On the replicator dynamics of lexical stress: accounting for stress-pattern diversity in terms of evolutionary game theory*

Andreas Baumann
Nikolaus Ritt
University of Vienna

This paper accounts for stress-pattern diversity in languages such as English, where words that are otherwise equivalent in terms of phonotactic structure and morphosyntactic category can take both initial and final stress, as seen in 'lentil – ho*tel, 'envoy – de*gree, 'researchN – re*searchN and 'accessV – ac*cessV. Addressing the problem in general and abstract terms, we identify systematic conditions under which stress-pattern diversity becomes stable. We hypothesise that words adopt stress patterns that produce, on average, the best possible phrase-level rhythm. We model this hypothesis in evolutionary game theory, predict that stress-pattern diversity among polysyllabic word forms depends on the frequency of monosyllables and demonstrate how that prediction is met both in Present-Day English and in its history.

1 Introduction

In this paper we examine whether there are circumstances under which the lexicon of a language will inevitably display stress-pattern diversity. The question is inspired by the situation that obtains in English, but we use English evidence primarily to explain the problem itself, to justify our approach and methodology, and to demonstrate our hypotheses. Instead of trying to account specifically for the facts attested in English, we conceptualise the problem as a general one, and accordingly model it in abstract terms.

* E-mail: ANDREAS.BAUMANN@UNIVIE.AC.AT, NIKOLAUS.RITT@UNIVIE.AC.AT.

For constructive comments on earlier versions of this paper we owe thanks to Donka Minkova, Elan Dresher, Heinz Giegerich, Kenny Smith and Simon Kirby, three anonymous reviewers and the editors of Phonology. Obviously, any remaining errors are ours.

The appendices mentioned in the paper are available as online supplementary materials at https://doi.org/10.1017/S0952675717000240.
When we say that English displays stress-pattern diversity, we mean that polysyllabic words can assume different stress patterns even though, in terms of phonotactic structure, morphological composition and morphosyntactic category they are either equivalent, as the examples in (1a), or even practically identical, as (1b).

(1) a. ['al.ge.bra]$_N$ vs. [ven.'de.tta]$_N$ b. ['de.fence]$_N$ vs. [de.'fence]$_N$
   ['Ca.na.da]$_N$ vs. [ba.'na.na]$_N$ [re.'search]$_N$ vs. [re.'search]$_N$
   ['con.cert]$_N$ vs. [con.'cern]$_N$ [re.'vert]$_N$ vs. [re.'vert]$_N$
   ['con.voy]$_N$ vs. [dis.'may]$_N$ [per.'fume]$_V$ vs. [per.'fume]$_V$
   ['en.voy]$_N$ vs. [de.'gree]$_N$ [ac.'cess]$_V$ vs. [ac.'cess]$_V$
   ['in.tern]$_N$ vs. [con.'cern]$_N$ ['a.zure]$_A$ vs. ['a.zure]$_A$
   ['len.til]$_N$ vs. [ho.'tel]$_N$ [ro.'bust]$_A$ vs. [ro.'bust]$_A$
   ['par.lay]$_N$ vs. [ar.'ray]$_N$
   ['cur.rent]$_A$ vs. [at.'tent]$_A$
   ['con.stant]$_A$ vs. [con.'tent]$_A$

This diversity is interesting because it goes against the fact that, for the most part, English words have unique stress patterns, and these patterns can be predicted quite precisely if one takes phonotactic structure, morphological composition and morphosyntactic category into account. This has been repeatedly demonstrated in a large body of research (see e.g. Chomsky & Halle 1968, Liberman & Prince 1977, Kiparsky 1979, Hayes 1982, Giegerich 1985, Anderson 1986, Kager 1989, Burzio 1994, Halle 1998, van Oostendorp 2005, Hyde 2007, 2011, Schane 2007). Thus, and although work on English word stress has reached a high level of descriptive adequacy and coverage, examples like the ones in (1) represent a residue of obvious irregularities that remain to be accounted for.

We introduce and defend two hypotheses. The first, and more fundamental one, is that constraints on word stress are essentially rhythmic, and apply primarily on the phrase level rather than on the word level itself. That stress patterns can nevertheless be predicted quite well by considering isolated words alone is because words anticipate and adapt to the phrases in which they occur. Secondly, and more specifically, we propose that stress-pattern diversity will establish itself stably among polysyllables when there is a sufficiently large proportion of monosyllables in speech. We derive the second hypothesis from the first one through a simulation in terms of an evolutionary game. In the main part of our paper, we motivate, describe and interpret that game in detail. Basically, it constructs polysyllabic words as ‘players’ and stress patterns as ‘strategies’. In each round, two words combine to form a phrasal sequence, sometimes with monosyllabic items between them. If the resulting phrase is rhythmically well-formed (observing, for example, the constraint that feet should be binary), the words earn a ‘reward’, and their stress patterns are stabilised.
If, on the other hand, the sequence is arhythmic, the stability of their stress patterns decreases.

First, however, we clarify how we conceptualise the problem. We are interested in the general conditions under which a language will admit and establish stress-pattern diversity in its lexicon. Such general conditions are different from the specific reasons for individual lexical items assuming their specific patterns, given that diversity is generally licensed. This means that we do not ask, for example, why *lentil* and *intern* take initial stress, or why *hotel* and *concern* take final stress. What we ask instead is why languages like English allow – or possibly require – a subset of lexical items to take initial stress in some cases and final stress in others.¹

As we shall see, our game predicts that diverse stress patterns will be stably established in the set of polysyllabic word forms, if utterances contain a sufficiently high number of monosyllables. Since English is indeed highly monosyllabic, it is clearly in line with that prediction. As our simulation is an evolutionary game, however, it also makes testable predictions about the ways in which stress-pattern distributions may change diachronically. We show that these predictions fit the historical development of English stress patterns well, and take this to corroborate our hypotheses.

Although the literature on stress assignment is vast, the question we address has not yet been discussed. However, our study builds on and complements various strands of extant research. Our own background is in natural phonology (cf. e.g. Donegan & Stampe 1979, 2009, Stampe 1979, Dressler 1985, Dziubalska-Kołaczyk 2009) and in theories of language as a culturally evolving system (cf. e.g. Lass 1997, Croft 2000, Blevins 2004, Ritt 2004, Brighton et al. 2005, Christiansen & Chater 2008a, b, Smith & Kirby 2008, Wedel 2009, Pierrehumbert 2012, Tallerman & Gibson 2012, McMahon & McMahon 2013).

Sharing views already expressed by Paul (1920) and revived in recent decades (see the references above), we regard constituents of linguistic competence – in this context the segmental and prosodic shapes of words – as being instantiated in populations of speaker minds. They owe their existence to successful transmission via communication and language acquisition. To be transmitted, lexical items (and, indeed, all constituents of competence) need to be articulated, perceived and processed in

¹ Crucially, the general question is not just a summary of the specific questions it appears to cover. It is qualitatively different, and belongs to a more fundamental level of explanation. Thus, even if one answers all the specific questions, one will not have answered the general one – nor vice versa. For instance, one might explain the final stress in *hotel* with reference to its French origin, and/or possibly the associations of elegance and luxury that a Romance stress pattern may evoke. Even if such an explanation were true (which may or may not be the case), it would not answer the deeper question of why English admits diverse stress patterns for words such as *lentil, hotel*, etc. Not all languages do. Polish, for example, has obligatory penultimate stress (also in *hotel*). Similarly, one will not have explained why *lentil* has initial stress and *hotel* final stress if one knows why English allows both patterns on words of that type.
discourse, and their transmission is subject to physiological and cognitive constraints on these processes.

Following a large body of linguistic and biological literature (see e.g. Hayes 1984, as well as Fitch 2013 and the references therein), we think that one class of constraints reflects a deeply rooted preference for linguistic utterances to be rhythmically structured. As indicated, our study rests on the idea that lexical word-stress patterns represent adaptations to rhythmic constraints on the phrases they form when uttered. The idea is not new. The impact of rhythmic, and ultimately utterance-based, constraints on the lexical and grammatical constituents of specific languages is demonstrated by, for example, Donegan & Stampe (1983), Kelly (1988, 1989), Schlüter (2005) and Fullwood (2014). Also, our paper relates to studies of rhythmically induced accent shifts, as in 'Tennessee' air and 'thirteen' men (cf. e.g. Liberman & Prince 1977, Hayes 1984). Unlike the latter, however, we do not restrict our focus to cases where phrase-level rhythm actually reverses lexical prominence relations. Instead, we explore the possibility that all lexical stress patterns normally reflect such constraints, and that this might explain why observable stress shifts represent the exception rather than the norm.

The revival of evolutionary approaches to language has motivated the increased use of quantitative methods developed for the study of dynamical systems in other domains, particularly evolutionary biology. Their fruitfulness has encouraged us to apply such a method to our own research question. The particular one we have chosen is evolutionary game theory, because our question is about general principles governing the distribution of stress patterns in the lexicon that result from interactions among the items that make it up, and evolutionary game theory is as an established tool for addressing issues of this kind (see e.g. Ross 2016).

The view that constituents of linguistic competence are shaped by factors whose immediate effect is on discourse rather than on cognitive representations of linguistic knowledge relates our work to utterance-based theories, as represented, for example, in the work of Bybee (e.g. Bybee et al. 1998, Bybee 2001, 2007) and Pierrehumbert (e.g. Altmann et al. 2009, Pierrehumbert 2012). Since our concern is not with individual words, however, but with global properties of the lexicon, frequency effects on specific items do not figure centrally in our study, nor does the question of whether words are cognitively represented as exemplars. In this respect, our study also differs from statistical investigations of stress-pattern diversity such as Domahs et al. (2014), which derives the stress patterns assumed by English compounds probabilistically but quite precisely from the interaction of a large variety of factors.

A recent study which is closely related to our own concerns – in terms of both approach and method – is Sonderegger & Niyogi (2010, 2013). They approach the functionally motivated contrast between stress patterns among English nouns and verbs in terms of dynamical systems modelling. Although we focus on stress-pattern diversity that is not morphosyntactically
motivated, the results of our model are in many respects compatible with their predictions (see §4.4).

We proceed as follows. First, in §2, we discuss existing approaches to word stress and why they do not account for stress-pattern diversity as attested in English. We argue that this is because established stress-assignment algorithms take isolated lexical items as their input, although the structures they build are essentially rhythmic, and rhythm is a property of phrases rather than words. We suggest an alternative view of stress assignment, in which rhythmic well-formedness constraints apply to phrases, rather than words. Words then adopt the patterns that satisfy these constraints best – on average – in all the phrases in which they occur. We also discuss some of the implications of a phrase-based theory of word stress and the predictions can be derived from it. We show that it makes the task of predicting the stress patterns of individual words highly complex, but allows us to derive hypotheses about the global distribution of stress patterns.

In §3 we explain why and how our research question can be addressed in terms of evolutionary game theory. We explain how phrase-based stress assignment can be construed as an evolutionary process, in which constraints on phrase-level rhythm select for a stable distribution of word-stress patterns in the lexicon, so that uttered phrases are maximally eurythmic. We also justify the choice of evolutionary game theory for modelling the dynamics of that process, deal with the abstractions that this requires and discuss their implications for interpreting the game as a plausible model of actual languages.

In §4, the main section of this study, we describe the game itself – from both linguistic and mathematical perspectives – and interpret its results. We derive the hypothesis that lexical phonotactics will license stress-pattern diversity if the lexicon contains a large proportion of monosyllables, and demonstrate how this can explain the diversity attested in English.

§5 shows that the diachronic predictions inherent in the model also fit the actual evolution of English word stress, thereby corroborating the validity of the model. §6 summarises our observations, and proposes possible directions for further research.

2 English word stress

The stress patterns of English polysyllables are assigned lexically. Knowing a word implies knowing which of its syllables to stress: in ‘father’ it is the first syllable, in ‘degree’ it is the last one, in ‘agenda’ the second, etc. In the majority of cases, English lexical stress is also ‘immobile’, i.e. the syllable that is lexically stressed normally emerges as prominent in all utterances of a word. However, observable stress shifts do occur in a non-negligible minority of items. These include words such as ‘fifteen’, which is stressed on the last syllable in isolation and in expressions like
she’s *fifteen*, but on the first syllable in phrases like ‘*fifteen* years’ (cf. also *Chin’ese* vs. *Chi’nese* whispers, *Ber’lin* vs. *Ber’lin* ‘Wall’, *Prin’cess* vs. *Prin’cess ‘Anne*, and *Tennes’see* vs. *Tenes’ssee* ‘air, *Pennsyl’vania* vs. *Pennsyl’vania* ‘Legis’lature’, etc.).²

Although English stress patterns are properties of lexical items, they are not unpredictable, but seem to be systematically related to other lexical properties, some of them phonotactic, others morphosyntactic. For many English words, the position of stress can be predicted if we know the number and the weight of their syllables, and whether they are nouns, verbs, adjectives, etc.

Existing theories of word stress provide many proposals for formalising these relations, and they differ from theory to theory (see again the references above). What most of them have in common, however, is that they consider lexical items in isolation. In order to capture the claim that some aspects of a word’s structure imply others, they take the former to be ‘underlying’ and the latter to be derived (by rules or ranked constraints). The stress pattern derived for a word always results from (i) its assumed phonotactic and morphological structure and its word class, and (ii) whatever derivational algorithm a model happens to use. Thus the stress pattern of a word is fully implied in its phonotactic and morphosyntactic properties.

Although existing models are sophisticated and quite adequate, none of them fits the facts fully. As the examples in (1) show, there is a substantial number of words that are fully equivalent in terms of syllable weight and morphological and syllabic structure, and belong to the same morphosyntactic categories, but are nevertheless stressed differently. It is obviously impossible to accommodate such word pairs in a model that derives stress patterns from the phonotactic structure and the word class of isolated lexical inputs. There is simply no way in which a deterministic algorithm that predicts a specific stress pattern for one specific type of lexical input can at the same time predict another one.

One way of dealing with the problem is to simply acknowledge a residue of unpredictability, and to say, for example, that stress placement distinguishes between ‘regular’ and ‘irregular’ words, just as past tense formation does among verbs (regular *-ed* vs. irregular *swim–swam, put–put*, etc.), or plural formation among nouns (regular *-s* vs. irregular *sheep–*sheep, *man–men, cactus–cacti*, etc.). One can then be satisfied with identifying productive rules of stress assignment, while accepting that not all words reflect them. This view is in fact widely held (for example by Hayes 1982), and finds support from the fact that at least some of the ‘irregular’ stress patterns occur in loanwords like *hotel, vendetta* and *banana*. While representing a way of dealing with diversity among stress

² Members of pairs such as ‘*formal* – for’mal + ity, con’structν – *constructν* and ‘acid – a’cid + ic* each have their own fixed stress pattern. The ‘shifts’ proposed, for example, in Lexical Phonology (e.g. Kiparsky 1979) or stratal OT are procedural metaphors for static relationships.
patterns, however, the acknowledgement that not all words are regularly stressed fails to explain why this should be the case.

This paper attempts to do this. As indicated, we pursue the hypothesis that word stress does not primarily reflect properties of words themselves, but instead represents an adaptation to constraints that apply to phrases. There are several reasons why this hypothesis appears promising.

First, consider occasional phrase-level stress shifts in cases such as fifteen vs. fifteen years, Chinese vs. Chinese whispers, etc. This type of stress-pattern diversity has received considerable attention, and there is agreement that it reflects a conflict between lexically assigned stress and rhythmic well-formedness constraints on phrases (cf. e.g. Kiparsky 1979, Hayes 1984, Kager 1989). Of course, such shifts represent positive evidence that phrase-level rhythm can affect word stress. Emphatically, however, this does not mean that words whose stress never shifts (i.e. the majority) are immune to rhythmic phrase-level constraints. It is just as possible, and in fact more likely, we think, that their lexical stress patterns already reflect these constraints so well that stress shifts are simply not required.

The plausibility of this interpretation is corroborated by the fact that the structures built by established stress-assignment algorithms are in fact rhythmic units. For example, consider the rule that figures in most accounts of English word stress. Although it comes in various versions and contexts, the formulation in (2) (based on Hayes 1982: 238) is as good as any.

(2) English Stress Rule (ESR)

Starting from the end of the word, move leftwards and stress the first syllable if it is heavy (i.e. if it ends in VV or VC). If it is not, stress the next one, irrespective of its weight.

In combination with ‘extrametricality’ conventions (by which word-final consonants (as in (3a)) and sometimes word-final syllables (as in (3b)) are ignored – more on this below), rules like (2) adequately derive stress for many English items, such as those in (3), where extrametricality is indicated.

(3) a. in’habit, con’tain   b. sta’lactite, a’genda, ‘algebra

Essentially, the rule builds left-headed feet that are binary either in terms of syllables (e.g. [\textit{happy}]_{\Sigma}, [\textit{alge}]_{\Sigma}bra, in[\textit{habi}]_{\Sigma}t) or in terms of moras (e.g. [\textit{cool}]_{\Sigma}, sta[lac]_{\Sigma}tite, a[gen]_{\Sigma}da). This is even more explicit in functionally equivalent OT constraints, in which the term ‘foot’ is used, as in (4) (Prince & Smolensky 2004: 56).

(4) FootBinarity (FtBin)

Feet are binary at some level of analysis ($\mu, \sigma$).
That word structure should at all be rhythmic, however, is puzzling, because words cannot establish rhythm by themselves (cf. Vennemann 1986). Rhythm is ‘characterized as the repetition of patterned sequences of elements, often varying in prominence’ (Hay & Diehl 2007: 113), and most words are too short for sequences to be repeated in them. Repetition does not occur below the phrase level. While Vennemann (1986) noted this, it is being increasingly recognised. Lee & Gibbons, for instance, report ‘converging evidence that the principle of rhythmic alternation applies across words’, and ‘assume … that the domain within which the principle applies is the phonological phrase’ (2007: 449). Thus, if English words appear to build trochees, the reason is probably that they do this in response to constraints on the phrases they form when combined.

A third argument involves the phenomenon of ‘extrametricality’, mentioned above. Rules or constraints like those in (2) and (4) are not enough to predict English word stress adequately. In addition, they require a set of conventions by which final consonants, as well as the final syllables of English nouns and some adjectives, are made ‘invisible’ to stress-assignment algorithms (see e.g. Liberman & Prince 1977, Hayes 1982, Hyde 2007, 2011).

Thus, in verbs such as torment and inhabit – for which the stress rule in (2) would predict tor‘ment (correctly), but *inh’abit – the final /t/ is extrametrical, i.e. it ‘does not count’. This leaves the structures tor/ment and in/ha.bi. Since the last syllable in in/ha.bi is now light, it is skipped by the stress rule, and stress falls on the previous syllable, i.e. ha, yielding in’habit, as required. Additionally, in nouns and a number of (in general derived) adjectives, all final syllables are extrametrical. Thus, in words such as agenda, diplomat, adjective and nominal, the stress rules see only a.gen, di.plo, a.djec.ti and no.mi. They apply normally to these structures, for example stressing the heavy final gen in a.gen, but not the light final plo in di.plo, yielding, as expected, a.gen, ‘diplomat, ‘adjective and ‘nominal.

While extrametricality conventions may be useful, however, the question of why they should be required at all has received little attention. We don’t think this question can be answered without taking the phrase level into account.\footnote{Thus Giegerich (1985) suggests that the ‘extrametricality’ assumptions required at the right edge of words might reflect the impact of their potential phrase-level contexts.}

Consider word-final consonants, for example. Normally, a syllable that ends in a consonant counts as heavy, and there is no obvious \textit{a priori} reason why a rule like (2), which is sensitive to weight, should ignore it. In phrases, however, a word-final consonant may be followed a vowel – as in inhabit every continent. In such cases, it can be resyllabified, yielding, for example, [\textipa{\textipa{m.hæ.bi.tev.ri}}] from [\textipa{\textipa{m.hæ.bit.ev.ri}}]. This reduces the weight of the word-final syllable, rendering it light in the case of inhabit. At the same time, the first syllable of \textit{every} is stressed, and functions as a foot-head in phases like \textit{every ‘continent}. Thus, if \textit{inhabit} were stressed...
on the final syllable, phrases could arise in which a syllable that was light by resyllabification (in this case [bI]) formed a monosyllabic foot, as in *[in.hæ.ˈbu.ˈtev.rI]. Such a sequence is ungrammatical in English, and would be highly arhythmic, and its occurrence is effectively prevented by the strategy of disregarding final consonants in word-stress assignment. Thus the otherwise puzzling convention of final-consonant extrametricality serves a clearly definable purpose, and since it serves it so well, it is likely that this is actually the reason for its existence (cf. Dennett 1987). This suggests that the function of lexical stress-assignment algorithms is to adjust word structures for their use in rhythmically structured phrases, and that rhythmic constraints apply on the phrase level.

A similar argument can be put forward for the extrametricality of final syllables in English nouns. At least since Middle English, nominal stems have been followed by syllabic inflectional endings significantly less often than verbal stems (see Ritt 2012: 163). Kelly (1988, 1989) demonstrates this for Modern English as well. If verbal roots are systematically followed by one more unstressed syllable than nouns, however, they will also, on average, be further from the next potential foot head than nominal roots. If phrases are subject to rhythmic well-formedness constraints, this means that feet will prefer to have the same number of syllables. In order to achieve this, of course, it makes perfect sense for the stress in verbal roots to fall further to the right than in nominal roots. Since this is what final-syllable extrametricality in nouns achieves, it is once again likely that that is its purpose: it helps words to anticipate and to satisfy a rhythmic constraint that applies on the phrase level, in this case the preference for feet to be isochronous.

In sum, there is considerable, if indirect, evidence that some of the most important constraints determining word-stress patterns in the English lexicon apply to phrases, and that words reflect them only indirectly, because they anticipate their effects and adopt the structures best adapted to them.

The next question is, however, what follows from the observation that word-stress patterns are likely to reflect phrase-level constraints. An obvious consequence is that the assignment of stress in any specific word must depend on the stress patterns assumed by the words in its phrasal context, and of course, this applies to those words as well, i.e. the dependence is mutual. A theory of phrase-based word-stress assignment will therefore have to take mutual interdependencies between items into account. This means that in order to predict the stress pattern of any single word we need – at least in principle – to consider all the lexical items in a language, as well as the frequency with which each co-occurs with any other. This renders the task of predicting the stress patterns of individual items computationally too complex to be practicable. On the other hand, however, computational tools such as agent-based simulation and evolutionary game theory have been specifically designed for modelling the dynamics of populations in which the historical stability of their members depends on their interaction. Although such models say little
about the fates of individuals, they make predictions about the relative fre-
quencies of subpopulations. Since the distribution of stress patterns in the
lexicon can be construed in terms of subpopulations and their relative fre-
quencies, it is clearly addressable by means of such mathematical models.

3 An evolutionary approach to the distribution of stress
patterns in the lexicon

To make the problem addressable by evolutionary game theory, we con-
ceptualise it in evolutionary terms. For the purpose, we take a strictly rep-
licator-based view. Following in this respect Dawkins (1976), Croft
(2000), Ritt (2004) and Pierrehumbert (2012), we think of the lexicon as
a population of transmittable constituents, which acquire evolutionary sta-
bility if they are faithfully transmitted through communication and lan-
guage acquisition. In discourse, words combine to form phrases, and
these phrases are subject to rhythmic well-formedness constraints. The
better and the more often the expression of a word satisfies these con-
straints, the fitter, or more evolutionarily stable, its stress pattern will
be. Since phrases normally involve more than a single word, the fitness
of the stress pattern on any word therefore depends on the stress patterns
assumed by the words it combines with.

To see what we mean by this, consider a hypothetical language that con-
stitutes only of disyllabic CVCV items, so that any utterance will be a
sequence of [CVCV]<sub>λω</sub> [CVCV]<sub>λω</sub> [CVCV]<sub>λω</sub> [CVCV]<sub>λω</sub> and so on. Assume,
for the sake of the argument, that there is only a single criterion of rhyth-
mical well-formedness: strict rhythmic alternation is good, and any devi-
ation from it is bad, no matter whether it involves clashes (σσσσ) or lapses (sss).
It is easy to see that in such a language all words will end
up either as trochees or as iambs. As soon as a majority of word tokens
assumes one of the two patterns, the other will produce suboptimal rhyth-
mic structures more often than not. Thus words are forced into absolute
conformity.

What is also easy to see, however, is that predictions will be more
difficult to make as soon as things become even slightly more complex.
What distribution of stress patterns among lexical items would a prefer-
ce for alternating rhythm predict, for example, if a language includes
not only disyllables, but also lexical monosyllables (which normally
carry stress) and monosyllabic function words (which are usually
unstressed)? Questions like this have no intuitive answer, and require
formal modelling.

One of the tools for modelling interactions is evolutionary game theory
(some previous applications in linguistics are referred to in §4). It is not the
only tool for the purpose, but it is highly transparent, and compares
favourably in this respect to agent-based computer simulations. At the
same time, it places considerable restrictions on the number of variables
that can be taken into account, and requires problems to be formulated
in the simplest possible terms. Otherwise, they become intractable by analytical modelling. While computer simulations are more powerful in this respect, the very general way in which our problem can be formulated allows us to choose the more transparent option.

Since we are interested in a principled answer to a general question, we derive predictions from our hypothesis by modelling an artificial mini-language, which is simplified in many respects. It contains only monosyllabic and disyllabic items. Major class items can be monosyllabic or disyllabic, and will always carry lexical stress. Function words, which are always monosyllabic, never do. No further distinctions are made among disyllabic major class items. All of them have the same phonotactic structure, can be used as nouns, verbs or adjectives, and do not differ in frequency or style. The phrases that our game considers involve disyllabic items that are either adjacent or separated by a single monosyllabic item. That item may be either a lexically unstressed function word or a lexically stressed major class word.\(^4\)

The only items that have a choice with regard to their stress pattern are major class disyllables. Since they are all of the same phonotactic, morphological and syntactic type, they represent a completely uniform population. We also assume no lexical solidarities of any kind. Instead, all disyllables are equally likely to occur with any of the others.

As far as the rhythmic quality of utterance sequences is concerned, we employ only a single criterion, requiring strict alternation of stressed and unstressed syllables.\(^5\) Any deviation from that ideal pattern is considered suboptimal.

In any single round of the game, two disyllabic ‘players’ meet in one of the three syntactically possible contexts, i.e. either next to one another, or with a single lexical or functional monosyllable between them. Each disyllabic player can assume either initial or final stress as its ‘strategy’. Next, the rhythmic quality of the resulting stress pattern is evaluated, and a ‘pay-off’ is distributed among the players. That pay-off then determines the ‘fitness’ (and thereby the ‘evolutionary stability’) of the chosen stress pattern. Evolutionary game theory thus allows us to calculate which stress patterns, or strategies, will become stably established in the population of disyllables. In principle, the game can have one of the following outcomes: (a) all disyllables choose initial stress, (b) all disyllables choose final stress or (c) a specific mix of initially and finally stressed words turns out to be stable. Clearly, the relevant question is whether, and under what conditions, outcome (c) is produced.

\(^4\) Although the syntax of our mini-language allows them, we disregard phrases in which disyllables are separated by longer sequences of monosyllables, because we consider the disyllables to be too distant to exert a strong rhythmic influence on one another.

\(^5\) In Optimality Theory, this criterion involves two constraints, namely *Clash (no adjacent strong beats on the grid) and *Lapse (no two adjacent unstressed syllables). See Kager (2005).
Note that, for the purposes of our discussion, the simplicity of the artificial language is clearly an advantage. It allows us to be certain that the evolutionary dynamics which our model predicts do not result from factors other than the ones we have modelled. In contrast, real language usage always reflects a large number of different factors, some system-internal, others not. For example, transitive verbs might prefer final rather than penultimate stress, because they are more likely than intransitives to be followed by a noun phrase beginning with an unstressed article (see Fullwood 2014). Likewise, the stress pattern of a specific word might be due to lexical solidarities or formulaic sequences in which it is involved. Thus, initially stressed ‘research’ might be motivated by the frequency of such phrases as ‘research exercise’ or ‘research outcome’. The rise of initially stressed ‘hotel’ might reflect the popularity of the song ‘Hotel California’. The unexpected final stress in words such as café and settee might reflect a desire to mark them as foreign. Furthermore, in natural languages, lexical items are hardly ever fully equivalent to one another, even in terms of their phonotactic shape. Instead, they are usually distinguished by fine-grained differences in syllable weight, segmental structure, etc. The point is that our model prevents such factors from interfering. In a sense, abstracting away from them is a radical way of controlling for them, and makes it possible to isolate our central question in a ‘virtual lab’.

4 The stress game: structure and analysis

The tool we apply to our problem is a model of ‘replicator dynamics’ (see e.g. Hofbauer & Sigmund 1998), a specific version of evolutionary game theory (Maynard Smith & Price 1973, Smith 1982). Applying evolutionary game theory to problems in diachronic phonology or morphosyntax is not new per se. For example, Nowak and colleagues (Nowak & Komarova 2001, Nowak et al. 2002, Mitchener 2003, Mitchener & Nowak 2004, Nowak 2006) have studied the replicator dynamics of generative grammars in detail. In their models, entire language systems are construed as replicators that compete for existence in a population of speakers. Similarly, Yang (2000) and Niyogi (2006) have also proposed inherently generative dynamical systems models based on various learning algorithms.

Item-based replicator models, on the other hand, have been developed less frequently, although the concept of linguistic replicators, or ‘linguemes’ (Croft 2000), has gained currency during the last two decades (see also Ritt 2004, Jäger & Rosenbach 2008, Baxter et al. 2009, McCrohon 2012). An example of an item-based model is that of Jäger (2008), who investigates the dynamics of phonemes in the vowel space. He construes exemplars, i.e. stored phonetic events (cf. Pierrehumbert 2001, Wedel 2006), as linguistic replicators. Another item-based model…
is Nowak’s model of word dynamics (Nowak 2000, Nowak et al. 2000, Solé 2011). In that model, however, linguistic replicators reproduce independently of one another, so that it cannot simulate the interactions that we are interested in.

Not all dynamical systems approaches to phonological change are based on evolutionary game theory. Wang et al. (2004), for example, study Lotka-Volterra type dynamics in the phonological evolution of a lexicon (thereby implicitly including the dynamics of stress patterns). In principle, their model could take word co-occurrence on the utterance level into account, but they are primarily interested in the ‘snowball effect’, i.e. the accelerating spread of phonological change through the lexicon. Therefore, they do not address the problem of stable diversity (see however Sherman’s 1975 related discussion of the diffusion of noun–verb stress alternation). A dynamical systems model that does deal specifically with stress-pattern diversity, on the other hand, is proposed by Sonderegger & Niyogi (2013). It is speaker-based rather than item-based, and models the effects of ‘mistransmission’, i.e. the probability that a finally stressed word is recategorised as initially stressed, and vice versa. The model indeed predicts a stable mix of stress patterns.

In short, the methods employed in our study are well established in linguistic research, but have not been applied to the specific problem we discuss here. In the following, we briefly introduce evolutionary game theory (§4.1), and describe the specific game we have developed (§4.2). In §4.3 we analyse and interpret the game.

4.1 Evolutionary game theory

Evolutionary game theory is an extension of game theory, developed for studying the dynamics of strategy distributions among populations in a series of abstract games (Maynard Smith & Price 1973, Hofbauer & Sigmund 1998, Nowak 2006). In each game in the series, two players meet, select strategies and interact. Depending on the combination of strategies, both players receive a ‘pay-off’. In traditional game theory, the set of possible strategies is finite, and the pay-off is a predefined function allocating numerical values to strategy combinations.

Like all mathematical models, games are neutral regarding their empirical interpretation. Thus – in spite of the name – ‘players’ do not have to represent animate agents (e.g. fund managers opting to either buy or sell stock, or animals deciding whether to fight or to flee). In our model, the players are major class disyllables of the form σσ. The strategies they can adopt are initial stress ‘σσ (strategy S₁), and final stress σσ (strategy S₂),
and the pay-offs they receive reflect the rhythmic well-formedness of the stress patterns they build when they combine (see §4.2 for details).

Games can be visualised in pay-off matrices. When the same two strategies ($S_1$ and $S_2$) are available to each of the two players, the pay-off matrix $A$ is as in (5).

\[
A = \begin{pmatrix}
    a_{11} & a_{12} \\
    a_{21} & a_{22}
\end{pmatrix}
\]

That is, both players get pay-off $a_{11}$ if they choose $S_1$; if a player using $S_1$ meets a player using $S_2$, the former gets $a_{12}$ while the latter gets $a_{21}$; and if both players choose $S_2$, both get $a_{22}$.

As indicated, evolutionary game theory models player populations that play a series of games, rather than individuals encountering each other only once. Individual players from the population meet one another randomly. They do not actually ‘choose’ their strategies, but ‘inherit’ them from their progenitors and pass them on to their offspring in turn. Pay-off is then interpreted as reproductive success or fitness (Nowak 2006: 49). Players who receive higher pay-offs than their opponents produce more offspring, and as a result the relative frequency of their strategy in the population rises.

Since individuals are fully characterised by their strategies, populations can be divided into disjoint subpopulations, one for each strategy. Assuming a very large population, we denote the proportion of $S_1$ players by $x$ and the proportion of $S_2$ players by $y = 1 - x$. In this paper, $x$ denotes the proportion of initially stressed disyllables, and $y$ the proportion of finally stressed ones.

Changes in the proportions of $x$ and $y$ over time are modelled in terms of replicator dynamics (see Appendix A for a mathematically explicit description of the model). For games with two strategies, the long-term evolution of the population can be predicted from the pay-off matrix $A$. In particular, the conditions for a stable mix of $S_1$ and $S_2$ players can be determined: if $a_{11} < a_{21}$ as well as $a_{12} > a_{22}$, i.e. if pairs of words benefit if their members have different stress patterns, the replicator dynamics will converge to a stable internal equilibrium $\hat{x}_{\text{int}}$, with a fraction of $\hat{x}_{\text{int}} S_1$ players and a fraction of $1 - \hat{x}_{\text{int}} S_2$ players.

In the context of our discussion, the most important question is whether and under what conditions there will be a stable internal equilibrium in which two different stress patterns coexist. It is worth pointing out, however, that our game will predict more than just this.

Since our model predicts the evolutionary trajectory of fraction $x$ deterministically (on the basis of system-internal selection processes), it also predicts, for each initial distribution of strategies, how it will evolve over time. Also, if at any point in time a change in $x$ brought about by ‘external’ factors (i.e. factors that are not captured by the model itself), the model predicts how the fraction $x$ will evolve afterwards. In the case of stress patterns, system-external factors that affect their relative frequencies in the
On the replicator dynamics of lexical stress  453

4.2 Formulation of the stress game

The players in the game are disyllables that can choose either initial stress ($S_1$) or final stress ($S_2$). Additionally, the model includes monosyllables, but since they have no alternative regarding their stress pattern, they are not players. Lexical monosyllables are stressed; functional ones are not. Our model does not contain trisyllables or longer items, because the point of our exercise is to prove a principle, and disyllables suffice for this purpose. Furthermore, there are not many instances of words of more than two syllables in English: about 92.76% ($\pm 0.02$, 95% confidence interval) of all utterance tokens are either monosyllabic or disyllabic (see Fig. 1a – calculations based on the CELEX database; Baayen et al. 1995).

lexicon include language contact and, in particular, the incorporation of foreign vocabulary. We come back to this below.
In the game, two lexical disyllables meet to form sequences (Fig. 1b). They can meet in the three different contexts in (6).

(6) a. Context $C_1$: $[\sigma\sigma][\sigma][\sigma\sigma]$
   'father and 'mother, 'Susan and Mi'chelle, de'grees of 'freedom

b. Context $C_2$: $[\sigma\sigma][\sigma][\sigma\sigma]$
   'father 'loves 'mother, 'Susan 'hates Mi'chelle, de'grees 'count 'nothing

c. Context $C_3$: $[\sigma\sigma]\sigma[\sigma\sigma]$
   'father's 'lover, 'Susan 'worries, for'get Mi'chelle

Crucially, the contexts are taken to vary independently of the strategies chosen by the players (i.e. there are no rhythm-based solidarities). The relative frequencies of the contexts in which disyllables meet therefore depend exclusively on the relative frequencies of lexical and functional monosyllables. That distribution is denoted by $p = (p_1, p_2, p_3)$, where $p_i \in [0, 1]$ for all $i$ and $p_1 + p_2 + p_3 = 1$ (see Appendix B, which explains geometrical representations of the probability distributions and how to read them). $p_1$ is the proportion of contexts involving a functional (un-stressed) monosyllable, $p_2$ is the proportion of contexts involving a lexical (stressed) monosyllable and $p_3$ is the proportion of contexts without monosyllables (Fig. 1c). By varying the composition of $p$ in the model, it is possible to observe its effects on the stable distribution of stress patterns.

Lexical stress is immobile and cannot shift. Thus both players enter the game with their strategies fixed in advance, and the resulting sequences satisfy phrase-level constraints to varying degrees. As the only constraint in our model is a preference for strict stress alternation, two types of violation can occur: sequences of two stressed syllables (clashes), and sequences of two unstressed syllables (lapses). Pay-off is then determined by the number of violations in a sequence. For example, the sequence 'robust re'search contains one violation (a lapse), while robust 'stress 'research contains two (clashes). Thus, reward is modelled as a decreasing function $\phi$ of the number of violations. We reward optimal rhythmicity (no violation) with $\phi(0) = \phi_{\text{max}} > 0$, and attribute $\phi(2) = 0$ to the rhythmically worst sequences that can occur in our game (i.e. sequences with two violations). Crucially, these definitions leave the question open of how the occurrence of a single violation is rewarded. It is not obvious that it should count as half as bad as two violations. As will be seen, the issue is important for the interpretation of the game. If the first violation subtracts more from the overall rhythmic quality of a sequence than the second one, then $\phi(1) < (\phi_{\text{max}}/2)$, as in Fig. 2. If, on the contrary, the second violation subtracts more than the first one, then $\phi(1) > (\phi_{\text{max}}/2)$. In any case, we denote the normalised difference between no violation and one violation as $\Delta := (\phi_{\text{max}} - \phi(1))/\phi_{\text{max}}$.

As indicated, we disregard sequences where disyllables are separated by more than one monosyllable. This keeps the complexity of the model manageable, and reflects the plausible assumption that the influence that two disyllables exert on each other decreases with the distance between them.
Another issue is the following: while rhythmicity is assigned to whole sequences, pay-off must be allocated to single players, and it is not self-evident that it should be divided equally. This is because the contributions of words to the rhythmic quality of a foot they build together may not be perceived as equal by speakers. English, for example, counts as trochaic, and feet are perceived as beginning with a peak, and extending to the right. In the case of both clashes and lapses, it is the left foot that is felt to be deficient. Thus, when rhythm is repaired by shift, it normally affects the first word, not the second (e.g. ‘Pennsylvania Legislature’ rather than ‘Pennsylvania Legislature’). We therefore divide pay-off into two shares \( \alpha \times \varphi(\cdot) \), and \((1 - \alpha) \times \varphi(\cdot)\), where \( \alpha \) lies within the unit interval. For the purpose of modelling stress placement in English, we attribute a larger share of the pay-off to the first word, and a smaller share to the second one, so that \( \alpha > 0.5 \).

For each encounter of two words, we consider both possible orderings for calculating the pay-off received by each of the two words. Thus, when \([\sigma\sigma]_A\) and \([\sigma\sigma]_B\) meet, pay-offs are first determined separately for each of the sequences \([\sigma\sigma]_A \ldots [\sigma\sigma]_B\) and \([\sigma\sigma]_B \ldots [\sigma\sigma]_A\), and then added together.

A single round of the game for fixed \( \pi \), \( \varphi \) and \( \alpha \) unfolds as in (7).

---

9 Whether the first or the second word undergoes shift in cases of stress clash seems to reflect language-specific conventions. For evidence see Kiparsky (1966).

10 In principle, of course, \( \alpha \)-values below 0.5 are possible as well, say for applications of the model to languages other than English.
4.3 Evolutionary dynamics of the lexicon

4.3.1 When mixed stress patterns are evolutionarily stable. Our game allows us to model the evolution of a lexicon whose items interact in a series of independent rounds, as in (7). To identify the conditions under which a mix of initially stressed and finally stressed words is evolutionarily stable, we calculate the pay-off matrix and analyse it in terms of evolutionary game theory, as outlined in Appendix A.

The entries of the pay-off matrix depend on the distribution \( \pi \), the rhythmicity function \( \varphi \) and the weight \( \alpha \). The matrix can be expressed as in (8).

\[
A^{\pi,\varphi,\alpha} = \begin{pmatrix}
\alpha_1^{\pi,\varphi,\alpha} & \alpha_2^{\pi,\varphi,\alpha} \\
\beta_1^{\pi,\varphi,\alpha} & \beta_2^{\pi,\varphi,\alpha}
\end{pmatrix}
\]

After pay-off determination (explicitly described in Appendix C) we find (9).

\[
\begin{align*}
\alpha_1^{\pi,\varphi,\alpha} &= \alpha_2^{\pi,\varphi,\alpha} = (\pi_1 + \pi_2)(\varphi_{\text{max}} - \Delta \varphi_{\text{max}}) + \pi_3 \varphi_{\text{max}}, \\
\beta_1^{\pi,\varphi,\alpha} &= \pi_3 \varphi_{\text{max}} + \pi_2 (1 - \alpha) \varphi_{\text{max}} + \pi_3 (\varphi_{\text{max}} - \Delta \varphi_{\text{max}}) \quad \text{and} \\
\beta_2^{\pi,\varphi,\alpha} &= \pi_1 (1 - \alpha) \varphi_{\text{max}} + \pi_2 \alpha \varphi_{\text{max}} + \pi_3 (\varphi_{\text{max}} - \Delta \varphi_{\text{max}})
\end{align*}
\]

By inserting the entries \( \alpha_{ij}^{\pi,\varphi,\alpha} (i, j = 1, 2) \) into formula (20) in Appendix A, it can be shown that the replicator dynamics do indeed exhibit an internal equilibrium, which can be simplified as (10).

\[
\hat{x}_{\text{int}} = \frac{\pi_2 + \alpha (1 - 2 \pi_2 - \pi_3) - \Delta (1 - 2 \pi_3)}{1 - \pi_3 - \Delta (2 - 4 \pi_3)}
\]

Notice that the size of the internal equilibrium no longer depends on the maximal reward \( \varphi_{\text{max}} \), but only on directly interpretable variables. It is evolutionarily stable, if \( \alpha_1^{\pi,\varphi,\alpha} < \alpha_2^{\pi,\varphi,\alpha} \) as well as \( \beta_1^{\pi,\varphi,\alpha} > \beta_2^{\pi,\varphi,\alpha} \) (Appendix A). In this case, both the enumerator and the denominator in the fraction above are negative (cf. (20)). By replacing these variables with the corresponding terms from the pay-off matrix derived above, and solving both equalities
for $\pi_3$, it can be shown that if $\Delta > \frac{1}{2}$, $x_{\text{int}}$ is indeed an internal evolutionarily stable strategy (ESS) as long as the set of inequalities in (11) is fulfilled (note that $\Delta > \frac{1}{2}$ ensures that both denominators are positive).

\[
\begin{align*}
\text{(11a)} & \quad \pi_3 < \frac{\Delta - \pi_3 + 2\alpha\pi_2 - \alpha}{2\Delta - \alpha} \\
\text{(11b)} & \quad \pi_3 < \frac{\Delta + \pi_3 - 2\alpha\pi_2 - (1 - \alpha)}{2\Delta - (1 - \alpha)}
\end{align*}
\]

This means that the proportion of sequences in which two disyllables follow each other immediately defines a threshold for the existence of internal ESSs. In other words, stress-pattern diversity among disyllables will only be stable if there is a sufficiently high number of monosyllabic items. If these conditions are fulfilled, however, it is inevitable.\(^{11}\)

4.3.2 The roles of rhythmicity and foot structure in the evolution of stress patterns. While the prediction that stress-pattern diversity depends on the number of monosyllables is the most striking result of our simulation, it has other interesting implications as well. Thus, its prediction of inevitable stress-pattern diversity seems also to depend on two other variables. First, it depends on the distribution of pay-off shares between the two disyllables in a phrase (discussed above in §4.2). Secondly, it depends on the relative weighting of the number of violations of rhythmic well-formedness in a phrase, i.e. on the question of how much worse two violations of rhythmic well-formedness conditions are in comparison to a single one. This dependence can be demonstrated in the following way.

(11a) and (b) define a region in the 2-simplex, in which the internal ESS can be located. Its size depends on the syntagmatic weighting parameter $\alpha$, as well as on $\Delta$, illustrated in Fig. 3.\(^{12}\) This figure displays a family of bifurcation diagrams that plot the equilibria of the evolutionary dynamics, i.e. the stable as well as unstable steady-state distributions of initially and finally stressed items. The vertical axis in each of the diagrams measures the proportion of initially stressed items, i.e. $x$. The transparent upper and lower triangular planes represent populations in which all disyllables are either initially stressed (upper plane) or finally stressed (lower plane). The surfaces between the upper and lower planes represent states in which mixed populations are stable.

Figure 3 arranges the diagrams in a composite plot. In this plot, the horizontal dimension corresponds to the weight, $\alpha$, which determines the pay-off distribution between the first and the second disyllable. As can be seen

\(^{11}\) Note that our model (reassuringly) confirms the intuitively obvious fact (see §3) that languages that have only disyllabic items will be either purely trochaic or purely iambic. This results when $\pi_3 = 1$.

\(^{12}\) A reviewer observes that the case of $\Delta = 1$ corresponds to the generative practice of making a binary distinction between grammatical and ungrammatical sequences, while disregarding different degrees of ungrammaticality. Surprisingly, it is precisely this case that admits the largest number of stable configurations of mixed stress patterns (the support for $x_{\text{int}}$ in the 2-simplex is maximal if $\Delta = 1$). Thus a strictly binary conception of grammaticality actually favours stress-pattern diversity in the lexicon.
in the plot, $\alpha$ determines the slope of the surface of internal ESSs. This is also evident by taking the directional derivative of $\hat{x}_{\text{int}}$ in (10) with respect to $\pi_2$, i.e. the ratio of constellations involving lexical monosyllables,
assuming that \( \pi_3 \) is fixed. Since for the denominator in (10) it holds that 
\[ 1 - \pi_3 - \Delta(2 - 4\pi_3) < 0 \]
if \( \hat{x}_{\text{int}} \) is an internal ESS, we have (12).

\[
(12) \quad \frac{d\hat{x}_{\text{int}}}{d\pi_2} = \frac{1 - 2\alpha}{1 - \pi_3 - \Delta(2 - 4\pi_3)} \leq 0 \text{ iff } \alpha \leq \frac{1}{2}
\]

Thus, if the first word gets a larger share than the second (i.e. \( \alpha > 0.5 \)), then an increase in lexical (and stressed) monosyllables leads to an increase in \( \hat{x}_{\text{int}} \), i.e. the proportion of initially stressed disyllables. This is intuitive: if the foot built by the first word in a sequence counts as more important, then the frequent occurrence of stressed monosyllables between disyllables will more strongly favour trochaic patterns among preceding disyllables than iambic patterns among subsequent ones. The reverse holds, of course, for \( \alpha < 0.5 \). If both words have an equal share, the only possible internal ESS is exactly \( \hat{x}_{\text{int}} = 0.5 \). Notably, \( \alpha \)-values close to 0.5 produce internal ESSs, no matter whether monosyllables are lexical and stressed or functional and unstressed. Conversely, if \( \alpha \) is close to 0 or 1, i.e. if one of the two disyllables receives (nearly) all of the pay-off, the dynamics are more likely to produce pure equilibria corresponding to uniform stress placement. This is also intuitive: if monosyllables affect only one of their disyllabic neighbours, selection will favour either trochees or iambs, depending on which of their neighbours they affect, and on whether there are more lexical or more functional monosyllables. As discussed in §4.2, the assumption that we consider most plausible for English is that a larger share of the pay-off should be attributed to the first of two disyllables in a phrase (corresponding to the last two columns in Fig. 3).

The second implication of (11) is that, for both inequalities to hold, a double violation must be perceived as less than twice as bad as a single one, i.e. \( \Delta > \frac{1}{2} \), or equivalently, \( \varphi(1) < \frac{1}{2} \times \phi_{\text{max}} \) (see Fig. 2). This is also illustrated by the vertical dimension in Fig. 3. The closer \( \Delta \) gets to 0.5, the smaller the support for \( x_{\text{int}} \), i.e. the smaller the region in which initial stress and final stress stably coexist. This behaviour of the model is also intuitive: in sequences with monosyllables, two violations will only occur when the disyllables have different stress patterns, as in \([\sigma\sigma][\sigma\sigma] \) (a double clash), or in \([\sigma\sigma][\sigma\sigma][\sigma\sigma] \) (a double lapse). When both disyllables have the same stress pattern, no more than a single violation can ever be produced (\([\sigma\sigma][\sigma\sigma][\sigma\sigma], [\sigma\sigma][\sigma\sigma][\sigma\sigma], [\sigma\sigma][\sigma\sigma][\sigma\sigma], [\sigma\sigma][\sigma\sigma][\sigma\sigma] \)). If double violations are treated as much worse than single violations, however, the reward earned by single-violation sequences will be so much higher than the one earned by double-violation sequences that stress-pattern diversity will never be stable, even if the number of monosyllables in a language is high.

The issue requires discussion. Consider first the intuitive and apparently most straightforward assumption that two violations are perceived as being exactly twice as bad as a single one: although it reflects the simple arithmetic truth that \( 2 \times 1 = 2 \), it is not really safer to assume that the perception of rhythmic well-formedness follows the simple rules of
multiplication than that it does not. Thus the alternatives deserve serious consideration. There are two of them, as a double violation may be perceived as being either more or less than twice as bad as a single one. As will be argued below, we think there are good reasons to think that the latter is the case.

By definition, sequences that produce double violations involve three adjacent syllables that are all either stressed or unstressed, as in (13b).

(13) a. 1 violation:  
*CLASH Mi'chelle’s ‘father
*LAPSE  ‘Susan and ‘Mike

b. 2 violations:  
*CLASH Mi'chelle’s ‘old ‘father
*LAPSE  ‘Susan and Mi’chelle

Crucially, the rhythmic quality of such sequences can be improved in a way that is not available to single violations, namely by demoting or lifting the middle syllable. Thus, Mi’chelle’s ‘old ‘father may optionally be pronounced (or perceived) as Mi’chelle’s ‘old ‘father, and ‘Susan and Mi’chelle as ‘Susan and Mi’chelle.13 That is to say, the demotion of lexically stressed old and the promotion of lexically unstressed and can restore rhythmic quality to some extent, albeit at the cost of backgrounding a content word, or foregrounding a functional one. Evidence that such adjustments do naturally occur are easily found in verse, for example (see also Attridge 1982). Consider the lines from Shakespeare in (14), in which lexical fair is demoted, and grammatical of promoted so that viable pentameters are produced.14

(14) a. A’rise, fair ‘sun, and ‘kill the ‘envious ‘moon  
Romeo and Juliet 2.1.5

b. ‘Now is the ‘winter ,of our ‘discon’tent  
Richard the Third 1.1.1

What is crucial about such repairs is that they are not equally available in the case of single violations. Thus, the single clash in Mi’chelle’s ‘father could clearly be optimised by shifting the stress in Michelle to the first syllable, yielding the perfectly alternating sequence ‘Michelle’s ‘father. However, the pay-off incurred by that sequence could no longer be attributed to the pattern Mi’chelle’s. The single lapse in ‘Susan and ‘Mike could of course be repaired by lifting and, but this would then produce a clash between lifted and and Mike. Thus any repair of single violations through rhythmic promotion or demotion would either affect the players

13 The extent to which rhythmicity represents an acoustic property of the speech signal or a perceptual construct is an interesting and not fully resolved issue, but does not affect our argument.

14 A reviewer pointed out that strategies of rhythmic demotion and promotion may themselves have emerged in response to the way in which English lexical stress patterns have come to be assigned. This is a fascinating hypothesis about co-evolution and co-adaptation among different components of linguistic systems. However, adding another variable to our model would increase its complexity significantly, so this idea will have to be taken up in later research.
themselves or produce another violation. Therefore, their effects cannot be mitigated in the same way as those of double violations.

We conclude that double violations can plausibly be considered as less than twice as bad than single ones, even though they should still count as worse, since (a) demotion and promotion are only optional, and (b) they come at the cost of (unnaturally) backgrounding a content word, or foregrounding a functional one. Thus, we assume that $\Delta > \frac{1}{2}$, and hence $\varphi(1) < \frac{1}{2} \times \varphi_{\text{max}}$.

4.3.3 Summary. We can now summarise our observations, and draw some initial conclusions. As we have seen, our game does indeed produce stable internal equilibria. Thus it predicts that, under specific conditions, stress-pattern diversity will inevitably be established in a lexicon, answering one of our central questions. If lexical stress assignment reflects constraints on the rhythmic quality of the utterances, then there are indeed conditions under which words of the same phonotactic structure and the same morphosyntactic class will necessarily display a variety of different stress patterns. This will happen irrespective of any possible further motivations, such as differences among words in terms of syllable weight, collocational preferences, etymological origins or stylistic values. Thus the problem of existing theories of word-stress assignment, i.e. that they invariably face a subset of words for which stress can be predicted only probabilistically, has received a principled account. In addition, our model has allowed us to derive the clear hypotheses in (15) about the conditions under which stress-pattern diversity will arise.

(15) Mixed stress patterns are evolutionarily stable if:
   a. $\Delta > \frac{1}{2}$, i.e. if two violations (i.e. a double clash or a double lapse) are perceived as less than twice as bad as a single violation, and
   b. $\pi_1$ is sufficiently small, i.e. if there are relatively few sequences of two neighbouring disyllables.

Furthermore, the following holds:
   c. In case a stable mixed equilibrium $\tilde{\xi}_{\text{int}}$, measuring the fraction of initially stressed items exists, then it increases (decreases) in the fraction of sequences involving monosyllabic lexical items if $\alpha > \frac{1}{2}$ ($\alpha < \frac{1}{2}$), i.e. if in a pair of two subsequent disyllables the first (second) gets a larger share of the received pay-off.

We have provided arguments for condition (a). Condition (b) also clearly holds in English, where the proportion of monosyllabic items exceeds the proportion of all polysyllables taken together (see Fig. 1). Finally, condition (c) also seems to hold in English: the first word in a sequence is perceived to be more responsible for its rhythmic quality than the second one, and receives a greater share of the pay-off. We therefore conclude that our initial hypothesis is plausible: word-stress assignment reflects an adaptation to constraints on the rhythm of phrases, and
this can explain stress-pattern diversity in the English lexicon. In order to corroborate the hypothesis, in §5 we consider whether the predictions it implies for the diachrony of stress pattern distribution match the evolution of word stress in English.

4.3.4 Related models: learning algorithms modelling mistransmission. Before we focus on the history of English word stress, however, we briefly compare the findings of our model with studies by Sonderegger & Niyogi (2010, 2013), in which English stress is also approached on the basis of dynamical systems theory. In contrast to our own game, they model the distribution of stress patterns in a population of speakers rather than in an evolving lexicon, using a learning algorithm that is driven by probability matching and focusing on the potential effects of mistransmission. Reassuringly, Sonderegger & Niyogi’s model converges with ours in many respects. This is particularly so in the case of $x_{\text{int}}$ being an ESS, where our model predicts the same as theirs if the respective pay-off differences between choosing the same strategy as the competing player and choosing the converse strategy are interpreted as Sonderegger & Niyogi’s mistransmission rates. In their model, $a$ is the rate of misinterpreting finally stressed items as initially stressed ones, and $b$ the rate of misinterpreting initially stressed items as finally stressed ones.\footnote{Note that, in contrast to Sonderegger & Niyogi’s (2013) model, the parameters $a$ and $b$ are conceptualised as rates rather than as probabilities, since the present replicator dynamics represent a dynamical system in continuous time. Moreover, it is worth pointing out that Sonderegger & Niyogi actually studied the word-stress dynamics of single lexical items (e.g. specific nouns and verbs), rather than entire sets of words of the same phonotactic and morphosyntactic type. This is, however, only a matter of reinterpreting the variables in the dynamical system, which is not a significant problem: instead of single words we can just as well look at classes of morphosyntactically and phonotactically equivalent items.} Sonderegger & Niyogi (2013: 277–278) show that there is a stable internal equilibrium at $x_\ast = b / (a + b)$, where $x_\ast$ measures the mean probability of a speaker using word-final stress. If we now set $a = (a_{12}^{\ast} - a_{21}^{\ast}) \times C > 0$ and $b = (a_{21}^{\ast} - a_{11}^{\ast}) \times C > 0$, where $C$ is some positive constant, we have (16).

\begin{equation}
(16) \quad x_\ast = \frac{a_{21}^{\ast} - a_{11}^{\ast}}{a_{12}^{\ast} - a_{22}^{\ast} + a_{21}^{\ast} - a_{11}^{\ast}} = 1 - \frac{a_{21}^{\ast} - a_{22}^{\ast}}{a_{12}^{\ast} - a_{11}^{\ast} + a_{21}^{\ast} - a_{11}^{\ast}}
\end{equation}

Assuming that the population of speakers, and thus also the population of utterances, is homogeneously mixed, it follows that $x_\ast = \bar{y}_{\text{int}} = 1 - \bar{x}_{\text{int}}$, so that, due to (20) in Appendix A, their model and ours yield equivalent equilibria.

Although the two models are equivalent in these respects, there are also differences. For example, Sonderegger & Niyogi (2013) say little about the factors that actually condition mistransmission rates. In contrast, our model makes very specific proposals about the conditions under which one stress pattern is transmitted more successfully than its alternative. In that sense, it not only corroborates Sonderegger & Niyogi’s findings,
but also elaborates them. Also, the game-theoretic approach taken here is more extensive, and covers the dynamics modelled by the mistransmission-based system as a special case among various possible scenarios. The strength of the replicator equation is that it can account not only for stable coexistence of stress patterns, but also for scenarios in which one of them becomes dominant, as well as for bistability among stress patterns (see Appendix A). To change the qualitative behaviour of their model so that bistability becomes possible, Sonderegger & Niyogi (2013: 279–280) need to resort to a different underlying learning mechanism, which in turn does not admit stable coexistence. Thus the game-theoretic approach taken here is more powerful in a number of ways, and can account for a variety of evolutionary dynamics with a single set of internal mechanisms. Which of the possible dynamics actually unfolds not only depends on properties of our model as such, but follows from differences between languages, or language states, such as the relative frequencies of monosyllabic and disyllabic words. As we shall see below, this makes it possible to derive testable, albeit general, hypotheses.

5 Further corroboration from diachronic evidence

Our model shows that the stability of stress-pattern diversity depends on the number of monosyllables in the lexicon. This implies the prediction that stress-pattern distribution will be affected by diachronic changes in the monosyllabicity of a language if such changes occur.

In the following, we show that the specific predictions that can be derived from our game match the long-term evolution of English word stress in a way that we consider encouraging. The history of English stress has of course been investigated widely, and in the context of this paper we can provide only a global summary.

Old English stress was uniformly root-initial (Minkova 1997: 137, Hutton 1998). In addition, Old English morphology was highly inflecting (at least in comparison to Middle and Modern English), so that it contained a large proportion of polysyllabic word forms. However, during later Old English, unstressed inflectional syllables underwent reduction and deletion processes, leading to a gradual but marked increase in the number of monosyllables. Thus what would have been a stable pure equilibrium in which initial stress dominated would at some point have become evolutionarily unstable, although initial stress may still have been the rule.

16 These are the scenarios in which one of the inequalities $a_{11}^{\pi_1, \alpha} > a_{21}^{\pi_1, \alpha}$ and $a_{13}^{\pi_1, \alpha} < a_{23}^{\pi_1, \alpha}$ is fulfilled (or both in the case of bistability). Crucially, there are settings for the defining parameters $\pi$, $\varphi$ and $\alpha$ that entail one or other inequality. For example, $\pi_1 = 1$ entails that final stress will dominate, $\pi_2 = 1$ leads to domination of initial stress and $\pi_3 = 1$ immediately implies bistability.

17 The same holds for other models of evolutionary language dynamics. For example, Yang’s (2000) speaker-based diachronic model driven by variational learning only allows for either initial stress or final stress to be stable and attracting, but cannot account for either bistability or stable coexistence.
If a pure equilibrium of initial stress is unstable, however, this means that a lexicon will cease to display pure initial stress, if it adopts words that display final stress. This is exactly what happened in English. Finally stressed loans from French and Latin were adopted in large numbers, starting in the wake of the Norman Conquest and continuing well into the Modern period (cf. e.g. Lass 1992, Minkova 1997, Dresher & Lahiri 2005). Although many early Romance loans quickly adopted the native initial stress pattern (Dresher & Lahiri 2015), they produced considerable surface variation, as can be seen for example in the works of Chaucer in stress doublets like those in (17).

(17) [ci.té]_N ~ [ci.'té]_N
    [di.vers]_A ~ [di.'vers]_A
    [hus.bond]_N ~ [hus.'bond]_N
    [A.pril]_N ~ [A.'prile]_N
    [out.lawe]_N ~ [out.'lawe]_N
    [ho.nour]_N ~ [ho.'nour]_N
    [for.tune]_N ~ [for.'tune]_N
    [ser.vise]_N ~ [ser.'vise]_N

Eventually, however, surface variability was once again reduced, and the English principle of assigning stress to words lexically and keeping its position fixed when the words are uttered reasserted itself. Items whose stress patterns had varied on the surface came to adopt one specific pattern in their lexical representation, and would maintain it in most of their realisations. Crucially, however, the stress patterns that came to be lexicalised did not all reflect the same stress-assignment principles, so that French loans like *cité* or *beauté* ended up with initial stress, i.e. as *city* and *beauty*, while others, such as *de'gree*, retained final stress. Ever since, the English lexicon has continued to contain items that are equivalent to one another both in terms of phonotactic structure and in terms of word class, but are nevertheless stressed differently. The only changes that have been observable during the modern period involve frequent, but not system-wide, shifts of stress towards the left, for example in words like *address* (< *ad'dress*), *balcony* (< *bal'cony*), *compact*<sub>A</sub> (< *com'pact*) and recent *hotel* (< *ho'tel*) (see also Minkova 1997).

As Fig. 4 demonstrates, the specific developments that English stress patterns have undergone correspond extremely well to the evolutionary dynamics predicted by our model for a language in which the number of (lexical) monosyllables increases over time. The bifurcation diagram in Fig. 4 is shown from two different perspectives. It can be read just like the diagrams in Fig. 3 above. The upper and lower planes represent

---

18 Dresher & Lahiri (2015) provide evidence that Early Modern loans from Latin had a greater impact on the principles of English stress assignment than Middle English ones from French. While revising established views in this respect, however, they do not dispute that the English lexicon has continuously displayed stress-pattern diversity since the Middle English period.
pure populations (initial stress on the upper plane, final stress on the lower one), and the twisted surface in the middle represents mixed populations. Dark grey areas represent evolutionarily stable equilibria; light grey areas unstable ones. What is particularly relevant in Fig. 4 are the tips of the two triangles that are marked $\pi_3 = 1$, i.e. $\pi = (0, 0, 1)$. They represent equilibria in which there are no monosyllables at all. As we
move away from them, the number of sequences with monosyllables increases: grammatical ones if we move towards \( p_1 = 1 \), and lexical ones if we move towards \( p_2 = 1 \).

Now, think of English as starting at point (1) on the upper planes in Fig. 4a and 4b, somewhere in the dark grey area. This point characterises a language with pure and evolutionarily stable initial stress, and corresponds to the situation that obtained in Old English. Next, assume a move of \( \pi \) away from \((0, 0, 1)\), i.e. assume that the number of monosyllables increases, as indeed happened in the development of Old English through the phonetic erosion of final syllables and the concomitant simplification of the inflectional system.\(^{19}\) At first, this does not produce variable stress, but as the language moves into the light grey area (2) on the upper surface, the state of pure initial stress becomes evolutionarily unstable. Being unstable, however, the dominance of pure initial stress will come to a catastrophic end as soon as a number of finally stressed items enter the population. Again, this is indeed what occurred in English when French and Latin loanwords were adopted in increasing numbers, from the beginning of the 12th century onwards. As soon as this happens, our model predicts that a mix of initially and finally stressed words should become evolutionarily stable, as the population drops to land near point (3) on the twisted surface inhabited by mixed populations (visible only in Fig. 4b).

Finally, our model predicts that, if the number of lexical monosyllables keeps increasing, the population will evolve steadily back towards the upper plane, i.e. towards a situation in which purely initial stress once again becomes the most frequently employed strategy for assigning stress (4). Once more, this seems to be exactly what has been going on in English during the last centuries: as monosyllabicity increased among lexical words (see Jespersen 1912, 1928), more and more polysyllabic words underwent stress shifts towards the left (\('address_N < ad'dress, etc.\).)

Since our model was originally designed to identify the conditions under which stress-pattern diversity will be stable in a language such as English, the fact that it also predicts the dynamics of its historical evolution represents an independent and rather strong corroboration of the hypotheses underpinning the model.

6 Conclusion and outlook

Focusing on English, we have tried to account for the existence of variable stress patterns among words of the same phonotactic and morphosyntactic types. Although we have dealt with word stress, we have departed from the tradition of deriving it from the properties of individual lexical items in

\(^{19}\) In functional words such as \textit{ane} ‘one’ \( > \text{a(n)} \), phonological reduction together with morphological simplification increased the number of unstressed monosyllables; in lexical words such as \textit{gode} \( > \text{god} \) it increased the number of stressed ones.
isolation. Instead, we have assumed that word stress reflects rhythmic constraints on utterance sequences rather than on individual items, and have modelled this approach in terms of game theory.

Our model specifies conditions under which stress-pattern diversity will inevitably emerge, and we have shown that these conditions do indeed hold in English. We have thereby accounted for a set of empirical facts that have previously been acknowledged, but not explained.

More specifically, we have revealed a causal connection between the degree to which languages require the stress patterns of polysyllabic items to be variable and the number of monosyllables in a language. Building on this, we have derived predictions about the evolutionary dynamics of stress systems, and shown that the predictions for English correspond well with the actual developments that English word stress has undergone during the last millennium.

Although our model remains abstract and general, we have demonstrated its explanatory power, and shown how it can enrich and deepen established theories of word stress. We have also demonstrated the general plausibility of our approach. Naturally, there is room for refinement and elaboration, in order to make the predictions our model implies more specific. For instance, we might investigate the potential effects of differential pay-offs for lapses and clashes, incorporate a greater variety of word structures in terms of syllable counts or allow for the possibility of specific interactions among different morphosyntactic classes. Even in its present state, however, our model is sufficiently robust to act as a basis for further research, and for providing accounts of similar problems in other languages.

REFERENCES

Altmann, Eduardo G., Janet B. Pierrehumbert & Adilson E. Motter (2009). Beyond word frequency: bursts, lulls, and scaling in the temporal distributions of words. *PLoS ONE* 4. Available at https://doi.org/10.1371/journal.pone.0007678.

Anderson, John M. (1986). Suprasegmental dependencies. In Jacques Durand (ed.) *Dependency and non-linear phonology*. London: Croom Helm. 55–133.

Attridge, Derek (1982). *The rhythms of English poetry*. London & New York: Longman.

Baayen, R. Harold, Richard Piepenbrock & Leon Gulikers (1995). *The CELEX lexical database*. Philadelphia: Linguistic Data Consortium. https://catalog.ldc.upenn.edu/ldc96l14.

Baxter, Gareth J., Richard A. Blythe, William Croft & Alan J. McKane (2009). Modeling language change: an evaluation of Trudgill’s theory of the emergence of New Zealand English. *Language Variation and Change* 21. 257–296.

Blevins, Juliette (2004). *Evolutionary Phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Boer, Bart de (2000). Self-organization in vowel systems. *JPh* 28. 441–465.

Boer, Bart de (2002). Evolving sound systems. In Angelo Cangelosi & Dominic Parisi (eds.) *Simulating the evolution of language*. London: Springer. 79–97.

Brighton, Henry, Kenny Smith & Simon Kirby (2005). Language as an evolutionary system. *Physics of Life Reviews* 2. 177–226.
Andreas Baumann and Nikolaus Ritt

Burzio, Luigi (1994). *Principles of English stress*. Cambridge: Cambridge University Press.

Bybee, Joan (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, Joan, Paromita Chakraborti, Dagmar Jung & Joanne Scheibman (1998). Prosody and segmental effect: some paths of evolution for word stress. *Studies in Language* 22, 267–314.

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

Christiansen, Morten H. & Nick Chater (2008a). Language as shaped by the brain. *Behavioral and Brain Sciences* 31, 489–509.

Christiansen, Morten H. & Nick Chater (2008b). Brains, genes, and language evolution: a new synthesis. *Behavioral and Brain Sciences* 31, 537–558.

Croft, William (2000). *Explaining language change: an evolutionary approach*. Harlow: Longman.

Dawkins, Richard (1976). *The selfish gene*. Oxford: Oxford University Press.

Dennett, Daniel C. (1987). *The intentional stance*. Cambridge, Mass.: MIT Press.

Domahs, Ulrike, Ingo Plag & Rebecca Carroll (2014). Word stress assignment in German, English and Dutch: quantity-sensitivity and extrametricality revisited. *Journal of Comparative Germanic Linguistics* 17, 59–96.

Donegan, Patricia J. & David Stampe (1979). The study of natural phonology. In Daniel A. Dinnsen (ed.) *Current approaches to phonological theory*. Bloomington: Indiana University Press. 126–173.

Donegan, Patricia J. & David Stampe (1983). Rhythm and the holistic organization of language structure. In John F. Richardson, Mitchell Marks & Amy Chukerman (eds.) *Papers from the parasession on the interplay of phonology, morphology, and syntax*. Chicago: Chicago Linguistic Society. 337–353.

Donegan, Patricia J. & David Stampe (2009). Hypotheses of Natural Phonology. *Poznań Studies in Contemporary Linguistics* 45, 1–31.

Dressler, B. Elan & Aditi Lahiri (2005). Main stress left in Early Middle English. In Michael Fortescue, Eva Skafte Jensen, Jens Erik Mogensen & Lene Schosler (eds.) *Historical linguistics 2003: selected papers from the 16th International Conference on Historical Linguistics, Copenhagen, 11–15 August 2003*. Amsterdam & Philadelphia: Benjamins. 75–85.

Dressler, B. Elan & Aditi Lahiri (2015). Romance loanwords and stress shift in English: a quantitative approach. Paper presented at the 2nd Edinburgh Symposium on Historical Phonology, University of Edinburgh. Available (August 2017) at http://homes.chass.utoronto.ca/~dresher/talks/Drescher-Lahiri-ESohPh_talk_pub.pdf.

Dressler, Wolfgang U. (1985). *Morphology: the dynamics of derivation*. Ann Arbor: Karoma.

Dziubalska-Kołaczyk, Katarzyna (2009). NP extension: B&B phonotactics. *Poznań Studies in Contemporary Linguistics* 45, 55–71.

Fitch, W. Tecumseh (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience* 7:68, 1–16.

Fullwood, Michelle A. (2014). Asymmetric correlations between English verb transitivity and stress. *BLS* 40, 125–138.

Giegerich, Heinz J. (1985). *Metrical phonology and phonological structure: German and English*. Cambridge: Cambridge University Press.

Halle, Morris (1998). The stress of English words 1968–1998. *LI* 29, 539–568.

Hay, Jessica S. F. & Randy L. Diehl (2007). Perception of rhythmic grouping: testing the iambic/trochaic law. *Perception and Psychophysics* 69, 113–122.

Hayes, Bruce (1982). Extrametricality and English stress. *LI* 13, 227–276.
Hayes, Bruce (1984). The phonology of rhythm in English. *LI* 15. 33–74.

Hofbauer, Josef & Karl Sigmund (1998). *Evolutionary games and population dynamics*. Cambridge: Cambridge University Press.

Hutton, John (1998). Stress in Old English, *giet ongean*. *Linguistics* 36. 847–885.

Hyde, Brett (2007). Non-finality and weight-sensitivity. *Phonology* 24. 287–334.

Hyde, Brett (2011). Extrametricality and non-finality. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*. Malden, Mass.: Wiley-Blackwell. 1027–1051.

Jäger, Gerhard (2008). Applications of game theory in linguistics. *Language and Linguistics Compass* 2. 406–421.

Jäger, Gerhard & Anette Rosenbach (2008). Priming and unidirectional language change. *Theoretical Linguistics* 34:2. 85–113.

Jespersen, Otto (1912). *Growth and structure of the English language*. 2nd edn. Leipzig: Teubner.

Jespersen, Otto (1928). Monosyllabism in English. *Proceedings of the British Academy* 14. 341–368.

Kager, René (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kager, René (2005). The factorial typology of rhythmic licensing constraints. *Phonological Studies* 5. 147–155.

Kelly, Michael H. (1988). Rhythmic alternation and lexical stress differences in English. *Cognition* 30. 107–137.

Kelly, Michael H. (1989). Rhythm and language change in English. *Journal of Memory and Language* 28. 690–710.

Kiparsky, Paul (1966). Über den deutschen Akzent. *Studia Grammatica* 7. 69–98.

Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *LI* 10. 421–441.

Lass, Roger (1992). Phonology and morphology. In Norman Blake (ed.) *The Cambridge history of the English language*. Vol. 2: 1066–1476. Cambridge: Cambridge University Press. 23–155.

Lass, Roger (1997). *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Lee, Ming-Wei & Julie Gibbons (2007). Rhythmic alternation and the optional complementiser in English: new evidence of phonological influence on grammatical encoding. *Cognition* 105. 446–456.

Liberman, Mark & Alan Prince (1977). On stress and linguistic rhythm. *LI* 8. 249–336.

Maynard Smith, J. & G. R. Price (1973). The logic of animal conflict. *Nature* 246. 15–18.

McCrohon, Luke (2012). The two-stage life cycle of cultural replicators. *Theoria et Historia Scientiarum* 9. 149–170.

McMahon, April M. S. & Robert McMahon (2013). *Evolutionary linguistics*. Cambridge: Cambridge University Press.

Minkova, Donka (1997). Constraint ranking in Middle English stress-shifting. *English Language and Linguistics* 1. 135–175.

Mitchener, W. Garrett (2003). Bifurcation analysis of the fully symmetric language dynamical equation. *Journal of Mathematical Biology* 46. 265–285.

Mitchener, W. Garrett & Martin A. Nowak (2004). Chaos and language. *Proceedings of the Royal Society of London B* 271. 701–704.

Mühlenbernd, Roland & Johannes Wahlhe (2016). Self-organization in sound systems: a model of sound strings processing agents. In Séan Roberts, Christine Cuskeley, Luke McCrohon, Lluis Barceló-Coblijn, Olga Feher & Tessa Verhoef (eds.) *Proceedings of the 11th International Conference on the Evolution of Language*. Available (August 2017) at [http://evolang.org/neworleans/papers/42.html](http://evolang.org/neworleans/papers/42.html).

Niyogi, Partha (2006). *The computational nature of language learning and evolution*. Cambridge, Mass.: MIT Press.
Nowak, Martin A. (2000). The basic reproductive ratio of a word, the maximum size of a lexicon. *Journal of Theoretical Biology* **204**, 179–189.

Nowak, Martin A. (2006). *Evolutionary dynamics: exploring the equations of life*. Cambridge, Mass. & London: Belknap.

Nowak, Martin A. & Natalia L. Komarova (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences* **5**, 288–295.

Nowak, Martin A., Natalia L. Komarova & Partha Niyogi (2002). Computational and evolutionary aspects of language. *Nature* **417**, 611–617.

Nowak, Martin A., Joshua B. Plotkin & Vincent A. A. Jansen (2000). The evolution of syntactic communication. *Nature* **404**, 495–498.

Oostendorp, Marc van (2005). Stress and metrical structure. Available (August 2017) at http://www.vanoostendorp.nl/pdf/051115.pdf.

Paul, Hermann (1920). *Prinzipien der Sprachgeschichte*. 5th edn. Halle: Niemeyer.

Piétrehumbert, Janet B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: Benjamins. 137–157.

Piétrehumbert, Janet B. (2012). The dynamic lexicon. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.) *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press. 173–183.

Ritt, Nikolaus (2004). *Selfish sounds and linguistic evolution: a Darwinian approach to language change*. Cambridge: Cambridge University Press.

Ritt, Nikolaus (2012). Middle English phonology in the digital age: what written corpora can tell us about sound change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.) *The Oxford handbook of the history of English*. Oxford: Oxford University Press. 81–86.

Ross, Don (2016). Game theory. In Edward N. Zalta (ed.) *Stanford encyclopedia of philosophy* archive (Winter 2016 edition). https://plato.stanford.edu/archives/win2016/entries/game-theory.

Schane, Sanford (2007). Understanding English word accentuation. *Language Sciences* **29**, 372–384.

Schlüter, Julia (2005). *Rhythmic grammar: the influence of rhythm on grammatical variation and change in English*. Berlin & New York: Mouton de Gruyter.

Sherman, Donald (1975). Noun-verb stress alternation: an example of the lexical diffusion of sound change in English. *Linguistics* **13**, 43–72.

Smith, John Maynard (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Smith, Kenny & Simon Kirby (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society of London*. B **363**, 3591–3603.

Solé, Ricard V. (2011). *Phase transitions*. Princeton: Princeton University Press.

Sonderegger, Morgan & Partha Niyogi (2010). Combining data and mathematical models of language change. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. 1019–1029.

Sonderegger, Morgan & Partha Niyogi (2013). Variation and change in English noun/verb pair stress: data and dynamical systems models. In Alan C. L. Yu (ed.) *Origins of sound change: approaches to phonologization*. Oxford: Oxford University Press. 262–284.

Stampe, David (1979). *A dissertation on Natural Phonology*. New York: Garland.

Tallerman, Maggie & Kathleen R. Gibson (eds.) (2012). *The Oxford handbook of language evolution*. Oxford: Oxford University Press.

Vennemann, Theo (1986). *Neuere Entwicklungen in der Phonologie*. Berlin: Mouton de Gruyter.
Wang, William S.-Y., Jinyun Ke & James W. Minett (2004). Computational studies of language evolution. In Chu-Ren Huang & Winfried Lenders (eds.) Computational linguistics and beyond. Taiwan: Academia Sinica. 65–108.

Wedel, Andrew (2006). Exemplar models, evolution and language change. The Linguistic Review 23, 247–274.

Wedel, Andrew (2009). Resolving pattern conflict: variation and selection in phonology and morphology. In James P. Blevins & Juliette Blevins (eds.) Analogy in grammar: form and acquisition. Oxford: Oxford University Press. 83–100.

Wolfram Research, Inc. (2016). Mathematica. Version 11.0. Software package. Champaign, Ill. http://www.wolfram.com/support.

Yang, Charles D. (2000). Internal and external forces in language change. Language Variation and Change 12. 231–250.

On the replicator dynamics of lexical stress