Morphology (2026) 36:8
https://doi.org/10.1007/s11525-026-09457-6

Modeling the role of prefixation in determining stress assignment in English verbs

Ingo Plag[1] · Sabine Arndt-Lappe[2] · Aaron Seiler[2] · Quentin Dabouis[3]

Received: 3 June 2025 / Accepted: 19 January 2026

Abstract
Most accounts of English stress assume that it is determined by a combination of morphological and phonological factors. Two main factors have been proposed to determine the patterns of observed stress in English verbs with no productive morphology: whether the verb is prefixed, and the weight of the final syllable. This paper tests whether stress assignment to English verbs can be accounted for without recourse to abstract morphological properties and whether, instead, alleged morphological effects emerge as a function of the statistical patterning of stress among recurrent forms in the lexicon. The paper implements a discriminative learning network (NDL), a well-established computational learning model that is trained using error-driven learning (Rescorla & Wagner, 1972; Baayen et al., 2011, see Arndt-Lappe et al., 2023 for application to stress). The model, which uses 1,948 verb types from Jones' pronunciation dictionary (Roach et al., 2006), works with phonotactic information only and does not have access to higher-level information about the morphological structure of words. The NDL model reaches the same level of accuracy as competing regression models that use the traditional predictors (syllable number, weight, morphological status of initial string). Closer inspection reveals that the NDL model brings about the alleged morphological and weight effects by a gradient association of phonotactics and stress. In the network, which has no explicit information about morphological structure, morphological effects emerge as a function of the recurrence of certain phonological strings. At a higher level, the results speak in favour of theories in which stress (and effects of weight and morphology) emerge from dynamically changing associations between representations in the lexicon.

Keywords Stress · Verbs · English · Prefixation · Discriminative learning · Morphology

1 Introduction

Stress in English verbs is usually assumed to be influenced by the weight of the final syllable and by transparent, productive morphology. However, since early generative

Extended author information available on the last page of the article

Published online: 02 February 2026

phonology, there is also another factor that has been proposed and that sometimes seems to prevail over other effects: opaque prefixation. Historically prefixed verbs like *contain, deduce* or *submit* have stress on their root, even in cases in which the final syllable is light. However, the formatives involved in those words are far from being consensually recognized by all morphologists, notably because they have little or obscure semantics (see discussion in Dabouis & Fournier, 2025).

In this paper we test the weight-based and morphology-based generalizations and predictions just described against a very large data set of 1,948 verbs, gleaned from Jones' *Pronouncing Dictionary* (Roach et al., 2006). In particular, we would like to understand the nature of the alleged prefix effects. Assigning prefix status to a string is sometimes straightforward (as in, for example, *un-natural*) but very often controversial (e.g. *re-legate*, *per-mit*), especially in those cases where the remaining part of the word is a bound root, not a word. Certainly those controversial prefix candidates seem to be genuine pieces of word structure as can be seen by several of their properties:

- They may be found in several, sometimes many, different words (e.g. *tain* in *attain, contain, detain, obtain*, *se-* in *select, seclude, secede, segregate*).
- They may show phonotactic sequences that are otherwise only attested in clearly morphologically complex words (e.g. /ksp/ in *expand, expire, explode* cp. *backspace, taxpayer, workspace*; see Hammond, 1999, §3.3, Guierre, 1990).
- Words that share an apparent constituent undergo similar changes when suffixed (e.g. *admit, permit, transmit* - *admission, permission, transmission*; Aronoff, 1976, pp. 12-15, Plag, 2018, pp. 24-26).

However, in the morphological literature, the problems with the alleged prefix effects are frequently discussed in the context of how to define the notion of 'morpheme', which is standardly defined as the smallest meaningful unit in morphological analyses (cf. Plag, 2018). The challenge is that the meaning of the units in question is often not clear, which may make the assumption of a morpheme problematic. Morphologists have found a way out of these terminological problems by positing 'formatives', "an overarching category that includes both morphemes and elements contributing to the construction of words whose semantic unity or function is obscure or dubious" (Bauer et al., 2013). Definitions of prefixal formatives are then often based on etymology (Dabouis & Fournier, 2023).

In this paper we explore the possibility that the alleged prefix effect is an effect of recurrent forms, irrespective of the question of whether there is some 'prefixal' meaning or etymology involved or not.

We will implement two kinds of analyses. The first is a regression tree analysis that uses traditional phonological predictors and morphological information to predict verb stress. The phonological predictors encode well-established aspects of syllabic structure (number of syllables, syllable weight), and the morphological predictor encodes whether the beginning of the word is

- a prefix (either transparent or opaque), or
- a string that is homophonous to a prefix, but is no prefix, or
- a string that is not homophonous with any prefix of English.

This first analysis will set the baseline against which we will test our alternative, recurrence-based hypothesis. A methodological challenge lies in distinguishing 'recurrent elements' from 'prefixes', especially 'opaque' prefixes. To eliminate any potential bias of our own concerning the morphological classification we used the coding provided by Wikimorph (Yarbro & Olney, 2021), a deep-learning-based tool for word segmentation that employs word embeddings for its classification. Crucially for us, Wikimorph takes into account semantic and etymological information when deciding on segmentation. The former is important to recognise transparent prefixes, the latter is important to recognise opaque prefixes.

As an alternative to traditional predictors we implement a discriminative learning model that uses only orthographic bigrams (i.e. sequences of two letters, and sequences of one letter and a word boundary symbol) as predictors of stress and does not have access to higher-level information about the morphological structure of words. Naive discriminative learning (NDL, Rescorla & Wagner, 1972; Baayen et al., 2011) is a well-established and computationally implemented learning theory developed in cognitive psychology that builds associations between different representations or events ('cues' and 'outcomes') through their co-occurrence or non-co-occurrence in learning events. In our implementation, the model works with phonotactic information as cues, and stress position as outcomes. For our analyses we use data from Jones' (Roach et al., 2006) pronunciation dictionary, with 1948 verb types.

Both kinds of models reach comparable, very high accuracy rates. And both models demonstrate that it is mainly the recurrence of phonological strings that brings about what has been interpreted as morphological effects.

A closer analysis of the association measures extracted from the NDL model reveals that differences between the encoded phonological and morphological categories are reflected in the activation profiles of pertinent words: the NDL model shows morphological and weight effects without having explicit information on these parameters. These effects emerge from a gradient association of phonotactic information (i.e. bigrams) and stress. At a higher level, these results speak in favour of usage-based theories of phonology (e.g. Bybee, 2006, 2010), in which stress (and, in our case, effects of weights and morphology) emerge from representations and their relations in the lexicon (Daelemans et al., 1994; Eddington, 2000; Arndt-Lappe et al., 2023).

The paper is structured as follows. In Sect. 2, we first give an overview of previous approaches to verb stress in English and discern open questions. This is followed by Sect. 3, in which we introduce the reader to discriminative learning and the NDL model used in this paper. In Sect. 4 we present our data and explain the different kinds of analyses we employ. The results are given in Sect. 5, which is followed by a discussion and conclusion in Sect. 6.

2 Verb stress and prefixation

In the generative literature since Chomsky and Halle (1968), it has been assumed that English stress is weight-sensitive. In those models, primary stress in verbs is assigned from the end of the word to the final syllable if it is heavy and to the penult

Table 1 Effects of syllable weight on stress according to the generative literature (angled brackets indicate extrametrical material)

Weight of relevant syllable	Nouns	Verbs
Heavy	*agén⟨da⟩, horí⟨zon⟩*	*cajó⟨le⟩, tormén⟨t⟩*
Light	*Améri⟨ca⟩, dísci⟨pline⟩*	*bóthe⟨r⟩, vómi⟨t⟩*

if the final syllable is not heavy. However, syntactic category also plays a role, as verbs and nouns pattern differently: for this weight-sensitive generalization to generate the correct outcome, it is necessary to ignore the final syllable of nouns and the final consonant of verbs. This has been formalized using extrametricality in Metrical Phonology (Hayes, 1982), and can be illustrated with the examples shown in Table 1.

However, in Chomsky and Halle (1968), it was shown that this generalization has two limitations. First, if the verb contains a historical prefix, stress falls on the final syllable regardless of its weight, i.e. even if the final syllable is light (e.g. *attách, commít, repél*). Second, in trisyllabic verbs with a heavy final syllable, stress actually usually falls on the antepenultimate syllable. This concerns suffixed verbs such as *cómplicate* or *clárify* but also simplex verbs such as *gállivant* or *mánifest*. For trisyllabic verbs, historical prefixation has been argued to interfere with weight-based stress assignment as verbs whose final syllable is a root have final stress, once again regardless of weight (e.g. *com-pre-hénd, inter-mít, intro-dúce*). Fournier (2007) takes a different approach and proposes that the stress behaviour of prefixed verbs should be related to Germanic root-initial stress assignment. He puts forward the generalization that prefixed verbs have stress on their root and that this does not interact with weight. Like weight-based stress assignment, this generalization applies differently to nouns, which display only small, though significant, differences between historically prefixed words and simplex words (Dabouis & Fournier, 2025; Guierre, 1979).

As noted by Dabouis and Fournier (2023), this second generalization has progressively disappeared from the mainstream generative literature, quite likely because the formatives involved in the construction of historically prefixed words do not fit standard assumptions about morphemehood. This is due to the fact that those formatives often have at best a rather vague meaning that can be found across the words that contain them. Dabouis and Fournier (2025) take on this issue, reviewing the available evidence that can allow one to determine whether words such as *attain, compel* or *submit* should be treated as simplex words. They discuss phonological, morphological and psycholinguistic evidence, and argue that such words cannot be treated as simplex words, although they differ from words with semantically transparent prefixes (e.g. *disconnect, misspell, rewrite*). Following previous proposals by Forster and Azuma (2000), Fournier (1996), Taft (1994), Wurm (1997), and Wurm (2000), they propose that it is the distributional recurrence of prefixes and roots that allows for their identification and makes them learnable despite their semantic opacity. They also show that many such historically prefixed words are not entirely opaque and share semantics on a sometimes quite abstract level such as "skeletal thematic information" (Baeskow, 2006). Similar proposals have been made for historically complex proper names in Dutch (Köhnlein, 2015) and English (Dabouis, 2023) (e.g. *Brix**ton**, Fleming**ton**, Shep**ton**, Washing**ton***).

All this raises a number of crucial questions in morphology and morpho-phonology, such as: What makes up a formative? Can it be identified in the absence of clear semantics, and given that language users do not know about their etymological status? If so, does this have to do with the recurrence of strings in the lexicon? Do other kinds of information such as phonotactics play a role?

In this paper, we use naive discriminative learning (NDL, Baayen et al., 2011), a computationally implemented learning model, to test to what extent stress patterns can be successfully captured without recourse to higher-level information about morphological structure. As a baseline, we compare this model with a statistical model that uses higher-order predictors (esp. syllable weight and morphological structure) on the same, comprehensive data set (data from Daniel Jones' pronunciation dictionary, Roach et al., 2006). If the NDL model is successful, this will provide evidence that the 'morphological' effects in verb stress are in fact learnable from recurrence only. A conceptually straightforward interpretation of this is that the morphological effects observed in verb stress indeed emerge from gradient associations between recurrent strings and certain stress patterns in the Mental Lexicon, in line with lexicon-usage-based models of morphology.

3 Discriminative learning

Discriminative learning theory is a well-established theory of learning in cognitive psychology (e.g. Rescorla, 1988a; Pearce & Bouton, 2001). The general cognitive mechanisms assumed in this theory have been shown to be able to model a number of important effects observed in both animal and human learning, for example the blocking effect (Kamin, 1969) and the feature-label ordering effect (Ramscar et al., 2010).

Discriminative learning has found its way into linguistics through the work of Harald Baayen and his associates, who have also developed computational implementations, called naive discriminative learning (NDL, e.g. Baayen et al., 2011) and linear discriminative learning (LDL, e.g. Baayen et al., 2019). Introductions to NDL have made their way into morphology textbooks (see Plag, 2018; Lieber, 2021), and applications of this approach to phenomena in different languages are numerous. Discriminative learning theory is based on the idea that learning results from exposure to informative relations among events in the environment. These relations are modeled as 'association weights'. The events associated with each other are called 'cues' and 'outcomes' and the associations between cues and outcomes are constantly updated. They increase if cues and outcomes are both present in the learning event, and decrease if the cues are present, but the outcome is not. The decrease of association weights in the case of the non-presence of an outcome also means that, in this framework, learning is partly error-driven. Mathematically, the association weights between cues and outcomes are calculated using the Rescorla-Wagner equations (Rescorla & Wagner, 1972; Rescorla, 1988a,b). The final association weights, reached at the end-state of the learning process, can also be directly computed by using the Danks equilibrium equations (Danks, 2003).

Discriminative networks have been implemented to model a wide range of phenomena, such as the prediction of inflectional and derivational forms (Baayen et al., 2018, 2019; Tomaschek & Ramscar, 2022; van de Vijver & Uwambayinema, 2022), the acoustic duration of suffixes and complex words, including nonce words (Stein & Plag, 2021; Schmitz et al., 2021), the articulatory movements across morphological boundaries (Saito et al., 2020), morphological disorders (Heitmeier et al., 2021) and participants' reaction times and decisions in experiments investigating morphological processing (Baayen et al., 2011; Plag et al., 2023).

An NDL-based analysis of stress assignment and stratification in English was proposed by Arndt-Lappe et al. (2023), who demonstrated that "neither directionality nor stratification need to be stipulated as a-priori properties of words or constraints in the lexicon" (p. 438).

With regard to the present research questions, an NDL network that is based on bigrams (or trigrams) can be used as a radical alternative to approaches that use rules or constraints that make reference to structural phonological and morphological information. An NDL network can be conceived as a part of a lexicon that is not a repository of lexical entries, but a lexicon in which lexical knowledge is ever-changing, distributed information concerning form and meaning, from which phenomena such as 'word' or 'morphological structure' are emergent and ephemeral.

4 Methodology

4.1 Data and coding

The data used in this study comes from Jones' *Pronouncing Dictionary* (Roach et al., 2006).[1] The dictionary has a total of 5,236 verbs for which it provides IPA-based transcriptions, along with their stress pattern. We started out with a cleaned file that excludes all verbs that were marked as "rare", "obsolete", or "non-British", as well as verbs that were treated as conversions (either from nouns or adjectives).

The reason for excluding converted verbs is that denominal and deadjectival verbs largely preserve the stress pattern of the source category and thus differ from non-converted verbs. Traditional descriptions of English stress give different generalizations depending on syntactic categories, and those generalizations are often implicitly restricted to monomorphemic words. Here, we are adding some complexity by integrating some morphologically complex words, and so including converted verbs would add further complexity to the data and would make the results more difficult to relate to the existing literature. Moreover, the data set used here contains a great number of converted verbs because it got its categorial information from the Macquarie Dictionary Delbridge et al. (2009), in which many items are tagged as having several categories, with some having only marginal verbal uses (e.g. *baby, patent, overnight, tapestry*). Out of these considerations, converted items were excluded so as to study the generalizations regulating stress placement in the core sublexicon of verbs, without interferences from other syntactic categories.[2] The cleaning steps are described in Dabouis and Fournier (2023).

[1] The data and analysis scripts for this study are available at https://osf.io/e6vc8.

[2] Admittedly, our approach here simplifies matters by assuming that syntactic categories are quite separate from one another. This might be argued to make predictions easier for any type of formal analysis (be it

From this file, we excluded all variants of pronunciation and also those variants that do not show in the spelling. That is, in cases where the dictionary gives two or more pronunciation variants, we settled on the variant that was marked as the "main pronunciation" in British English. After cleaning, the data set contained 2,311 verbs (N = 2,311).

Primary stress position was extracted from the dictionary and coded from the right edge of the word (following traditional descriptions) as "preante" for preantepenult stress, "ante" for antepenult stress, "pen" for penultimate stress, and "fin" for final stress, respectively. From now on, this set will also be referred to as the "full" data set. For different kinds of analyses we derived other data sets from the full data set (see below). Being the largest possible data set, it was used to train the NDL network. All other data sets were derived from this data set (see below).

The coding of morphology is a notorious problem. As discussed in Sect. 2, morphological complexity is gradient, morphological boundaries may vary in their strengths, and it is not always clear whether a given word should be classified as morphologically simplex or complex. Morpheme-based approaches to morphological complexity face the problem of setting up clear criteria for assuming the presence of an affix. Well-known problematic cases include historically Latinate prefixes such as *ad-, con-, in-, re-*. These prefixes sometimes have a clearly discernible meaning, sometimes not, so that it is not always clear whether one should assume the presence of a prefix or not. Depending on the researcher's definition of morpheme, the pertinent strings can be either treated as being morphologically relevant or not (see, for example, Plag, 2018, Sect. 2.1 for further discussion). The different approaches to determining morphological complexity often lead to different classifications of the same data, with the consequence that proposed morphological effects may appear controversial to begin with.

In order to avoid introducing our own bias into the classification of verbs as being either prefixed or simplex, we used the morphological classification as found in the database Wikimorph (Yarbro & Olney, 2021), a large-scale database that comprises around 450,000 entries derived from a deep learning-based model. Wikimorph is particularly useful for our purposes as it is the only non-lexicographic source to our knowledge that provides etymological segmentations, which are especially useful for the investigation of opaque prefixes.[3] For each keyword in the database, Wikimorph provides a morphological segmentation, along with etymological information about the decomposed units. Our coding of morphological structure proceeded as follows: First, we extracted all verbs from the Wikimorph database that are also contained in the full (i.e., the dictionary) data set. This reduced the data set to 1948 verb types (which will be termed the 'Wikimorph' data set).

traditional-structural or NDL-based) than for the language users, who may not always be certain which verbs are converted from nouns. Future research might consider including everything that is verbal in nature or even the whole lexicon (though the latter has already been explored in Arndt-Lappe et al. (2023), with different research questions).

[3] The Wikimorph data set is based on the English Wiktionary XML dump file. The data set contains morphemes (from English and root languages), part-of-speech tags, and definitions for each morpheme. The authors also trained a GPT-2 model on this data set to morphologically break down any English word. See Yarbro and Olney (2021) for further details.

In a second step, each verb was assigned to one of four prefixal categories based on the morphological and etymological information provided by Wikimorph: verbs were classified as `transparent` if Wikimorph provided a morphological boundary for the prefix and the base was a free form. Examples are *be-little*, *en-force* and *re-fill*. If, on the other hand, Wikimorph provides a morphological boundary for the prefix, but the base does not correspond to a free form, the item was coded as `opaque` - examples for the latter category include *ab-duct*, *con-fer* and *re-legate*. Thus, when we talk about 'transparent' and 'opaque' prefixes in the remainder of this paper, it is important to bear in mind that our definition of 'transparency' is in fact based on the morphological status of both the prefix and the base. We deliberately employ a more restrictive definition than some of the literature which refers to the morphological status of the prefix only. The reason is that a less restrictive transparency definition is notoriously difficult to operationalize (compare, e.g., the potential contribution of *de-* in *deduce, demolish, determine*).

Finally, in case Wikimorph does not provide a morphological boundary for the prefix but the initial string is one that is commonly found in prefixed words, the verb was coded as `no_prefix`, as in *alarm, second* or *realize*. The fourth category comprises verbs whose initial syllable is not homographic with any prefix of English, such as *blemish, cancel* and *usurp*. Such words were coded as `no_prefix_string`.[4]

It should be noted at this point that there are some issues with the classification in Wikimorph. There is a set of cases where Wikimorph treats words as unprefixed where some analyses might assume the presence of a prefix, in words such as *arrive, aspire, decrease, digress, elapse, erode, export, inflate, object, perfect, process, recite* or *sublime*. In another set of cases, Wikimorph did not detect a prefix even though most linguists would agree that these words contain a prefix. This concerned, for instance, words such as *disoblige, ensure, imprint*. However, we kept all Wikimorph codings. Using Wikimorph as a classifier has two major advantages: First, it rules out any potential bias of our own that might otherwise be introduced into our classification, especially in cases where several different analyses are possible. Second, taking segmentation information from the Wikimorph model allows us to systematically operationalize the distinction between `opaque` and `transparent`.

In order to study potential effects of syllable weight, we created a weight-informed subset as follows. First, we removed a number of suffixed words from the Wikimorph data set because suffixes can blur the effects of syllable weight, as some systematically shift stress to a specific syllable while others are stress-preserving, irrespective of the weight structure of their final syllable. Some affixes like *-ate* impose prosodic restrictions on their bases or derivatives, such that stress lands systematically on the prefix in trisyllables (e.g. *ácclimate, cómpensate, dédicate, émigrate, ímplicate, óbfuscate, prómulgate, súbjugate*). Therefore, we removed all items with a stress-affecting suffix (e.g. *demonstrate, envision, glorify, luminesce*) and words with a stress-neutral (potential) suffix attached to a free base (e.g. *crackle, fatten, handle, shorten*). We also left out compounds (e.g. *air-cool, crosscut, handcraft, highjack, windsurf*), as those words also follow different generalizations which are not related to weight (see discussion in Dabouis and Fournier, 2023: §3.5). Note that Wikimorph

[4]The distribution of the four categories is given in Sect. 5.1.

Table 2 Distribution of number of syllables, weights of the last syllable and stress in the Weight data set (N = 1,414)

NUMSYLL	2	3	4
N	1198	215	1
WEIGHT	H	L	
N	906	508	
STRESS	antepenult	penult	final
N	14	409	991

was not used for the identification of the suffixed words or compounds, it was only used for detecting prefixes. Those words were identified manually (for details, see Dabouis and Fournier, 2023: §2.3). Then, we added the coding of weight for the last syllable for the remaining verbs (see below).

This reduced the data set to 1414 verb types with the number of syllables, weights of the last syllable and stress distribution as shown in Table 2. We call this data set the 'Weight data set'.

Length information was taken from the CELEX Lexical Database (Baayen et al., 1996), given as the number of syllables in each verb, and weight information was derived from the transcriptions in Roach et al. (2006). Weight information was taken over from Dabouis and Fournier (2023), who based their weight analysis mainly on Hammond (1999). Each final syllable was assigned a mora count based on the segments it contains based on the following assumptions (for further details and justification, see Dabouis and Fournier, 2023: §4.1). The rationale of the coding is to take into account all assumptions made in the literature in connection to the idea that stress assignment is weight-based.

- /ə/ is moraless.
- Final unstressed /i/ and /əʊ/ are monomoraic (but final stressed /əʊ/ is treated like the other long/tense vowels).
- Short/lax vowels (/æ/, /ɛ/, /ɪ/, /ɒ/, /ʊ/, /ʌ/) are monomoraic.
- Long/tense vowels (/aɪ/, /iː/, /uː/, /ɑː/, /ɜː/, /ɔː/, /əʊ/) are bimoraic, except /aʊ/ and /ɔɪ/, which are trimoraic.
- The last consonant is moraless.
- Preconsonantally, sonorants and /s/ are moraless if unstressed and have one mora if stressed, to the exception of /ŋ/, which is bimoraic.
- Preconsonantally, other consonants are treated as monomoraic.

Each item ended up with a mora count between 0 and 3 for its final syllable, which we then turned into a binary coding: `heavy` for final syllables with two or more moras, and `light` for final syllables with fewer than two moras.

For the analysis of the Weight data set the predictor variables and their coding can be summarized as follows.

- MORPHOLOGY. This predictor encodes the status of the prefixal morphology, as shown in the top panel of Fig. 2.
- NUMSYLL. This variable encodes the number of syllables for each verb.
- WEIGHT. This variable encodes the weight of the rime of the last syllable, with `L` for light and `H` for heavy rimes.

Table 3 Overview of the different data sets and the information coded for the respective data set

Name	size	coding
Full data set	2,311	word, stress position
Wikimorph data set	1,948	word, stress position, morphological structure, number of syllables
Weight data set	1,414	word, stress position, morphological structure, number of syllables, weight of final syllable

Table 3 gives an overview of the different data sets.

4.2 Statistical modeling

The choice of an appropriate statistical model partly depends on the nature of the variables. We are dealing with nominal predictor variables as predictors (encoding MORPHOLOGY and WEIGHT), a numerical predictor (LENGTH) and a nominal response variable (STRESS), and it is possible that there are interactions between the predictor variables. With this constellation of variables, tree-building algorithms such as conditional inference trees are a method of choice. These algorithms work through all predictors and partition the data into subsets that differ significantly in their distribution of the response variable from other subsets. Each of the subsets is then characterized by a particular constellation of the values of the predictor variables.

We used conditional inference trees, as implemented in the R packages `party` and `partykit` (Hothorn et al., 2009; Hothorn & Zeileis, 2009). Conditional inference trees are binary trees that are built top-down. The feature that best predicts the response variable splits the data set up, and this is repeated as long as it improves the predictive power of the model. From the top to the bottom of the tree the subsets become increasingly structurally homogeneous. As we will see, an important advantage of this statistical method is that it deals with multinomial response data in an easily interpretable way, which is most welcome in our case, where we predict an outcome with different values (e.g. `antepenult, penult, final`).

Unlike the regression trees, which had access to all structural predictors coded, the NDL network was trained on bigrams only, with no information about more abstract phonological or morphological structure. We then used statistical methods to test to what extent the NDL-derived stresses are correlated with our structural predictors. The rationale and specifications of the different discriminative and statistical models are explained as we go along.

4.3 Implementation of the NDL model

We used the `ndl` package (Arppe et al., 2015) in R with the `estimateWeights` and `estimateActivations` functions to compute the association weights between cues and outcomes using Danks' equilibrium equations (Danks, 2003). Figure 1 depicts the structure of an NDL network as implemented in this study, with the bigrams as cues and stress positions as outcomes.

In line with previous research (e.g. Arndt-Lappe et al., 2023) we use orthographic bigrams because this makes the task of finding the right stress position more difficult

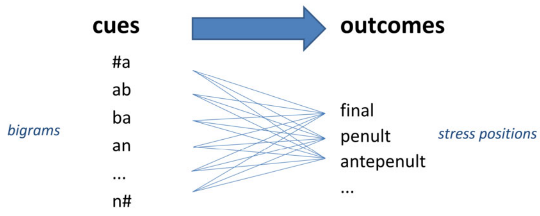

Fig. 1 NDL network with bigrams as cues and different stress positions as outcomes

Table 4 Activation weights of bigrams and stress positions (rounded). Hash marks symbolize the left edge of the word

bigram	preantepenult	antepenult	penult	final
#a	−0.026	0.018	−0.159	0.61
#b	−0.069	0.028	−0.600	1.08
#c	−0.002	−0.079	0.007	0.53

Table 5 Activation weights of whole words (based on their bigrams) and stress positions. Observed stress position and stress position as predicted by the model are given in the two rightmost columns

verb	preante-penult	antepenult	final	penult	observed	predicted
abase	0.01	0.04	1.14	−0.19	final	final
abash	0.00	−0.01	0.91	0.10	final	final
abate	0.01	0.03	1.12	−0.16	final	final
abbreviate	0.00	1.00	0.00	0.00	antepen.	antepen.
abdicate	0.00	1.00	0.00	0.00	antepen.	antepen.

for the NDL model than using phonological representations.[5] In many English words, vowel quality is predictive of stress. For instance, schwas are never stressed, and the letter <a> may represent (unstressed) schwa, but also stressed /æ/ or stressed /eɪ/. Using orthographic representations also avoids making potentially controversial assumptions about underlying phonological vowel qualities.

After training the network, each bigram has a specific association weight to each of the stress positions. This is illustrated in Table 4. Negative associations may also be generated and indicate that cue and outcome tend not to co-occur.

To arrive at associations between whole words and stress position, one adds up the association weights of the bigrams of a given word. Each word thus has association weights with each possible stress position, and by choosing the stress position with the highest association weight, the researcher can use the NDL model to predict the stress position for this word. The results of such calculations are illustrated in Table 5. The two rightmost columns give the observed and predicted stress positions.

Readers might wonder how a model that does not know about syllables would be able to pick up effects that are dependent on the number of syllables, for example to

[5] We also implemented a model using trigrams, but this led to overfitting (with an accuracy of 99.6 percent), due to the very high number of different trigrams (2,628 as against 461 different bigrams).

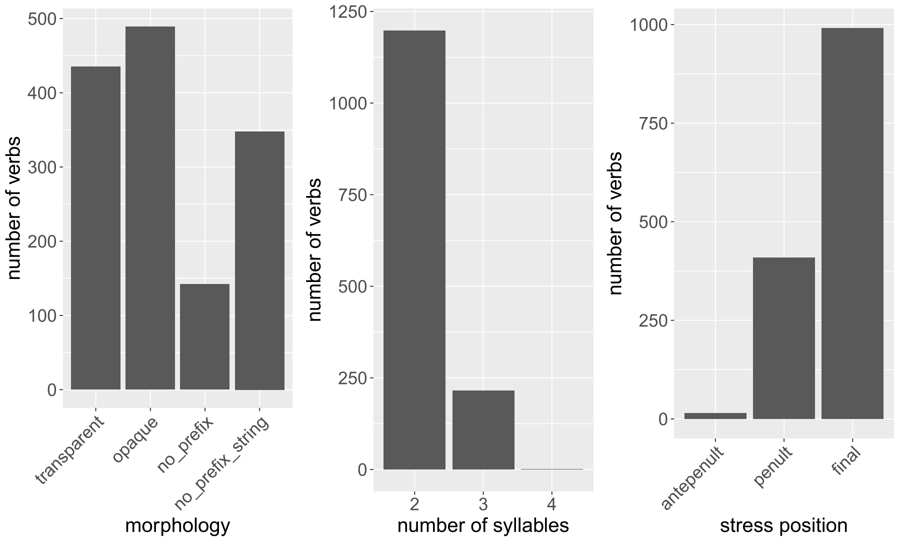

Fig. 2 Descriptive statistics of the Weight data set (N = 1414)

avoid predicting antepenultimate stress in disyllablic words. The answer follows from the representations that are used. If the model makes accurate predictions about stress that reflect differences in the number of syllables of the target words, this is because the associations of the bigrams add up in such a way that the resulting association for a particular type of stress (e.g. antepenult stress) is smaller than those for the other possible stresses (i.e. on the penult or final). If the model is able to pick up an effect of number of syllables, this must mean that words of different length differ in their distribution of bigrams. For example, the phonological makeup of disyllables (in terms of distribution of bigrams) must be different from that of tri- or quadrisyllables.

5 Results

5.1 Predicting stress using structural predictors

We will first give an overview of the Weight data set, then present the results of the regression tree analysis. Finally we will turn to the NDL model. In Fig. 2 we present the distribution of morphological, syllable count and stress position information in the Weight data set. There are 41 different prefixes attested in the data.

Figure 3 shows the resulting tree. The terminal nodes give the majority choice (i.e. 'a' for antepenult, 'p' for penult, and 'f' for final stress) and the proportions of words with the respective stress in this subset, in the same order, i.e. antepenult at the top and final at the bottom. Taking the majority choices in each terminal node as the predicted outcome for that subset, one can compute the overall accuracy of the model's predictions. The tree model has a very high accuracy of 94 percent in the

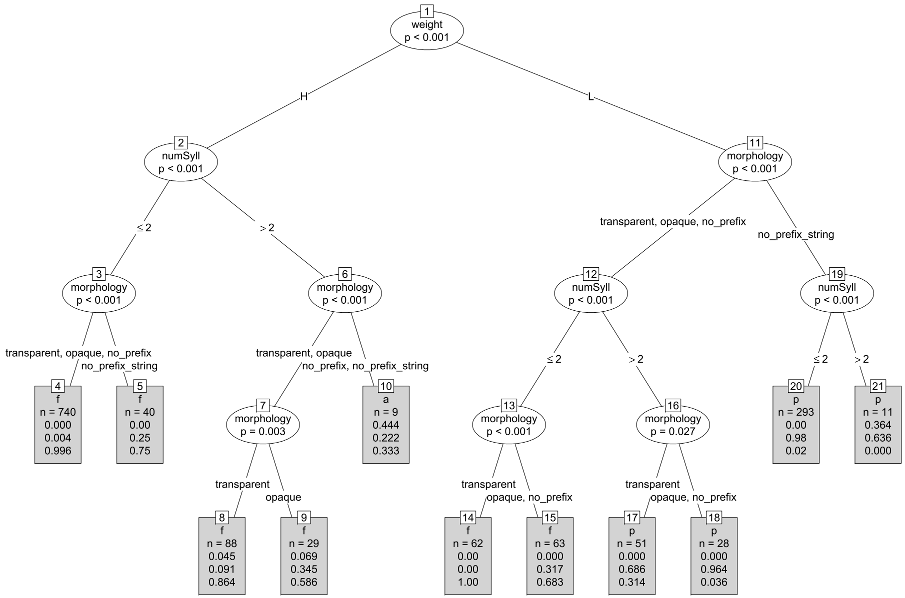

Fig. 3 Conditional inference tree for the Weight data set. Abbreviations represent the majority choice in each subset: ‘a’ = antepenult, ‘p’ = penult, ‘f’ = final. The decimal numbers in the gray boxes give the percentages of antepenult, penult and final stresses (from top to bottom)

sample, with high precision for final (0.97) and penult (0.87) stress, but low precision for antepenult stress (0.28).[6] Let us look at the results in more detail.

The top node (node 1) refers to WEIGHT, which means that the weight of the final syllable is the most influential factor in determining stress position. For words with a heavy final syllable ('H', left branch originating from node 1), final stress is always the majority choice (with the exception of the nine words in node 10), but the proportion of final stresses among subsets of the data varies by number of syllable. Among disyllabic words there is a near-categorical decision (see node 4, only 0.4 percent errors, i.e. 3 out of 740 verbs) for those verbs that start in a string that either is a prefix, be it opaque (e.g. *announce, concern, deceive, rebel*) or transparent (e.g. *amount, condense, demist, rebuild*), or is spelled like a prefix (e.g. *appeal, confine, detain, reckon*). Among the verbs that begin in a string that is not homographic with a prefix (node 5; e.g. *blaspheme, frequent, lament, rampage*) there is about a quarter of words that have penultimate stress. This difference means that the alleged morphological effect is purely based on the recurrence of the word-initial string of letters. We will henceforth refer to such effects as 'form-based'.

With longer words, however, the effect might be genuinely based on transparent morphology, in the sense of our definition (see nodes 8, 9 and 10): The proportion of final stresses decreases monotonically (from almost 90 percent to some 30 percent) with decreasing morphological complexity: transparent prefix (e.g. *condescend, redirect, underlay*) > opaque prefix (e.g. *coincide, reconcile, supervene*) > simplex verb (e.g. *occupy, manifest*). Note, however, that there are only nine non-prefixed verbs in this subset (node 10). We conclude that the main morphological effect among long verbs with heavy final syllables is the transparency effect among prefixed verbs. Unprefixed long verbs are very few in number, but their stress behaviour is in line with the latter effect.

In sum, then, disyllables and longer verbs differ in the nature of the morphological effect in the subset of verbs with heavy final syllables. With disyllables it is purely form-based, in longer verbs it seems to depend on transparency.

Let us turn to the verbs with light final syllables ('L', right branch off node 1). In this subset, MORPHOLOGY and NUMSYLL also both play a role. Node 11 shows that there is a form-based morphological effect present in this subset, too. Verbs that do not start in a prefix string (e.g. *anger, caress, happen, manoeuvre, scurry*) behave differently from verbs that start in a string that either is a prefix (e.g. *beget, discover, overcome, repel, repossess*) or could be one (e.g. *attack, express, interpret, progress*). The former subset (node 19) prefers penult stress, with disyllables showing an almost categorical choice and longer verbs showing more than one third of antepenult stresses.

The verbs starting in prefixal strings (node 12) show a morphological transparency effect: In disyllabic verbs (node 13), a transparent prefix goes together with categorical final stress, while verbs with opaque prefixes or with initial prefix-homographic strings have about one third penultimate stresses. With longer verbs (node 16) we find the same kind of effect, but this time verbs with transparent prefixes have penult stress in 69 percent of the cases, while verbs with opaque prefixes or with initial prefix-homographic strings have almost categorical penultimate stress.

[6] We also performed a ten-fold cross-validation which yielded an accuracy of 93.6 percent.

Table 6 Distribution of observed stresses vs. stresses as predicted by the NDL model (N = 2,311). Majority predictions are shaded (preantepenult has two equally preferred majority decisions)

observed	predicted			
	preantepenult	antepenult	penult	final
preantepenult	3	3	1	0
antepenult	0	425	15	10
penult	0	19	1131	38
final	0	14	87	565

The most important finding of this analysis is that morphology does play a role in stress assignment, but there are two types of effects of a different nature. One is a transparency effect in the sense of the definition adopted in this paper, the other is an effect of recurrent word-initial segmental strings which, crucially, are independent of meaning.

5.2 Predicting stress using an NDL model

In this subsection we present the results emerging from our NDL model. We trained the model on the Full data set, as the model does not need any other information than the orthographic make-up of a word and its stress pattern. By using the largest data set, we also simulate best the learning process, like a language learner who makes use of all available cues and outcomes. The resulting network of associations can then be inspected, or used for deriving predictions, for any data set.

First, we will look at the performance of the model when used as a classifier, as described in Sect. 3. This means that the highest of the three association weights for a given bigram determines the stress position. For this classification task we can use the largest possible data set (Full data set, N = 2,311).

Taking the highest activation for a particular stress position as the model's decision to assign stress to this position, we end up with the distribution of predicted vs. observed stresses as shown in Table 6. The NDL model arrives at an accuracy of 0.92, with highest precision for final (0.95) and antepenultimate stress (0.94), high precision for penultimate stress (0.85) and moderate precision for preantepenultimate stress (0.43).

We also performed a ten-fold cross-validation to see how the model performs on unseen data. This was done in such a way that for each fold of held-out data, we used the 90 percent remaining data as the training set. In this setup, the average cross-validated accuracy was 85.6 percent (ranging from 78.8 to 88.3 percent). The cross-validation results mean that the NDL model is quite able to generalize to new data.

In order to test the NDL model's sensitivity to morphological and phonological properties (i.e. weight of the final syllable and number of syllables), we used the

Weight data set as a test set. For this data set, the accuracy of the bigram-based NDL model is also very high (0.93).[7]

In a second step we then fitted a conditional inference tree to predict the NDL-derived stresses on the basis of WEIGHT, NUMSYLL and MORPHOLOGY. This tree thus models the relation between the three structural properties on the one hand and the stress as predicted on the basis of bigrams on the other hand. If the model finds significant relations, this would mean that the NDL model incorporates these effects even though it is not given explicit information about the respective properties of the words.

The conditional inference tree is given in Fig. 4. The tree shows effects of all three predictor variables. The hierarchy of the effects is not exactly the same as in the analysis of the observed stresses, but this is expected due to the differences between the observed stresses and the stresses as predicted by the NDL model. Morphology emerges as the most important predictor (node 1), with a purely form-based effect (node 13 vs. node 2). For verbs with no initial prefixal string, weight is decisive (node 13). Heavy final syllables go together with a majority of final stresses (node 14), while verbs with light final syllables strongly favour penultimate stress (node 15).

For verbs that begin in a string that could be or is a prefix, length, weight and morphology play a role. The two-syllable verbs (branching off node 2) prefer final stress (nodes 4, 6, 7) to varying degrees depending on weight (node 3) and morphology (nodes 5, 6, 7). For longer verbs, weight is also important. While verbs with a heavy final syllable prefer final stress (node 9), verbs with a light final syllable prefer penult stress, and this tendency towards penult stress is more pronounced for verbs with opaque prefixes (node 12 vs. node 11).

Let us now compare the stress predictions of the regression tree with traditional variables (see Fig. 3) with the stress predictions of the bigram-based NDL model. The decisions by the two models overlap in 90 percent of the cases, and the success of the predictions of the two models does not differ much by morphology, number of syllables and final syllable weight, as is illustrated in Fig. 5, top to bottom, the structure-based model on the left, the bigram-based NDL model on the right. The size of the tiles is proportional to the number of data points with the respective specification.

5.3 Morphological effects in a model without morphology

We will now explore in more detail how the NDL model is able to learn about morphology, so that the effects we see in nodes 1, 5 and 10 can emerge. We start with a number of testable expectations that relate morphological structure to the activation weights of the NDL model. If the NDL model is able to learn the role of morphology, we expect the following:

[7] In an alternative setup, we trained the NDL model on the Weight data set. In this setup, the NDL model reaches an even higher accuracy of 0.95 (and 0.89 in cross-validation). We decided to use the NDL model trained on the Full data set because this seems theoretically preferable and empirically more conservative. Furthermore, it works against the NDL model.

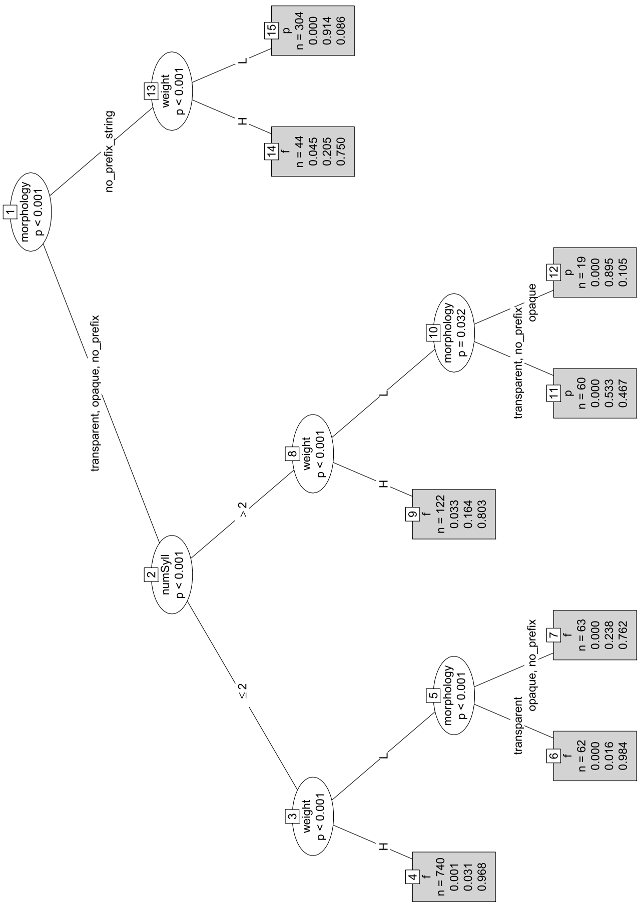

Fig. 4 Conditional inference tree for the Weight data set, fitted to the stress positions as predicted by the NDL model. Abbreviations represent the majority choice in each subset: ‘a’ = antepenult, ‘p’ = penult, ‘f’ = final. The decimal numbers in the gray boxes give the percentages of antepenult, penult and final stresses (from top to bottom)

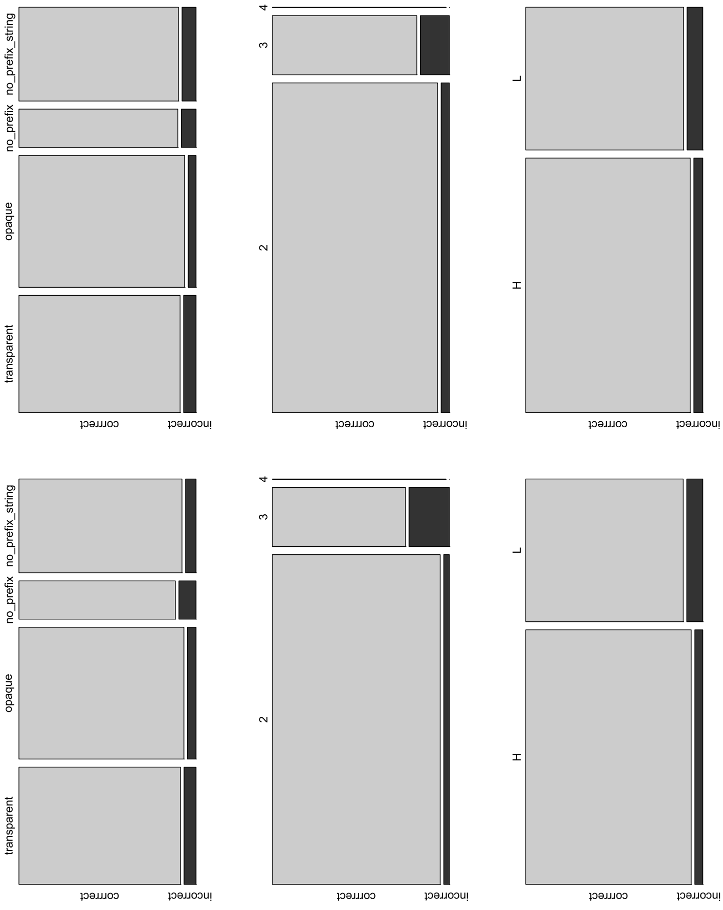

Fig. 5 Mosaic plots comparing the correctness of the structure-based model vs. the bigram-based NDL model. The left column of panels shows the structure-based model, the right column the bigram-based NDL model. The top row shows correctness by morphology, the second row by number of syllables, and the third row by weight of the final syllable

- Activations should vary by prefix. If the recurrence of the association of segmental form and stress is what the model picks up, the fact that prefixes are not recurrent to the same degree entails that they should have different activations.
- Prefixed words should have low activations for the first syllable because the prefixes should be unstressed.
- For the same reason the activations of the first syllables should be lower for prefixed words than for simplex words.

As we will see, all expectations are in accordance with what we find in the model. Figure 6 depicts the mean activation for each stress position by prefix (given the small number of only four preantepenult stresses, we do not show the activations for these verbs). We can see that affixes indeed vary considerably. For instance, *ante-* has the second highest activation among all prefixes for antepenult stress and a negative activation for penult stress, while *post-* has a high activation for final stress (like many other prefixes) but a medium activation for penult stress and a negative activation for antepenultimate stress.

Figure 7 demonstrates that indeed the prefixed words have significantly lower stress activations for the first syllable. The value `simplex` unites the values `no_prefix` and `no_prefix_string` we used above.[8]

To summarize, we can say that all three predictions from above are confirmed. This means that the NDL model shows morphological effects although the model has no explicit morphological information. This may indicate that morphological effects emerge as a by-product of the association of recurrent phonological strings with stress patterns and are thus epiphenomenal in nature.

6 Discussion and conclusion

In this paper we investigated two main questions: First, what is the role of morphological structure (specifically, of prefixes) in stress assignment to English verbs? In particular, we tested the idea that effects of morphology might be purely form-based in nature, arising through recurrent word initial phonological strings, independent of their relation to meaning. Second, we investigated whether stress assignment can be modeled using a discriminative learning model that has no a-priori information about syllable weight, number of syllables or morphology, and operates solely on the basis of bigrams and their associations to the three stress positions. What did we learn?

With regard to the nature of the alleged morphological effects in stress assignment to English verbs, we saw that most of these effects are effects of recurrent orthographic strings. That is, verbs that start in a string that is not homographic with a prefix behave differently from those verbs that start in a string that is homographic with a prefix and could be a prefix or not. This finding is consistent with several studies that have used pseudo-words to study stress assignment and that have reported

[8]There is also a difference between transparent and opaque prefixes, but only with trisyllabic words. The interpretation of this difference is unclear, but has a parallel in the structure-based tree model. Nodes 7 and 16 represent subsets of trisyllabic words that also show a difference between words with transparent vs. opaque prefixes.

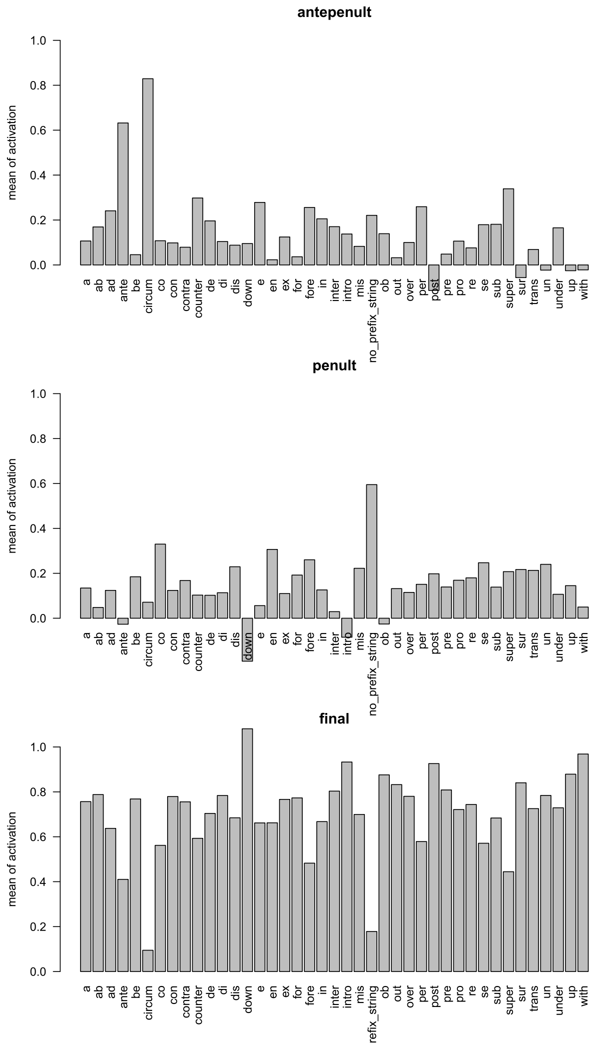

Fig. 6 Mean activation by prefix for each stress position (without the four verbs with preantepenult stress)

effects of prefixation on stress decisions based only on orthographic strings that can be prefixes, including differences between individual opaque prefixes of the type we find reflected in the NDL activations in this study (Rastle & Coltheart, 2000; Ktori et al., 2016, 2018; Gabillet et al., 2024; Treiman et al., 2021; Turcsan & Herment, 2015). Our findings also complement the evidence reported for suffixes in Arndt-Lappe et al.'s study (2023), which has shown that NDL activations also reflect differences in stress behaviour between stress-shifting and stress-preserving suffixes in the lexicon.

Interestingly, the regression tree analyses also revealed a transparency effect whereby transparent forms behaved differently from opaque forms, or from opaque and non-prefixed forms. This aligns with experimental studies of pseudo-verbs that report effects of semantic transparency in stress assignment (e.g. Gabillet et al., 2024).

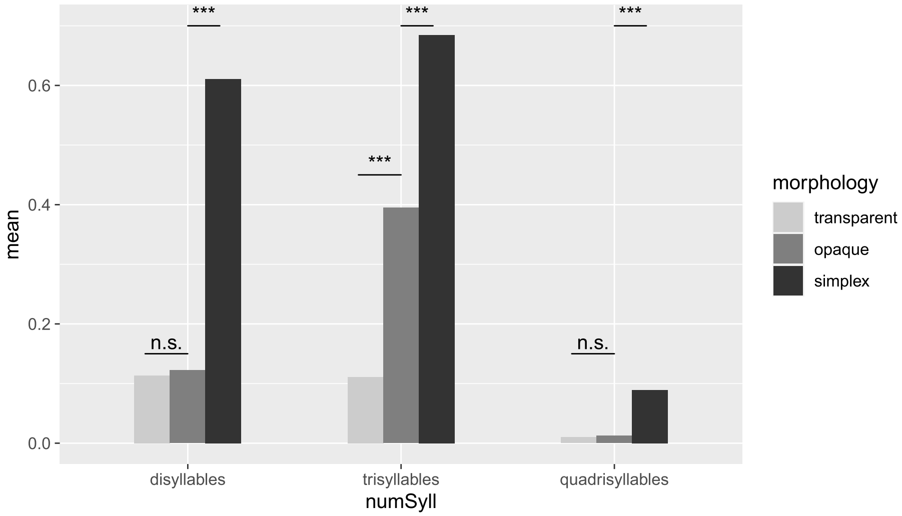

Fig. 7 Mean activation of first syllable by verb length and morphology. The horizontal lines indicate the (non)-significance of contrasts (Wilcoxon tests). Abbreviations: *** = $p < 0.001$, n.s. = $p > 0.05$

Comparing the discriminative learning network's predictions with those of a statistical model that uses the traditional structural predictors weight and morphology, we found that the NDL model shows a comparable accuracy in predicting stress assignment as the statistical model that uses traditional predictors. This success of the NDL model is in line with similar NDL studies of stress such as Arndt-Lappe et al. (2023) for English, and Tomaschek et al. (2023) for German, as well as with recent studies exploring English verb stress with Analogical Modeling, a different computational algorithm (Arndt-Lappe et al., 2025; Seiler & Arndt-Lappe, 2025). The high predictive power of the NDL model indicates that phonotactic information alone may be sufficient to understand stress assignment. Closer inspection of the NDL model revealed that phonological and morphological effects emerge from the association weights even though the model does not have explicit information of syllable structure (number of syllables and weight) and morphology.[9]

While generally consistent with prior studies, our findings also diverge from previous work in one important respect. In particular, the discrepancy in accuracy between the current verbs-only model and Arndt-Lappe et al.'s (2023) model trained on all grammatical categories merits some discussion. Unlike in their study, NDL achieved high accuracy when predicting stress position as counted from the right word edge in our simulations. By contrast, Arndt-Lappe et al.'s simulation, which

[9]This paper was primarily interested in comparing two kinds of architecture. The effect of morphology could also be tested holding the architecture constant and simply comparing two NDL models, one with just bigrams as predictors, and the other with bigrams plus morphological information. The accuracy of the only bigrams-based model is, however, already so high (0.95, see footnote 7) that including morphological information on top cannot be expected to change much. And indeed, accuracy rises only slightly from 0.95 to 0.97 (0.89 to 0.91 for cross-validated models). This rise is statistically not significant ($\chi^2 = 87{,}808$, $p = 1$, $df = 1$ (non-cross-validated), $\chi^2 = 151{,}976$, $p = 1$, $df = 1$ (cross-validated), χ^2-square test).

comprised 33,407 English lemmas across all grammatical categories, exhibited substantially lower accuracy for right-edge stress classification (i.e., 59 percent accuracy on bigrams; compare this also to 71.4 percent accuracy for classification from the left edge and 72.1 percent for classification in terms of the stressed vowel).

These differences may arise from two sources. The first is that Arndt-Lappe et al.'s data set also contained suffixed words, which means that this data set was less controlled to begin with, and thus subject to more uncontrolled variability. As a second reason for the differences, we suggest that they likely reflect fundamental differences in the structural makeup of words across lexical categories, with important implications for how these words are stressed. English nouns and verbs differ not only in stress patterns but also in morphological structure: nouns exhibit a relatively greater number and variety of stress-preserving suffixes compared to verbs (see, for example, the overviews in Bauer et al., 2013, Chaps. 10-13), which makes stress position arguably less uniform when counted from the right word edge and therefore harder to predict. Verbs, by contrast, have been shown to be structurally more uniform. In a large-scale dictionary study, Dabouis and Fournier (2023) report that verbs – especially disyllabic ones – have more morphology on the left than on the right edge, while about two thirds of longer verbs are suffixed with only a small set of suffixes; related figures from Guierre (1979) reported in Dabouis and Fournier (2025) further indicate that items without productive morphology are prefixed in as many as 77 percent of verbs but far less frequently in other categories. These structural tendencies suggest that the form-based effects we observe are likely to be most pronounced in verbs. In sum, the differences in stress predictability between verbs and other syntactic categories are at least partly due to differences in the phonological and morphological make-up of these classes.

In the study reported here, the role of meaning was not investigated directly, as our data was only categorized based on the morphological status of bases. However, it has been pointed out by Dabouis and Fournier (2025) that certain historical prefixes and bound roots may still have recurrent meaning, although they can be quite vague (e.g. 'back' in *rebound, reflect, repel*, 'breaking' in *abrupt, interrupt, rupture*). Similarly, Baeskow (2006) argues that the verbs with bound roots such as *duce, ceive* or *mit* have similar argument structures, which he calls "skeletal thematic information" (after Lieber (2004)). More insights into the synchronic status of historical prefixes and roots might be gained by exploring their properties using methods such as distributional semantics (Widdows & Cohen, 2010), as shown recently by Carlson and Crosson (2025).

Our results also provide a challenge for theories of phonology that restrict their interface with morphology to transparent constructions. In this respect, our findings are in line with previous work such as Dabouis and Fournier (2025), Dabouis (2023), Köhnlein (2015) or Mascaró (2016), who have made proposals to capture the phonological effects of semantically impoverished formatives.

Funding information Open Access funding enabled and organized by Projekt DEAL. This research was carried out in the context of the project 'English Root Stress across Frameworks' (grant 469696746), jointly funded by the Deutsche Forschungsgemeinschaft (Sabine Arndt-Lappe) and the Agence Nationale de la Recherche (Quentin Dabouis). This funding is gratefully acknowledged.

Declarations

Competing interests The authors declare no competing interests.

References

Arndt-Lappe, S., Schrecklinger, R., & Tomaschek, F. (2023). Stratification effects without morphological strata, syllable counting effects without counts–modelling English stress assignment with naive discriminative learning. *Morphology*, *33*(4), 433–457.

Arndt-Lappe, S., Ganster, T., & Seiler, A. (2025). Explaining analogy in word-formation: The role of lexical network structure. In R. Huyghe, M. Prudent, & J. Salvadori (Eds.), *Proceedings of the fifth international workshop on resources and tools for derivational morphology (derimo)* (pp. 11–20). Fribourg: University of Fribourg.

Aronoff, M. (1976). *Word formation in generative grammar (Vol. 1)*. Cambridge: MIT Press.

Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., Sering, T., & Shaoul, C. (2015). ndl: Naive Discriminative Learning [Computer software manual].

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). *Celex2*. Philadelphia: Linguistic Data Consortium.

Baayen, R. H., Milin, P., Filipović Durdević, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438–481.

Baayen, R. H., Chuang, Y.-Y., & Blevins, J. P. (2018). Inflectional morphology with linear mappings. *The Mental Lexicon*, *13*(2), 230–268.

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*, *2019*(1), 1–39.

Baeskow, H. (2006). A revival of romance roots. *Morphology*, *16*(1), 3–36. https://doi.org/10.1007/s11525-006-0001-7.

Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711–733.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Carlson, M. T., & Crosson, A. C. (2025). The synchronic status of historical bound roots in the mental lexicon. *The Mental Lexicon*. https://doi.org/10.1075/ml.24032.car.

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

Dabouis, Q. (2023). Phonological evidence for morphological complexity in English proper names. *Anglophonia*, *36*. https://doi.org/10.4000/11qbh.

Dabouis, Q., & Fournier, J.-M. (2023). The stress patterns of English verbs: Syllable weight or morphology? In N. Ballier, J.-M. Fournier, A. Przewozny, & E. Yamada (Eds.), *New perspectives on English word stress* (pp. 154–191). Edinburgh: Edinburgh University Press.

Dabouis, Q., & Fournier, J.-M. (2025). Opaque morphology and phonology: Historical prefixes in English. *Journal of Linguistics*, *61*(2), 231–263. https://doi.org/10.1017/S002222672400015X.

Daelemans, W., Gillis, S., & Durieux, G. (1994). The acquisition of stress: A data-oriented approach. *Computational Linguistics*, *20*(3), 421–451.

Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology*, *47*, 109–121.

Delbridge, A., Bernard, J. R. L., Blair, D., Butler, S., Peters, P., & Yallop, C. (Eds.) (2009). *Macquarie dictionary, Autralia's national dictionary online*. Macquarie Library.
Eddington, D. (2000). Spanish stress assignment within the analogical modeling of language. *Language*, *76*(1), 92–109.
Forster, K. I., & Azuma, T. (2000). Masked priming for prefixed words with bound stems: Does submit prime permit? *Language and Cognitive Processes*, *15*(4–5), 539–561. https://doi.org/10.1080/01690960050119698.
Fournier, J.-M. (1996). La reconnaissance morphologique. In *8ème Colloque d'Avril sur l'anglais oral* (pp. 45–75). Villetaneuse: Université de Paris-Nord, CELDA, diffusion APLV.
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.
Gabillet, M., Seiler, A., Arndt-Lappe, S., & Dabouis, Q. (2024). Stress assignment in English pseudowords: On the role of weight, prefixation and syntactic category. Paper presented at the 21st Old World Conference in Phonology at the University of Leipzig (14-16 February).
Guierre, L. (1979). Essai sur l'accentuation en anglais contemporain: Eléments pour une synthèse. Unpublished doctoral diss., Université Paris-VII.
Guierre, L. (1990). Mots composés anglais et agrégats consonantiques. In *5ème colloque d'avril sur l'anglais oral* (pp. 59–72). Villetaneuse: Université de Paris-Nord, CELDA, diffusion APLV.
Hammond, M. (1999). *The phonology of English: A prosodic optimality-theoretic approach* (J. Durand, Ed.). Oxford: Oxford University Press.
Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry*, *13*(2), 227–276.
Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2021). Modeling morphology with linear discriminative learning: Considerations and design choices. Submitted for publication. Retrieved from https://arxiv.org/abs/2106.07936.
Hothorn, T., & Zeileis, A. (2009). Partykit: A toolkit for recursive partytioning. Retrieved from http://R-Forge.R-project.org/projects/partykit/.
Hothorn, T., Hornik, K., Strobl, C., & Zeileis, A. (2009). Party: A laboratory for recursive partytioning. Retrieved from http://CRAN.R-project.org/package=party.
Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York: Appleton-Century-Crofts.
Köhnlein, B. (2015). The morphological structure of complex place names: The case of Dutch. *Journal of Comparative Germanic Linguistics*, *18*(3), 183–212. https://doi.org/10.1007/s10828-015-9075-0.
Ktori, M., Tree, J. J., Mousikou, P., Coltheart, M., & Rastle, K. (2016). Prefixes repel stress in reading aloud: Evidence from surface dyslexia. *Cortex*, *74* (October 2015), 191–205. https://doi.org/10.1016/j.cortex.2015.10.009.
Ktori, M., Mousikou, P., & Rastle, K. (2018). Cues to stress assignment in reading aloud. *Journal of Experimental Psychology*, *147*(1), 36–61. https://doi.org/10.1037/xge0000380.
Lieber, R. (2004). *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
Lieber, R. (2021). *Introducing morphology* (3rd ed.). Cambridge: Cambridge University Press.
Mascaró, J. (2016). Morphological exceptions to vowel reduction in central Catalan and the problem of the missing base. *Catalan Journal of Linguistics*, *15*, 27–51. https://doi.org/10.5565/rev/catjl.184.
Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, *52*(1), 111–139.
Plag, I. (2018). *Word-formation in English* (2nd ed.). Cambridge: Cambridge University Press.
Plag, I., Heitmeier, M., & Domahs, F. (2023). German nominal number interpretation in an impaired mental lexicon: A naive discriminative learning perspective. *The Mental Lexicon*, *18*(3), 417–445.
Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*(6), 909–957.
Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language*, *42*, 342–364. https://doi.org/10.1006/jmla.1999.2687.
Rescorla, R. A. (1988a). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience*, *11*(1), 329–352.
Rescorla, R. A. (1988b). Pavlovian conditioning: It's not what you think it is. *The American Psychologist*, *43*(3), 151–160.
Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Roach, P., Setter, J., Esling, J. H., & Jones, D. (2006). *Cambridge English pronouncing dictionary* (17th ed.). Cambridge: Cambridge University Press.
Saito, M., Tomaschek, F., & Baayen, R. H. (2020). Relative functional load determines co-articulatory movements of the tongue tip. In *Proceedings of the 12th international seminar on speech production*. New Haven, CT. (217).
Schmitz, D., Baer-Henney, D., & Plag, I. (2021). The duration of word-final /s/ differs across morphological categories in English: Evidence from pseudowords. *Phonetica*, *78*(5–6), 571–616.
Seiler, A., & Arndt-Lappe, S. (2025). Morphology, analogy and lexical distributions: The case of English verb stress. Paper presented at the *14th Mediterranean Morphology Meeting (MMM)*. Zadar, 12–14, 6.
Stein, S. D., & Plag, I. (2021). Morpho-phonetic effects in speech production: Modeling the acoustic duration of English derived words with linear discriminative learning. *Frontiers in Psychology*, *12*, 678712.
Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*(3), 271–294. https://doi.org/10.1080/01690969408402120.
Tomaschek, F., & Ramscar, M. (2022). Understanding the phonetic characteristics of speech under uncertainty-implications of the representation of linguistic knowledge in learning and processing. *Frontiers in Psychology*, *13*, 754395.
Tomaschek, F., Domahs, U., & Domahs, F. (2023). Modelling German word stress. *Glossa: A Journal of General Linguistics*, *8*(1).
Treiman, R., Wolter, S., & Kessler, B. (2021). How sensitive are adults to the role of morphology in spelling? *Morphology*, *31*(3), 261–271.
Turcsan, G., & Herment, S. (2015). Making sense of nonce word stress in English. In J. A. Mompean & J. Fouz (Eds.), *Investigating English pronunciation: Current trends and directions* (pp. 23–46). Hampshire: Palgrave Macmillan.
van de Vijver, R., & Uwambayinema, E. (2022). A word-based account of comprehension and production of Kinyarwanda nouns in the Discriminative Lexicon. Linguistics Vanguard. Retrieved from https://psyarxiv.com/78a6e/.
Widdows, D., & Cohen, T. (2010). The semantic vectors package: New algorithms and public tools for distributional semantics. In *Proceedings - 2010 IEEE 4th international conference on semantic computing, ICSC 2010* (pp. 9–15). https://doi.org/10.1109/ICSC.2010.94.
Wurm, L. H. (1997). Auditory processing of prefixed English words is both continuous and decompositional. *Journal of Memory and Language*, *37*(3), 438–461. https://doi.org/10.1006/jmla.1997.2524.
Wurm, L. H. (2000). Auditory processing of polymorphemic pseudowords. *Journal of Memory and Language*, *42*, 255–271. https://doi.org/10.1006/jmla.1999.2678.
Yarbro, J. T., & Olney, A. M. (2021). Wikimorph: Learning to decompose words into morphological structures. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *Artificial intelligence in education* (pp. 406–411). Cham: Springer.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Ingo Plag[1] · Sabine Arndt-Lappe[2] · Aaron Seiler[2] · Quentin Dabouis[3]

I. Plag
ingo.plag@uni-duesseldorf.de

S. Arndt-Lappe
arndtlappe@uni-trier.de

A. Seiler
seiler@uni-trier.de

Q. Dabouis
quentin.dabouis@uca.fr

[1] Institut für Anglistik und Amerikanistik, English Language and Linguistics, Heinrich-Heine-Universität Düsseldorf, 40204 Düsseldorf, Germany

[2] Fachbereich II, Anglistik, Universitätsring 15, D-54296 Trier, Germany

[3] Université Clermont Auvergne, Département d'Anglais, 34 avenue Carnot, 63000 Clermont-Ferrand, France