which this could be done. Speakers could just assume the weight-based generalisation, or they could adopt a simple generalisation regarding prefixed words according to which prefixes should not get primary stress. This second option has some support from Guierre (1979), as he found that, in his dataset, 76% of prefixed disyllabic words have final stress while only 7% of non-prefixed disyllabic words do. This could explain the relatively high rate of final stress in prefixed disyllabic nouns (19% in Guierre’s dataset, as opposed to 7% for non-prefixed nouns). Another option would be for speakers to assume, as asserted by Moore-Cantwell (2020), that ‘verbs prefer final stress’. In the dataset used in section 4.2, we found final stress in 64% of disyllabic verbs and in 30% of longer verbs (36% if the verbs in *-ize* are excluded). The evidence from our study to support this last option is rather weak. Once again, if we consider Guierre’s (1979) data, 92% of disyllabic nouns have initial stress while only 20% of verbs do, but he also shows that only 9% of nouns are prefixed while 77% of verbs are. As we have argued in this paper, the explanation for the greater tendency of verbs to have final stress seems to lie partly in their morphological characteristics. Moore-Cantwell supports her assertion by quoting three papers, Sereno and Jongman (1995), Sonderegger and Niyogi (2013) and Kelly and Bock (1988). Let us briefly go through the findings of these papers and see to what extent they can be used as evidence that ‘verbs prefer final stress’.

First, let us note that all the cited papers only deal with disyllabic words and that, if any preference for final stress is to be observed for verbs, it has to be restricted to disyllabic verbs,



as found in our data. Sereno and Jongman (1995) report acoustic differences between homophonic disyllabic verb-noun pairs (for example *answer*), with a preference for trochaic patterns for nouns and iambic patterns for verbs. However, their data sample is very restricted, and their results did not reach statistical significance. Their study was recently replicated by Lohmann (2020) with more extensive data and he did not confirm their findings. Sonderegger and Niyogi (2013) report on the diachronic evolution of homographic verb-noun pairs (for example *convict*, *concrete*, *exile*). They report that most stress changes lead to pairs in which the noun has initial stress and the verb has final stress. However, they note that most pairs do not change, as long as they are in ‘stable states’ (both initial stress, both final stress or noun with initial stress and verb with final stress), the only unstable state being when the noun has final stress and the verb has initial stress. However, the vast majority of the pairs which they study are prefixed, so we can shed doubts on whether their results can be extended to non-prefixed words. Moreover, they observe that words which share a prefix tend to evolve in similar directions, which they comment as being ‘particularly interesting because they suggest that it is a shared morphological prefix rather than simply shared initial segments which correlates with trajectory similarity’. Finally, Kelly and Bock (1988) report that the rhythmic context in which words appear might explain the stress difference between nouns and verbs, as they find that nouns tend to appear more often in trochaic-biasing contexts than verbs. They also report evidence from nonword experiments that rhythmic context and syntactic category are indeed determining for the placement of stress in disyllabic pseudowords. However, there are several limitations to this study and what can be drawn from it. First, there was no control of the structure of pseudowords, and some of them could be assimilated to compounds (for example *pomset*, *ransfoe*, *solray*), prefixed words (for example *pernew*, *covact*, *premit*), while others contained final sequences which are found in real prefixed words (for example *ponset*, *rendict*, *deltain*) and yet others contained double vowels in their last syllable (for example *delpéen*, *ronvoon*, *bontoon*), which are known to be associated to final stress (see Chapter 2, this volume). It is unclear how this lack of control of the stimuli used affected the results. The second limitation of the paper is the size of the difference which Kelly and Bock report. In their second experiment, they report 75% of trochaic patterns for pseudowords in verbal contexts, against 86% for pseudowords in nominal contexts (trochaic-biasing and iambic-contexts confounded). Although the difference does go in the expected direction, nouns being more likely to be trochaic than verbs, this difference is nowhere near that found in the lexicon: the authors themselves report a dictionary study in which 94% of disyllabic nouns were found to be trochaic while only 31% of disyllabic verbs were. In her dissertation, Moore-Cantwell (2016) reports that speakers’ productions tend to reflect the probabilistic distribution of patterns found in the lexicon. The considerable gap between the dictionary data and the production data should be accounted for. Our hypothesis is that the difference can be explained because of the morphology of the stimuli used by Kelly and Bock, as only 12% of their stimuli were prefixed, and therefore we would indeed expect less final stress than what is found in the lexicon, which contains a greater proportion of prefixed words.

There are other studies which have sought to test whether disyllabic verbs do prefer final stress as compared to nouns. The results appear to confirm this hypothesis, although the effects of syllable structure were also found to be significant. However, the effects of prefixation were not tested thoroughly enough (Turcsan and Herment 2015) or not at all (Guion et al. 2003). Therefore, it is possible that speakers, or at least some of them, assume that verbs should have final stress more often than nouns, and such a tendency could be one of the factors explaining the existence of alternating noun/verb pairs such as *récord* ~ *recórd* (see Chapter 3, this volume), but in the light of the findings reported in this paper, it seems to us that prefixation should be controlled for along the other factors tested in previous studies. Moreover, the

findings reported by Kelly and Bock (1988) raise interesting research questions: could it be that the context in which individual words usually appear determine their stress patterns? Could some of the exceptions to the generalisations studied in this paper be accounted for in that way? Castanier (2016: 611) notes that this could be the case of a word like *ally* which may have seen its stress pattern shift to first-syllable stress as an effect of stress shift in expressions such as *allied forces*.

## 5 Conclusion

In this paper, we have shown that stress assignment in ‘monomorphemic’ verbs has been described through two generalisations, one referring to semantically opaque prefixation and another one referring to the weight of the final syllable. We saw that these two generalisations have known different fates in the literature, as the former has only been maintained in a specific strand of literature: the framework introduced by Guierre, while the latter has become the standard generalisation on verb stress in the generative literature. The reference to semantically opaque elements is controversial and words as *contain* or *submit* have often been treated as morphologically simplex. We have put forward arguments from phonology, phonotactics and psycholinguistics which support the idea that the constituents in these words are ‘morphologically relevant’, even though they may not be called morphemes. Consequently, we argued that it is legitimate to include semantically opaque constituents among the possible determiners of stress placement.

Therefore, we have sought to evaluate the empirical validity of the two generalisations through a large dictionary-based study on English verbs in British English. In order to narrow our study down to the relevant morphological categories, we first gave a detailed description of stress placement in verbs. The first question we have had to deal with is that of conversion, as a large proportion of our dataset was found to be verbal uses, mainly of nouns, rather than actual verbs. These were left aside from the present study and we then detailed the methods used to classify the remaining 3,550 verbs morphologically. The main striking fact which emerged from this morphological classification is the large proportion of left-edge constructions in verbs, which in itself is an argument to assume that speakers may be more sensitive to prefixation in verbs, even if it is semantically opaque.

We gave a detailed descriptive account of the stress pattern of English verbs based on our extensive dataset of 3,350 verbs. To our knowledge, this is the most detailed description to date with a systematic quantification of each category, of exceptions and of stress regularities. Then we detailed how the weight of the final syllable was coded in our data to measure the empirical validity of the two generalisations. A striking result is that the two generalisations make identical predictions for a large majority of the data and can only be compared within a relatively marginal proportion. Overall, we found that both generalisations have similar efficiencies. In order to tease apart these two generalisations, we detailed the theoretical cost of each generalisation as well as its empirical scope beyond the present dataset. We argue that the results reported in this paper constitute strong support in favour of the inclusion of semantically opaque constituents among morphologically relevant units for phonological computation. We showed that it is possible to keep both generalisations for a more accurate description of the stress pattern of English verbs. Finally, we discussed other possible analyses which speakers might make and have reviewed some of the evidence and limitations for each of the three alternatives considered.

Dabouis, Q., Fournier, J.-M. (2023) “The Stress Patterns of English Verbs: Syllable Weight and Morphology”, *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.

## References

- Abasq, V. (2007), ‘Préfixation et particules adverbiales en anglais contemporain : Etude du comportement accentuel’, PhD dissertation, Université de Tours.
- Anderson, S. R. (1992), *A-Morphous Morphology*, Cambridge: Cambridge University Press.
- Arndt-Lappe, S. (2011), ‘Towards an exemplar-based model of stress in English noun–noun compounds’, *Journal of Linguistics*, 47(3), 549–585.
- Aronoff, M. (1976), *Word Formation in Generative Grammar*, Cambridge, MA: MIT Press.
- Bauer, L., Lieber, R. and I. Plag (2013), *The Oxford reference guide to English Morphology*, Oxford: Oxford University Press.
- Bell, M. and I. Plag (2013), ‘Informativity and analogy in English compound stress’, *Word Structure*, 6(2), 129–155.
- Bermúdez-Otero, R. (2018), ‘Stratal phonology’, in S. J. Hannahs and A. R. K. Bosch (eds.), *The Routledge Handbook of Phonological Theory*, Abingdon: Routledge, pp. 100–134.
- Booij, G. and J. Rubach (1984), ‘Morphological and prosodic domains in Lexical Phonology’, *Phonology Yearbook*, 1, 1–27.
- Bram, B. (2011), ‘Major Total Conversion in English: The Question of Directionality’, PhD dissertation, Victoria University of Wellington.
- Burzio, L. (1994), *Principles of English Stress*, New York: Cambridge University Press.
- Burzio, L. (2007), ‘Phonology and phonetics of English stress and vowel reduction’, *Language Sciences*, 29(2–3), 154–176.
- Carr, P. (1999), *English Phonetics and Phonology*, Oxford: Basil Blackwell.
- Castairs-McCarty, A. (2002), *An Introduction to English Morphology: Words and Their Structure*, Edinburgh: Edinburgh University Press.
- Castanier, J. (2016), ‘L’évolution accentuelle du lexique anglais contemporain appréhendée à travers les dictionnaires de prononciation (XVIIIe - XXIe siècles)’, PhD dissertation, Université de Poitiers.
- Chomsky, N. and M. Halle (1968), *The Sound Pattern of English*, New York: Harper & Row.
- Collie, S. (2007), ‘English stress preservation and Stratal Optimality Theory’, PhD dissertation, University of Edinburgh.
- Cruttenden, A. (2014), *Gimson’s Pronunciation of English*, 8th edn, Oxon: Routledge.
- Dabouis, Q. (2016), ‘L’accent secondaire en anglais britannique contemporain’, PhD dissertation, University of Tours.
- Dabouis, Q. and J.-M. Fournier (2019), ‘On the role of morphology, syllable structure, frequency and spelling in English vowel reduction’, poster presented at the 27th MFM Conference held at the University of Manchester on 23-25 June 2019.
- Dabouis, Q., J.-M. Fournier and I. Girard (2017), ‘Ternarity is not an issue: Secondary stress is left edge marking’, paper presented at the MFM25 Fringe Meeting – PTA Dataset Workshop at the University of Manchester.
- Delbridge, A. et al. (eds.) (2009), *Macquarie Dictionary, Australia’s National Dictionary Online*, Macquarie Library.
- Domahs, U., I. Plag and R. Carroll (2014), ‘Word stress assignment in German, English and Dutch: Quantity-sensitivity and extrametricality revisited’, *The Journal of Comparative Germanic Linguistics*, 17(1), 59–96.
- Fournier, J.-M. (1996), ‘La reconnaissance morphologique’, in *8ème Colloque d’Avril sur l’anglais oral*, Villetaneuse: Université de Paris-Nord, CELDA, diffusion APLV, 45–75.
- Fournier, J.-M. (2007), ‘From a Latin syllable-driven stress system to a Romance versus Germanic morphology-driven dynamics: In honour of Lionel Guierre’, *Language Sciences*, 29, 218–236.
- Fournier, J.-M. (2010), *Manuel d’anglais oral*, Paris: Ophrys.

- Dabouis, Q., Fournier, J.-M. (2023) "The Stress Patterns of English Verbs: Syllable Weight and Morphology", *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.
- Fudge, E. (1984), *English Word Stress*, London: G. Allen and Unwin.
- Giegerich, H. J. (1999), *Lexical Strata in English: Morphological Causes, Phonological Effects*, Cambridge: Cambridge University Press.
- Giegerich, H. J. (2009), 'The English compound stress myth', *Word Structure*, 2(1), 1–17.
- Goldwater, S. and M. Johnson (2003), 'Learning OT constraint rankings using a maximum entropy model', in J. Spenser, A. Eriksson and Ö. Dahl (eds.), *Proceedings of the Workshop on Variation within Optimality Theory*, Stockholm: Stockholm University, pp. 111–120.
- Guierre, L. (1979), 'Essai sur l'accentuation en anglais contemporain : Eléments pour une synthèse', PhD dissertation, Université Paris-VII.
- Guierre, L. (1984), *Drills in English Stress Patterns: Ear and Speech Training Drills and Tests for Students of English as a Foreign Language*, 4th edn, Londres: A. Colin - Longman.
- Guierre, L. (1990), 'Mots composés anglais et agrégats consonantiques', in *5ème Colloque d'Avril sur l'anglais oral*, Villetaneuse: Université de Paris-Nord, CELDA, diffusion APLV, 59–72.
- Guion, S. G. et al. (2003), 'Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words.', *Language and speech*, 46(Pt 4), 403–427.
- Halle, M. and S. Keyser (1971), *English Stress: Its Form, Its Growth, and its Role in Verse*, New York: Harper & Row.
- Halle, M. and J.-R. Vergnaud (1987), *An Essay on Stress*, Cambridge, MA: MIT.
- Hammond, M. (1999), *The Phonology of English: A Prosodic Optimality-Theoretic Approach*, Oxford: Oxford University Press.
- Hammond, M. (2003), 'Frequency, cyclicity, and optimality', University of Arizona, Available at: <http://www.u.arizona.edu/~hammond/kslides.pdf> (Accessed: 2 September 2014).
- Harris, J. and E. Gussmann (1998), 'Final codas: Why the west was wrong.', in Cyran, E. (ed.), *Structure and Interpretation: Studies in Phonology*, Lublin: Folium, pp. 139–162.
- Hayes, B. (1982), 'Extrametricity and English stress', *Linguistic Inquiry*, 13(2), 227–276.
- Hayes, B. (1989), 'Compensatory Lengthening in Moraic Phonology', *Linguistic Inquiry*, 20, 253–306.
- Ben Hedia, S. and I. Plag (2017), 'Gemination and degemination in English prefixation: Phonetic evidence for morphological organization', *Journal of Phonetics*, 62, 34–49.
- Van Heuven, W. V. J. et al. (2014), 'Subtlex-UK: A new and improved word frequency database for British English', *Quarterly Journal of Experimental Psychology*, (67), 1176–1190.
- Hyman, L. (1985), *A Theory of Phonological Weight*, Dordrecht: Foris.
- Jones, D. (2006), *Cambridge English Pronouncing Dictionary*, 17th edn, Cambridge: Cambridge University Press.
- Kager, R. (1989), 'A metrical theory of stress and destressing in English and Dutch', PhD dissertation, University of Utrecht.
- Kahn, D. (1976), 'Syllable-based generalizations in English phonology', PhD dissertation, MIT.
- Katamba, F. and J. Stonham (2006), *Morphology*, Basingstoke, New York: Palgrave Macmillan.
- Kelly, M. H. and J. K. Bock (1988), 'Stress in time', *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 389–403.
- Kiparsky, P. (1982), 'From Cyclic Phonology to Lexical Phonology', in H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations I*, Dordrecht: Foris, pp. 131–175.

- Dabouis, Q., Fournier, J.-M. (2023) “The Stress Patterns of English Verbs: Syllable Weight and Morphology”, *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.
- Ktori, M. et al. (2016), ‘Prefixes repel stress in reading aloud: Evidence from surface dyslexia’, *Cortex*, 74(October 2015), 191–205.
- Ktori, M., P. Mousikou and K. Rastle (2018), ‘Cues to stress assignment in reading aloud’, *Journal of Experimental Psychology*, 147(1), 36–61.
- Lieberman, M. and A. Prince (1977), ‘On stress and linguistic rhythm’, *Linguistic Inquiry*, 8(2), 249–336.
- Lieber, R. (2004), *Morphology and Lexical Semantics*, Cambridge: Cambridge University Press.
- Lohmann, A. (2020), ‘No acoustic correlates of grammatical class: A critical re-examination of Sereno and Jongman (1995)’, *Phonetica*, 77(6), 429–440.
- Marchand, H. (1969), *The Categories and Types of Present-Day English Word-Formation*, Munich: C.H. Beck.
- Marslen-Wilson, W. D., M. Bozic and B. Randall (2008), ‘Early decomposition in visual word recognition: Dissociating morphology, form, and meaning’, *Language and Cognitive Processes*, 23(3), 394–421.
- McCarthy, J. J. and A. S. Prince (1996), ‘Prosodic Morphology 1986’, *Linguistics Department Faculty Publication Series*, University of Massachusetts, Amherst, 13.
- McMahon, A. (2001), ‘Review *The Phonology of English: A Prosodic Optimality-Theoretic Approach* by of Hammond (1999), Oxford: Oxford University Press’, *Phonology*, 18, 421–426.
- Moore-Cantwell, C. (2016), ‘The representation of probabilistic phonological patterns : Neurological , behavioral , and computational evidence from the English stress system’, PhD dissertation, University of Massachusetts Amherst.
- Moore-Cantwell, C. (2020), ‘Weight and final vowels in the English stress system’, *Phonology*, 37(4), 657–695.
- Mudgan, J. (2015), ‘15. Units of word-formation’, in Müller, P. O. et al. (eds.), *Volume 1: Word-Formation: An International Handbook of the Languages of Europe*, Berlin, Boston: Mouton de Gruyter, pp. 235–301.
- Pastizzo, M. J. and L. B. Feldman (2004), ‘Morphological processing: A comparison between free and bound stem facilitation’, *Brain and Language*, 90(1–3), 31–39.
- Pater, J. (2000), ‘Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints’, *Phonology*, 17, 237–274.
- Plag, I. (1999), *Morphological Productivity: Structural Constraints in English Derivation*, Berlin: Mouton de Gruyter.
- Plag, I. (2003), *Word-Formation in English*, Cambridge: Cambridge University Press.
- Plag, I. (2006), ‘The variability of compound stress in English: Structural, semantic and analogical factors’, *English Language and Linguistics*, 1(10), 143–172.
- Plag, I. et al. (2008), ‘The role of semantics, argument structure, and lexicalization in compound stress assignment in English’, *Language*, 84(4), 760–794.
- Plag, I., Kunter, G. and S. Lappe (2007), ‘Testing hypotheses about compound stress assignment in English: A corpus-based investigation’, *Corpus Linguistics and Linguistic Theory*, 2(3), 199–232.
- Prince, A. and P. Smolensky (1993), ‘Optimality Theory: Constraint interaction in Generative Grammar’, Manuscript, Rutgers University and University of Colorado.
- Raffelsiefen, R. (1993), ‘Relating words: A model of base recognition, Part I’, *Linguistic Analysis*, (23), 3–161.
- Raffelsiefen, R. (1999), ‘Diagnostics for prosodic words revisited: The case of historically prefixed words in English’, in T. A. Hall and U. Kleinhenz (eds.), *Studies on the Phonological Word*, Current Issues in Linguistic Theory 174, Amsterdam: John

Dabouis, Q., Fournier, J.-M. (2023) “The Stress Patterns of English Verbs: Syllable Weight and Morphology”, *New Perspectives on English Word Stress* (Ballier, N., Fournier, J.-M., Przewozny, A., Yamada, E., eds.), Edinburgh : Edinburgh University Press, 154-191.

- Benjamins Publishing, pp. 133–201.
- Raffelsiefen, R. (2007), ‘Morphological word structure in English and Swedish: The evidence from prosody’, *Fifth Mediterranean Morphology Meeting*, pp. 209–268.
- Rastle, K. and M. Coltheart (2000), ‘Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords’, *Journal of Memory and Language*, 42, 342–364.
- Rastle, K. and M. H. Davis (2008), ‘Morphological decomposition based on the analysis of orthography’, *Language and Cognitive Processes*, 23(7–8), 942–971.
- Roach, P. (2000), *English Phonetics and Phonology: A Practical Course*, 3rd edn, Cambridge: Cambridge University Press.
- Selkirk, E. O. (1980), ‘The role of prosodic categories in English word stress’, *Linguistic inquiry*, 11(3), 563–605.
- Selkirk, E. O. (1982), *The Syntax of Words*, Cambridge, MA: MIT Press.
- Selkirk, E. O. (1984), *Phonology and Syntax: The Relation between Sound and Structure*, *Current studies in linguistics series*, Cambridge: MIT Press.
- Sereno, J. A. and A. Jongman (1995), ‘Acoustic correlates of grammatical class’, *Language and Speech*, 38(1), 57–76.
- Siegel, D. C. (1974), ‘Topics in English morphology’, PhD dissertation, MIT.
- Siegel, D. C. (1980), ‘Why there is no = boundary’, in M. Aronoff and M.-L. Kean (eds.), *Juncture*, Saratoga: Anma Libri, pp. 131–134.
- Sonderegger, M. and P. Niyogi (2013), ‘Variation and change in English noun/verb pair stress: Data and dynamical systems models’, in A. C. L. Yu (ed.), *Origins of Sound Change*, Oxford: Oxford University Press, 262–284.
- Stockwell, R. and D. Minkova (2001), *English Words: History and Structure*, Cambridge: Cambridge University Press.
- Szpyra, J. (1989), *The Morphology-Phonology Interface: Cycles, Levels and Words*, London: Routledge.
- Taft, M. (1994), ‘Interactive-activation as a framework for understanding morphological processing’, *Language and Cognitive Processes*, 9(3), 271–294.
- Treiman, R. et al. (2020), ‘Cues to stress in English spelling’, *Journal of Memory and Language*, 112.
- Trevian, I. (2003), *Morphoaccentologie et processus d’affixation de l’anglais*, Bern: Peter Lang.
- Trevian, I. (2007), ‘Stress-neutral endings in contemporary British English: An updated overview’, *Language Sciences*, 29(2–3), 426–450.
- Trevian, I. (2015), *English Suffixes: Stress-assignment Properties, Productivity, Selection and Combinatorial Processes*, Bern: Peter Lang.
- Turcsan, G. and S. Herment (2015), ‘Making sense of nonce word stress in English.’, in J. A. Mompean and J. Fouz (eds.), *Investigating English Pronunciation: Current Trends and Directions*, Hampshire, UK: Palgrave Macmillan, pp. 23–46.
- Wells, J. C. (2008), *Longman Pronunciation Dictionary*, 3rd edn, London: Longman.
- Wennerstrom, A. (1993), ‘Focus on the prefix: Evidence for word-internal prosodic words’, *Phonology*, 10(2), 309–324.
- Wenszky, N. (2004), *Secondary Stress in English Words*, Budapest: Akadémiai Kiadó.