Segment count and weight in *y*-adjective comparatives: Inroads that bite off more than one can chew![1]

DEBORAH CHUA

National Institute of Education, Nanyang Technological University

Author correspondence:

Deborah Chua

National Institute of Education

Nanyang Technological University

NIE3-B1-19, 1 Nanyang Walk, Singapore 637616

Singapore

Email: debchua2004@gmail.com; deborah.chua@nie.edu.sg

[1] I would like to thank Professor Laurie Bauer for his insightful feedback and suggestions on previous drafts of this article and for his help in verifying some of my transcriptions. I am grateful to Professor Paul Warren for having introduced me to mixed effects modelling. I would also like to thank two reviewers of *English Language and Linguistics* for their helpful comments on versions of this article. The initial data and thinking that led to this article would not have come about without my PhD candidature at Victoria University of Wellington (VUW), New Zealand, and I am grateful to VUW for having supported my candidature.

ABSTRACT

Adjectival syllable count, often used to predict English comparatives *more* versus *-er*, is of little help in predicting the comparatives of adjectives ending in <y>, pronounced /i/, here called the *y*-adjectives. Examples of y-adjectives include *silly* and *worthy*. This article considers whether the phonemic segment count (segment count) and penultimate syllable weight (penultimate weight) of *y*-adjectives may serve as alternatives to syllable count in predicting *more* versus *-er*. The segment count and penultimate weight of relevant *y*-adjective tokens from a set of diachronic corpora are studied, alongside the tokens' morphological complexity and period of occurrence in two separate, parallel sets of mixed effects models. Syllabification principles for penultimate weight coding differentiate the two sets of modelling. Findings converge on segment count as a predictor of the comparative form, while the role of morphological complexity remains less clear, emerging significantly from one set of modelling but not the other. A rethinking of adjectival length based on segment count is advanced for our understanding of *y*-adjective comparatives. Discussed also are downstream implications of variant syllabification theories on accounts of *y*-adjective comparatives, together with insights shed on morpho-phonological intersections and the potential place of English *y*-adjective comparatives within the ambit of English alternations.

KEYWORDS

English comparatives, Phonemic segment, Syllable weight, Word length, Diachrony

1 INTRODUCTION

When work began on this article, the goal was to investigate whether two novel factors would account for the English COMPARATIVE FORMS (*more*, *-er*) of a group of adjectives. Ending in an orthographic <y>, pronounced /i/, these are called *Y*-ADJECTIVES. Examples include *silly*, *worthy*, *lazy* and *friendly*. Factors, otherwise known as predictors, of interest are the number of PHONEMIC SEGMENTS a *y*-adjective comprises (SEGMENT COUNT) and the WEIGHT (light, heavy) of its PENULTIMATE SYLLABLE (PENULITMATE WEIGHT). At the heart of several arguments in this article is a conception of adjectival length based on phonemic segmentation. Hence, to foreground this, I have opted to use the term, phonemic segments, rather than the term, PHONEMES, on its own. Segment count and penultimate weight have not been studied in past accounts of comparative alternation. As this article will show, their study highlights the value word length defined by segment count has for an understanding of *y*-adjective comparatives and furthers discussion related to the SYLLABLE UNIT for this understanding. Moreover, where the syllable units that derive penultimate weight assessments of *y*-adjectives are, themselves, subject to different theories of syllabification, the capacity of accounts of *y*-adjective comparatives to foreground the non-congruence between these theories shows up in the present work. Called into play correspondingly is the subtlety of morphological complexity in advancing our understanding of *y*-adjective comparatives. This subtlety, previously attributed to frequencies of comparative constructions in user cognition (Chua 2016: 177–8, 2019: 397), may now be purported to arise also from subscribed syllabification principles. This renders the morphological factor as possibly subject to downstream implications of non-congruent theories of syllabification. It is in the ways above and more that the present work bites off a larger chunk of the scholarship than intended from the outset. Next, I will explain how formerly studied factors set the scene for foregrounding segment count and penultimate weight as potential predictors of *y*-adjective comparatives.

2 Linguistic unit volume in English comparative accounts

The potential value of segment count and penultimate weight for an understanding of *y*-adjective comparatives is traceable to an intrinsic interest in the volume of linguistic units in several accounts of English comparatives. Often evoked to explain the tilting of adjectives in favour of one comparative alternative or another, these units may include a word's SYLLABLE(S) (Jespersen 1949: 347; Schibsbye 1965: 134; Zandvoort & Van Ek 1977: 188; Quirk *et al*. 1985: 461–462; Palmer *et al*. 2002: 1583–1584; Carter & McCarthy 2006: 439; Hilpert 2008: 407), its STRESS DISTRIBUTION (Kruisinga 1932: 62; Curme 1947: 220; Jespersen 1949: 350; Zandvoort & Van Ek 1977: 189; Quirk *et al*. 1985: 462; Palmer *et al*. 2002: 1583; Mondorf 2003: 278), its MORPHOLOGICAL CONSTITUENT(S) (Leech & Culpeper 1997: 355; Mondorf 2003: 283, 2009: 141; Hilpert 2008: 407; Chua 2016: 71, 2018: 480), and its distributions of PREMODIFICATION (Leech & Culpeper 1997: 367; Lindquist 1998: 127, 2000:

132) and COMPLEMENTATION (Mondorf 2003: 262; Hilpert 2008: 407). There are, of course, several other factors evoked in the literature to explain English comparative alternation. However, the ones above are noted because, together, they demonstrate that it is not uncommon – across phonological, morphological and syntactic considerations alike – for accounts of English comparative forms to be underpinned by a reference to the relative volume of linguistic units.

At the phonological level, by sheer count of syllables, we may say that a word with two or more syllables denote more volume than a monosyllabic word. When we speak of volume moreover, a correspondent notion is bulk; just as more syllables denote more volume, it may denote more bulk. The association of insufficient 'bulk' with phonologically unstressed units (Haspelmath 2008: 18) makes it reasonable therefore to include phonological stress as an indicator of linguistic unit volume, i.e. relatively more stress indicates relatively more volume, all other things being equal. Relatively more MORPHEMES in a word may also denote relatively more volume. In saying this, however, important to note is that a word's morpheme count is not always formally transparent and ascertainable by the amount of surface form material alone. Let us take, for example, the *y*-adjectives *silly* /sɪl.i/ and *lucky* /lʌk.i/. *Silly* comprises one morpheme, i.e. *silly* cannot be further broken down into meaningful parts. *Lucky*, on the other hand, comprises two morphemes, i.e. the meaning of *luck* and the attribute of experiencing this *luck*. Although this is so, the two words have the same amount of orthographic and phonetic surface form material – four phonemes and five letters each. The point here is that while with syllables as a linguistic unit, more volume necessarily denotes more surface forms, with morphemes, more volume, denoted by more morphemes, need not necessarily turn up more surface forms. It turns up, instead, in morphologically complex rather than simple forms. Beyond a word's span, phrasal-syntactic structures with infinitival or prepositional complements, or premodification, visibly denote more volume than parallel structures without. The linguistic units mentioned – syllables, stress distribution, morphological constitution, complementation and premodification – have all been proposed as potential predictors of English comparatives (Hilpert 2008: 407; Mondorf 2009: 64–8, 72–5). Garnered from them then is a sense that differences in the volume of linguistic units impact the choice between *more* and *-er*. It becomes easy to see as such how my present interest in segment count and penultimate weight coheres with the way the scholarship has conventionally thought about English comparative alternation. That is, when I think of this alternation as potentially predictable by whether a *y*-adjective has more or fewer phonemic segments, or a heavier or lighter (Lass 1992: 68; Hyman 2003: 5) penultimate syllable, I am thinking of the alternation in terms of differences in the volume of linguistic units. The volume, in this case, is specified through segment count and penultimate weight rather than, say, though syllables or morphological constitution.

3 Why phonemic segment count?

Word length based on indicators other than syllable count are not novel in predicting between morphological and phrasal alternatives. Character count was a proxy, for example, for word length where more -'*s* and *of*-genitives were found, respectively, with longer possessums and longer possessors (Ehret *et al*. 2014: 276). The mention of the English genitive here may raise a question on whether English comparative and genitive alternations are comparable in terms of seeking alternative indicators of word length (beyond syllable count) to explain English comparatives. It is true that while they both alternate between morphological and phrasal forms, English comparatives and genitives are dissimilar in some ways. For example, while the way genitives -'*s* and *of* order the possessum and possessor feature in how the length of possessums and possessors impact the choice of genitive form (Ehret *et al*. 2014: 276), an ordering constraint is not likewise prominent in the way adjective length has been studied to impact the choice of comparative form. It might come across somewhat in arguments for end-weight effects (Mondorf 2009: 100), where *more*, "by creating a heavier constituent" than *-er*, is hypothesised to be favoured "in end position". Even then, relevant findings have shown that the condition of positioning/ordering on the comparative is "[q]uite tellingly…weakest for disyllabics ending in <y>". These disyllabics comprise, in part, the adjective group of interest in the present work, though my categorisation of *y*-adjectives might be noted to also include those that comprise more than two syllables.

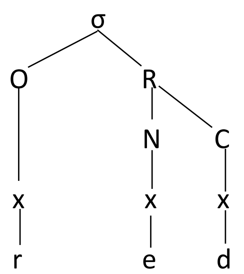
The point here is that it seems more fair to do otherwise than to engage in an argument on whether sufficient similarity exists between English comparative and genitive operations to justify a thinking of the former as potentially predictable by word length indicators alternative to syllable count. No other work at present has placed comparative alternation alongside genitive alternation in reference to word length determinants or otherwise, and hence, there are no grounds a priori to jump the gun and say just because there is an ordering constraint in genitive alternation that intersects with the word length predictor, we have to find a comparable constraint in comparative alternation, even if by coercion through end-weight, before one can propose as follows: word length indicators alternative to syllable count have before explained the choice between English morphological–phrasal alternatives, i.e. in genitive alternation, so there is a possibility these indicators might explain the choice between these alternatives in English comparative alternation. In other words, the case with English genitives ought not to obstruct this proposal on grounds of limited comparability between English comparative and genitive alternations because we simply have no evidence that these grounds matter. On the contrary, if it turns out that a redefinition of word length from syllable count does matter for the choice between comparative *more* and -*er*, knowledge is advanced of the contribution of redefinitions of this type in understanding both English comparatives and genitives. The uptake of some form of word length redefinition for its potential in predicting between comparative *more* and -*er* is, thus, reasonable.

The more important question is whether, as an alternative to syllable count, adjectival length ought, for the purposes of this study, to be served by character or segment count. Character count seems to present as a good candidate in the first instance, since it has a precedence in affecting the morphological–phrasal alternation in the English genitive (Ehret *et al*. 2014: 276). High correlations between word character and syllable counts have, moreover, been reported (Ehret *et al*. 2014: 276, citing Wolk *et al*. 2013: 395), pointing to the former as a close approximate of the latter. However, grounds exist to support the use of segment count as an indicator of adjectival length in the present work. First, the use of segment count to predict comparative forms follows intuitively from the phonemic/phonological character of many previous comparative form predictors (Hilpert 2008: 407). It follows also from the nature of the data used in this article, namely, seven corpora of British stage comedies spanning between the 17th and 20th centuries (more on this later). Stage comedies are written to be spoken more than they are to be read, so segment count (as a derivative of phonemic segments), more so than character count (as a derivative of orthographic characters), would reflect more authentically the way the comedies would have been received by the populace during the periods in which they were written.

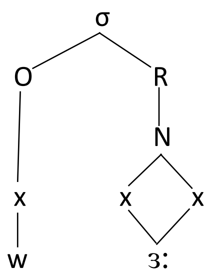
It is useful to point out that in segmenting phonemes to derive segment count in this article, DIPHTHONGS and LONG VOWELS will be taken to comprise two segments each, following Carr's (1999: 70–1) assignment of these constituents into separate SKELETAL TIERS (or TIMING SLOTS) within the

NUCLEUS of a syllable, as opposed to a single skeletal tier occupied by a short vowel. Examples (1), (2) and (3), respectively, of the syllables /red/ (from /red.i/ *ready*), /wɜː/ (from /wɜː. ði/ *worthy*) and /weɪ/ (from /weɪ.ti/ *weighty*) illustrate this. In the examples, σ represents the syllable, O, the onset of the syllable, R, the rime of the syllable, which comprises the nucleus N and the coda C, and x, the skeletal tier/timing slot to which Carr refers. As seen in examples (2) and (3), respectively comprising the long vowel /ɜː/ and the diphthong /eɪ/ as their syllable nucleus (N), /ɜː/ and /eɪ/ each branches – whether in or out – from two skeletal tiers x, whereas in example (1), the single short vowel /e/ branches from only one skeletal tier. Graphically, this shows, following Carr (1999), of long vowels and diphthongs comprising two phonemic segments each, as against the composition of one phonemic segment in short vowels.

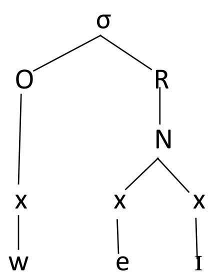
(1)



(2)

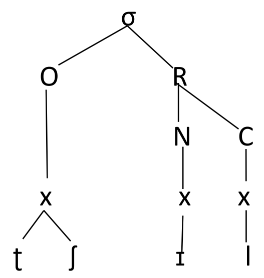


(3)



AFFRICATES, though, will be taken as a single phonemic segment, since they have been documented to ‘behave like single segments’, ‘occupy[ing] a single unit of timing’ (Carr 1999: 71). As example (4) illustrates with the syllable /ʧɪl/ (from / ʧɪl.i/ *chilly*), the affricate /ʧ/ branches only from one single skeletal tier x, unlike the long vowel /ɜː/ and the diphthong /eɪ/, respectively, in examples (2) and (3).

(4)



4 WHY PENULTIMATE WEIGHT?

While word length redefinitions underscore segment count as a potential predictor of *y*-adjective comparatives, a recognition of *y*-adjective syllable units, where segment count does not, renders penultimate weight a potential predictor. The fact, for example, that *witty* /wɪti/ has four segments does not also signal that *witty* has two syllables; there are other adjectives, such as *cross* /krɒs/, which have four segments but are monosyllabic. It helps little as well to turn to syllable count for a recognition of syllable units, as, mostly disyllabic, *y*-adjectives vary minimally in this count. More helpful would be an alternative way of encoding syllable unit variation, and penultimate weight works because like syllable count, it rests on a delimitation of SYLLABLE BOUNDARIES.

To exemplify this, let us refer to *y*-adjectives *goody*, *healthy* and *speedy*, and their phonemic transcriptions from two dictionaries (see Table 1).

Table 1. *Examples of penultimate weight variation in y-adjectives.*

Y-adjectives	(A)		(B)	
	Phonemic transcription	Rime structure^	Phonemic transcription	Rime structure^
goody	/gʊd.i/	VC	/gʊd.i/	VC
healthy	/hel.θi/	VC	/helθ.i/	VCC
speedy	/spi:.di/	VV	/spi:d.i/	VVC

Notes:
(A) Phonemic transcriptions based on *Cambridge Dictionary* online (Cambridge University Press 2020); syllable boundaries are marked by periods (.).
(B) Phonemic transcriptions based on *Longman Pronunciation Dictionary* (Wells 2000); syllable boundaries are marked by periods (.).
^Refers to the rime structure of only the penultimate syllable of a phonemic transcription; V stands for single vowel; C stands for a single consonant; VV stands for a long vowel, following the classification of long vowels (and diphthongs) to comprise two phonemic segments (see Section 3 above).

The adjectives exemplified in Table 1 all comprise two syllables. They also do not differ in having a final OPEN SYLLABLE comprising the same nucleus, /i/. Of interest is the variation in their penultimate syllable, specifically the weight of this syllable (SYLLABLE WEIGHT). Syllable weight here and elsewhere are taken to exclude the syllable ONSET (Hayes 1981; Hammond 1997; Giegerich 2009), and to depend 'solely on the properties of [the] RIME [my emphasis]' (Hyman 2003: 6). While there are syllable weight accounts that incorporate the onset, namely, MORA-derived weight accounts (Gordon 2002b: 5), which have some morae constituted, in part, by onsets (Hyman 2003: 16; Bauer 2012: 117–18), an onset-exclusive weight remains justified by the ONSET CREATION RULE (OCR) (Hyman 2003: 15–16). This rule specifies that the WEIGHT UNIT (WU) of a [+consonant] segment associates with its right [-consonant] segment, consequently deleting onset weight. Given this, and the association of syllable weight with phonemic segments (Gordon 2002b: 2), Table 1 shows *goody* to have a vowel-consonant (VC) rime in its penultimate syllable, and, depending on the dictionary source of the phonemic transcription, *healthy* is shown to have a vowel-consonant (VC) or vowel-consonant-consonant (VCC) rime, and *speedy*, to have a vowel-vowel (VV) or vowel-vowel-consonant (VVC) rime. If we accept (see the relevant argument later in Section 5.2) that a heavy syllable, compared to a light one, is derived from a complex nucleus or CODA – where complexity is defined by a VV and/or a CC rime (Lass 1992: 68), *speedy* differs from *goody* in having a heavy, rather than light, penultimate weight. *Healthy* may be the same or different in penultimate weight to either, depending on whether we take its rime to be VCC – where its penultimate weight is heavy, or VC – where its penultimate weight is light. Differences as above justify an encoding of syllable unit variation through penultimate weight.

As illustrated with *healthy* /helθ.i/ (or /hel.θi/), a *y*-adjective's penultimate weight may be heavy or light depending on where the boundary lies between the penultimate and final syllable. Found in variant dictionary transcriptions, for example, /hel.θi/ in *Cambridge Dictionary* online (Cambridge University Press 2020) versus /helθ.i/ in *Longman Pronunciation Dictionary* (Wells 2000), these different syllable boundaries stem from different syllabification principles. In the case of *healthy*, the MAXIMAL ONSET PRINCIPLE (MOP; Carr 1999: 74; Schlüter 2009: 169), anticipating coda spillovers to following onsets, have /θ/ syllabified as the onset of the final syllable in /hel.θi/. On the other hand, the syllabification of /θ/ in /helθ.i/ is justified by conditions in Wells (2002) for the retention as codas of syllable-final consonants that might have an alternative conception as onsets of the following syllable. In *healthy*, the penultimate syllable is relatively more stressed, which attracts /θ/ if 'consonants are syllabified with the more strongly stressed of two flanking syllables' (Wells 2002). /θ/ is part of the morpheme, *health*, meaning that if 'consonants belong to the syllable appropriate to the morpheme of which they form a part' (Wells 2002), it should be retained with the penultimate syllable in *healthy*. In addition to *healthy*, other *y*-adjectives exist that have penultimate syllable coda consonants that may be alternatively conceived as final syllable onsets because of the differential syllabification principles between the MOP and Wells (2002). My way around this, barring a removal of syllable unit considerations, is to see whether findings converge from dual sets of analyses – one with penultimate weight data informed by the MOP (dataset-MOP), and another with this data informed by Wells's (2002) syllabification conditions (dataset-Wells). The datasets used are uploaded to: https://osf.io/9eqrg/.

5 Segment count and penultimate weight in comparative alternation

5.1 *Data description*

To determine whether segment count and penultimate weight account for comparative forms, 253 tokens (54 types) of comparative *more* and *-er* *y*-adjectives were examined. These *y*-adjectives were obtained from seven diachronic corpora compiled by the author from a selection of British English stage comedy excerpts published between the 17th and 20th centuries. As several excerpts were obtained via institutional access in Victoria University of Wellington from *Literature Online* (Proquest 1996–2013), with terms governing them solely for personal or internal use, the corpora cannot be made publicly available. Nevertheless, the diachronic slant of this data lends the advantage that any corresponding conclusions drawn about English *y*-adjective comparative formation would have a built-in consideration of the passage of time.

Each of the seven corpora represented a time span of 50 years and comprised comedy excerpts published within those 50 years. Time periods correspondent to the seven corpora are: 1601–1650 (Period 1); 1651–1700 (Period 2); 1701–1750 (Period 3); 1751–1800 (Period 4); 1801–1850 (Period 5); 1851–1900 (Period 6); and 1901–1950 (Period 7). Each corpus comprises approximately 288,000 words (Chua 2018: 471). The selection of comedies for the corpora was guided by the goals of achieving an approximate measure of consistency in the word counts for each corpus, in the number of playwrights whose plays were included in the corpus, and in the word counts tagged to each playwright – see Chua (2016: 77–79) for a documentation of the compilation process and Chua (2016: 197–203) for a list of the comedies (and their playwrights) included in the corpora. The *y*-adjectives extracted for examination in the present work, and in previous works based on the seven corpora (Chua 2016, 2018), were ones found with English comparative forms *more* or *-er* within a single 50-year period and/or across multiple 50-year periods of the data.

5.2 *Data coding*

Each *y*-adjective of the 253 studied comparative tokens was coded for its number of phonemic segments (segment count) and whether its penultimate syllable was heavy or light (penultimate weight). Since *y*-adjectives are non-variant in their word-final /i/, segment count excluded this /i/. As above, diphthongs and phonologically long vowels in *y*-adjectives were taken to comprise two segments each, and affricates to comprise one segment (Carr 1999: 70–1). Given diachronic data, transcriptions that derive segment count and penultimate weight considered period-relevant phonemic make-up – see Dobson (1968), Cruttenden (1994) and MacMahon (1998). For some *y*-adjectives, parts of their phonemic transcription required a period-referenced variation from an otherwise contemporary transcription. For example, *courtly* from a token of *more courtly* was transcribed /kɔə(oə)ɹtli/ rather than present-day UK English /kɔ:tli/ because it was found in a 17th century corpus. 18th century was when post-vocalic /r/ loss occurred (Cruttenden 1994: 75), and since 'a large number of RP [received pronunciation] /ɔ:/ result from the loss of post-vocalic /r/ in the eighteenth century…via such stages as [ɔə] or [oə]' (Cruttenden 1994: 111), 17th-century *courtly* is likely pronounced /kɔə(oə)ɹtli/. Where applicable, contemporary transcriptions draw on the *Cambridge Dictionary* online (Cambridge University Press 2020) and the *Longman Pronunciation Dictionary* (Wells 2000), depending on the syllabification principles that inform penultimate weight for a specific analysis. In view of linguistic economy (Whitney 1868: 28), /p/ optionality in the *Longman* transcription of *empty* /empti/ is transcribed without /p/ in the dataset that relies on Wells's (2002) syllabification principles for penultimate weight coding. Where this /p/ is not similarly optional in the *Cambridge* transcription, *empty* is one segment count fewer for dataset-Wells than for dataset-MOP.

For assessing penultimate weight, the literature presents at least two ways. As a structural property of a syllable's rime, this assessment may be guided by rime complexity. Here, a LIGHT SYLLABLE occurs where 'neither the nucleus nor the coda is complex' (V and VC rimes), and a HEAVY SYLLABLE occurs where 'either the nucleus or coda (or both) is complex' (VV, VVC and VCC rimes) (Lass 1992: 68) – the double-V represents either diphthongs or phonologically long vowels. Syllables where both nucleus and coda are complex (VVCC rime) are SUPERHEAVY. A second school of thought agrees that V rimes are light, and broadly-speaking, that VV, VVC, VCC and VVCC rimes are heavy, but VC rimes are light or heavy (Ryan 2016: 721), depending on a language's obstruent-to-sonorant and voiceless-to-voiced 'ratio[s] of coda consonants' (Gordon 2002a: 74). Syllable weight here is guided by a rime's energy more than its complexity, so that a language with fewer sonorant than obstruent, and voiced than voiceless, codas have VC rimes denote a light syllable. Conversely, a language with more sonorant than obstruent, and voiced than voiceless, codas have VC rimes denote a heavy syllable.

The fact that English is of interest is not particularly helpful for deciding whether *y*-adjective penultimate weight should be assessed based on rime complexity or rime energy. Even as the language's VC rimes are deemed to represent light syllables (Lass 1992: 68), 'some extent' (Hyman 2003: 5) of their heaviness is claimed (Ryan 2016: 721). The grey area though in taking English VC rimes as a heavy syllable based on rime energy alone carves out space for supporting these rimes as signals of a light syllable. For this, the English coda consonant inventory in Table 2 helps. Here, we see fewer possible sonorant than obstruent codas in the language (6 sonorants versus 15 obstruents), and fewer possible voiced than voiceless codas (13 voiced versus 8 voiceless).

Table 2. *English coda consonant inventory*.

Consonant feature	Consonants	Consonant count
sonorant	ŋ, n, r, m, l, w	6
obstruent	p, d, z, t, v, k, b, g, ʔ, f, ɵ, ð, s, ʃ, ʒ	15
voiced	ŋ, n, d, r, z, v, m, l, b, g, ð, ʒ, w	13
voiceless	p, t, k, ʔ, f, ɵ, s, ʃ	8

Note: With the exception of /h/, which is not found as a coda consonant in English, entries follow from the view that 'English permits stops, fricatives, nasals and liquids as coda consonants' (Yavas & Core 2001: 37).

While a light syllable in English VC rimes is justified by the relatively higher count of obstruent codas in Table 2, a heavy syllable in them is justified by the relatively higher count of voiced codas. This ambiguity is common across languages (Ryan 2016: 728), but a way around it for English is to observe that its obstruent coda count is 2.5 times higher than its sonorant coda count, while its voiced coda count is only 1.6 times higher than its voiceless coda count. Proportion-wise, the case obstruent codas present for having English VC rimes signal light syllables is stronger than the case voiced codas present for having these rimes signal heavy syllables. This coheres with a rime complexity view of syllable weight (Lass 1992: 68), where a VC rime, given its simpler structure than a VV, VVC, VCC or VVCC rime, represents a light syllable. In penultimate weight coding therefore, *y*-adjectives with a V or VC rime in their penultimate syllable were deemed light. *Y*-adjectives with a VV, VVC, VCC or VVCC rime in their penultimate syllable were deemed heavy. Although in some places, a VVCC rime deems a syllable superheavy (Lass 1992: 68), *y*-adjective tokens with this rime structure in their penultimate syllable are far and few in between in my datasets – one such token for dataset-MOP, and four such tokens for dataset-Wells. To avert a drastic imbalance in token count for the possible range of penultimate weights, which will pose data convergence issues for statistical analyses, tokens with a superheavy penultimate syllable were collapsed with those of a heavy penultimate syllable.

5.3 *Findings*

Segment count and penultimate weight were analysed with variables formerly found to predict *y*-adjective comparatives (Chua 2016: 113–15, 2018: 480), namely, the corpus period in which *y*-adjective comparative tokens are found (period) and the morphological complexity of the relevant *y*-adjectives (morphology). Period (1–7) and segment count (2–7) are continuous variables, and morphology (complex, simple) and penultimate weight (light, heavy) are binary variables. Form (more, -er) was included to identify the comparative forms of *y*-adjectives. Item, which differentiates between *y*-adjective lexemes, was included to allow ‘any fluctuation between *more* and *-er*’ (Chua 2018: 477) to be lexically specified (Palmer *et al*. 2002: 1583). With period, segment count, morphology and penultimate weight as independent variables (IVs), form as the dependent variable (DV), and item as a random effect, a series of MIXED EFFECTS MODELS (MEMs) were fitted using the glmer function from the lme4 package (version 1.1-9) (Bates *et al*. 2018) in R (version 3.5.2) (R Core Team 2018). Model comparisons were perfomed using the anova function from the lme4 package. To permit period and segment count effects to differ for *y*-adjectives, random by-item slopes for these were included where no non-convergent models resulted.

As noted above (Section 4), dataset-MOP and dataset-Wells were separately analysed. Initial mixed effects modelling of both datasets included interactions between:

- period and morphology;
- period and penultimate weight;
- morphology and penultimate weight;
- morphology and segment count; and
- penultimate weight and segment count.

A shift towards comparative *-er* for *y*-adjectives with time (Kytö & Romaine, 1997: 344), and a purported *-er* bias with morphological simplicity (Leech & Culpeper 1997: 355; Mondorf 2003: 283; Hilpert 2008: 407; Chua 2018: 466), justifies an interaction between period and morphology in the first instance, to detect whether any -*e*r bias towards later time periods is primarily found with simple *y*-adjectives (Chua 2018: 479). Period and penultimate weight interaction considers the possibility of differential weight effects over time associated with potential diachronic subjectivity in the phonemic make-up of *y*-adjectives. Interactions between morphology and penultimate weight, between morphology and segment count, and between penultimate weight and segment count consider the potential of overlaps between morphological and weight, between morphological and segment count, and between segment count and weight effects. These overlaps (if any) are worthwhile to note because they would show whether, and if so, to what extent morphological complexity, penultimate weight and segment count, as indicators of the volume of linguistic units, are reinforcing (or otherwise) in predicting *y*-adjective comparatives. It may be the case, for example, that any bias towards *more* or *-er*, with say, a light penultimate weight, is found only in *y*-adjectives with a relatively lower segment count, given that light weight tokens concentrate in *y*-adjectives of 4 segments or less for both datasets (see Table 3).

Table 3. *Cross-tabulation of token counts of y-adjectives between penultimate weight and segment count.*

Segment count	Dataset-MOP		Dataset-Wells	
	Penultimate weight		Penultimate weight	
	Heavy	Light	Heavy	Light
3	48	116	50	115
4	20	29	29	19
5	31	3	33	1
6	4	0	4	0
7	2	0	2	0

Note: Segment count excludes the non-variant word-final /i/ of *y*-adjectives.

In neither dataset is inclusion, in initial modelling, of one or more of the interactions above pre-empted by any high correlation of the centered IVs in interaction – see Table 4 with no value >0.7 (Clark & Randal 2011: 60).

Table 4. *Correlation matrix of all IVs (centered) proposed for inclusion in MEMs of dataset-MOP*.

Dataset-MOP	morphology	segment count	penultimate weight
segment count	-0.44182486		
penultimate weight	0.53231163	-0.43074211	
Period	0.01930207	-0.04307304	-0.13219278
Dataset-Wells	morphology	segment count	penultimate weight
segment count	-0.44717369		
penultimate weight	0.58261071	-0.48589843	
Period	0.01930207	-0.04257026	-0.0972182

Note: Centred IVs are used in this correlation matrix because centering changes the binary IVs into numeric predictors, which a correlation matrix requires.

Table 5 presents results from the step-wise modelling of dataset-MOP.

Table 5. *Effects considered for the comparative forms of y-adjectives indicated from seven mixed effects models; Model 4-MOP accepted as most explanatory, dataset-MOP.*

	Estimate	Standard Error (SE)	z-value	p-value
Model 1-MOP				
period	0.10630	0.12877	0.826	0.409080
penultimate weight	2.28046	2.55375	0.893	0.371864
morphology	-5.56004	3.25402	-1.709	0.087513
segment count	-1.25706	0.27396	-4.589	4.46e-06***
period:morphology	0.07095	0.20317	0.349	0.726929
morphology:penultimate weight	0.25652	1.09444	0.234	0.814689
penultimate weight:segment count	-0.82986	0.65517	-1.267	0.205282
morphology:segment count	1.72121	0.87560	1.966	0.049327*
period:penultimate weight	0.17483	0.19439	0.899	0.368444
Model 2-MOP				
period	0.1790	0.1008	1.777	0.075602
penultimate weight	2.9704	2.4104	1.232	0.217837
morphology	-6.4148	3.1187	-2.057	0.039701*
segment count	-1.2454	0.2730	-4.562	5.07e-06***
period:morphology	0.1587	0.1757	0.903	0.366304
morphology:penultimate weight	0.4954	1.0454	0.474	0.635567
penultimate weight:segment count	-0.8313	0.6437	-1.291	0.196556
morphology:segment count	1.8170	0.8512	2.135	0.032793*
Model 3-MOP 3				
period	0.23399	0.08192	2.856	0.004285**

penultimate weight	2.91987	2.42172	1.206	0.227934
morphology	-5.53165	2.92550	-1.891	0.058646
segment count	-1.25782	0.27523	-4.570	4.87e-06***
penultimate weight:morphology	0.24092	0.99654	0.242	0.808971
penultimate weight:segment count	-0.80820	0.64685	-1.249	0.211500
morphology:segment count	1.78626	0.85324	2.094	0.036303*
Model 4-MOP				
period	0.23221	0.08142	2.852	0.00434**
penultimate weight	3.13137	2.28624	1.370	0.17079
morphology	-5.36248	2.88643	-1.858	0.06319
segment count	-1.24394	0.26817	-4.639	3.51e-06***
penultimate weight:segment count	-0.85394	0.62561	-1.365	0.17226
segment count:morphology	1.78883	0.86713	2.063	0.03912*
Model 5-MOP				
Period	0.22479	0.08096	2.776	0.0055**
penultimate weight	-0.05277	0.42827	-0.123	0.9019
Morphology	-4.40224	2.82529	-1.558	0.1192
segment count	-1.40259	0.26213	-5.351	8.76e-08***
penultimate weight:morphology	0.58872	0.96495	0.610	0.5418
morphology:segment count	1.41460	0.78342	1.806	0.0710
Model 6-MOP				
Period	0.22009	0.08017	2.745	0.00605**
penultimate weight	0.04759	0.39451	0.121	0.90398
Morphology	-3.56776	2.46401	-1.448	0.14763

segment count	-1.38451	0.25924	-5.341	9.26e-08***
morphology:segment count	1.30314	0.75832	1.718	0.08571
Model 7-MOP				
Period	0.21840	0.07889	2.768	0.00564**
Morphology	-3.54497	2.45614	-1.443	0.14893
segment count	-1.39176	0.25229	-5.517	3.46e-08***
morphology:segment count	1.30175	0.75807	1.717	0.08594

*p<.05, **p<.01, ***p<.001.

Notes: Factors separated by a colon indicate two-way interactions. For example, period:morphology indicate a two-way interaction between period and morphology. Interactions included generate also simple effects within those interactions. Lower-order significant effects within higher-order ones are not further analysed (De Rosario-Martinez 2015: 6)

From Table 5, the initial model, Model 1-MOP, finds a significant interaction between morphology and segment count (estimate=1.721, SE=0.876, z=1.966, p<.05), and so does Model 2-MOP (estimate=1.817, SE=0.851, z=2.135, p<.05). Though the interaction between period and penultimate weight was dropped in Model 2-MOP to see whether significance might obtain for period and/or penultimate weight as simple/independent effects, none was found. Models 1-MOP and 2-MOP do not significantly differ (chi-square chi=0.8206, df=1, p>.05), so Model 2-MOP, with fewer parameters/factors and hence, relative simplicity, is accepted over Model 1-MOP. Period, previously found to be significant from the same data (Chua 2016: 115, 2018: 482), was included only as a simple effect in Model 3-MOP and is again found significant (estimate=0.234, SE=0.082, z=2.856, p<.01), alongside a significant interaction between morphology and segment count (estimate=1.786, SE=0.853, z=2.094, p<.05). Model 3-MOP does not differ significantly from Model 2-MOP (chi-square chi=0.8398, df=1, p>.05) and is accepted over the latter. To see whether significance might obtain for penultimate weight, interactions between penultimate weight and each of morphology and segment count were dropped in turn, respectively, in Models 4-MOP and 5-MOP. Neither Model 4-MOP (chi-square chi=0.0573, df=1, p>.05) nor Model 5-MOP (chi-square chi=1.6694, df=1, p>.05) differs significantly from Model 3-MOP, accepting the former two over Model 3-MOP for their relative simplicity. It is Model 4-MOP, though, that finds the interaction between segment count and morphology significant (estimate=1.789, SE=0.867, z=2.063, p<.05). As this significance resonates with the morphological factor in comparatives noted in other works (Leech & Culpeper 1997: 355; Mondorf 2003: 283, 2009: 141; Hilpert 2008: 407; Chua 2016: 115. 2018: 482), Model 4-MOP is accepted over Model 5-MOP on theoretical grounds. The fair number of works just cited is deemed sufficient to constitute these grounds, even if not all of them focus on *y*-adjectives, and even if some, e.g. Leech and Culpeper (1997), refer to the morphological factor indirectly thorough the concept of suffixation. It is precisely because *y*-adjective comparatives have not been extensively studied that this work and Chua (2016, 2018) came about, so it is not reasonable to expect all prior scholarship on English comparative alternation to have dealt with *y*-adjectives exclusively before we may appeal to these for the retention of morphological considerations in data analyses. To see whether further model simplification would better explain the data, the interaction between penultimate weight and segment count was dropped from Model 4-MOP. Resultant alternatives modelled with and without random by-item slopes do not, between them, differ significantly (chi-square chi=3.8817, df=5, p>.05), accepting the simpler alternative without random by-item slopes, Model 6-MOP. Although Models 4-MOP and 6-MOP do not significantly differ (chi-square chi=1.9624, df=1, p>.05), technically justifying an acceptance of Model 6-MOP, Model 6-MOP shows no significance of morphology while Model 4-MOP does (in interaction with segment count). Like with the case between Models 4-MOP and 5-MOP therefore, Model 4-MOP is accepted over Model 6-MOP on theoretical grounds, i.e. it resonates with previous works that note the morphological factor in comparative alternation (in *y*-adjectives) (Leech & Culpeper 1997: 355; Mondorf 2003: 283, 2009: 141; Hilpert 2008: 407; Chua 2016: 115. 2018: 482). Where a further simplification of Model 6-MOP to Model 7-MOP by dropping penultimate weight does not yield any significant morphological effect as well, Model 4-MOP is decidedly accepted as most explanatory of dataset-MOP.

Table 6 presents results from the step-wise modelling of dataset-Wells.

Table 6. *Effects considered for the comparative forms of y-adjectives indicated from six mixed effects models; Model 7-Wells accepted as most explanatory, dataset-Wells.*

	Estimate	Standard Error (SE)	z-value	p-value
Model 1-Wells				
period	0.13798	0.11304	1.221	0.222208
penultimate weight	2.78502	2.65184	1.050	0.293617
morphology	-4.94171	3.22243	-1.534	0.125144
segment count	-1.27327	0.27427	-4.642	3.44e-06 ***
period:morphology	0.02421	0.21159	0.114	0.908905
morphology:penultimate weight	-0.04292	1.05988	-0.040	0.967695
penultimate weight:segment count	-0.99358	0.73602	-1.350	0.177034
morphology:segment count	1.67354	0.91170	1.836	0.066413
period:penultimate weight	0.16743	0.20156	0.831	0.406161
Model 2-Wells				
period	0.1822	0.1002	1.819	0.068849
penultimate weight	3.2338	2.5736	1.257	0.208916
morphology	-5.4557	3.1449	-1.735	0.082779
segment count	-1.2673	0.2741	-4.622	3.79e-06 ***
period:morphology	0.1235	0.1706	0.724	0.468951
morphology:penultimate weight	0.1153	1.0239	0.113	0.910325
penultimate weight:segment count	-0.9431	0.7192	-1.311	0.189774
morphology:segment count	1.6786	0.8925	1.881	0.060011
Model 3-Wells				
period	0.18302	0.09996	1.831	0.067103

penultimate weight	3.32475	2.45615	1.354	0.175850
morphology	-5.34869	3.02618	-1.767	0.077149
segment count	-1.26082	0.26766	-4.710	2.47e-06 ***
period:morphology	0.11976	0.16711	0.717	0.473562
penultimate weight:segment count	-0.96165	0.70304	-1.368	0.171358
morphology:segment count	1.67530	0.90201	1.857	0.063269
Model 4-Wells				
period	0.22656	0.08056	2.812	0.004921 **
penultimate weight	3.25333	2.43355	1.337	0.181267
morphology	-4.87964	2.92172	-1.670	0.094894
segment count	-1.27435	0.26854	-4.745	2.08e-06 ***
penultimate weight:segment count	-0.94353	0.69864	-1.351	0.176845
segment count:morphology	1.66539	0.89223	1.867	0.061965
Model 5-Wells				
period	0.218138	0.079680	2.738	0.00619**
penultimate weight	-0.004845	0.416510	-0.012	0.99072
morphology	-3.292695	2.452791	-1.342	0.17946
segment count	-1.392635	0.264164	-5.272	1.35e-07***
morphology:segment count	1.221191	0.754574	1.618	0.10558
Model 6-Wells				
period	0.21827	0.07889	2.767	0.00566**
morphology	-3.29515	2.44348	-1.349	0.17748
segment count	-1.39173	0.25230	-5.516	3.46e-08***
morphology:segment count	1.22129	0.75443	1.619	0.10549

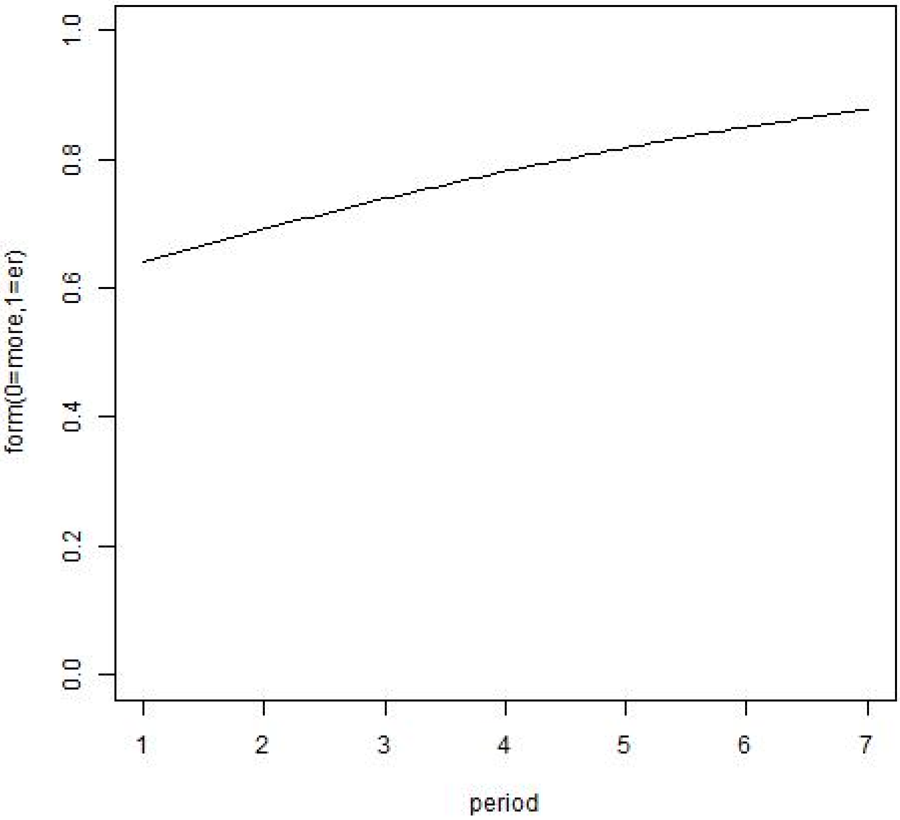
Model 7-Wells				
period	0.22323	0.08207	2.720	0.006527**
morphology	0.65618	0.49966	1.313	0.189097
segment count	-1.14714	0.27378	-4.190	2.79e-05***

*p<.05, **p<.01, ***p<.001.

Notes: Identical to notes below Table 5.

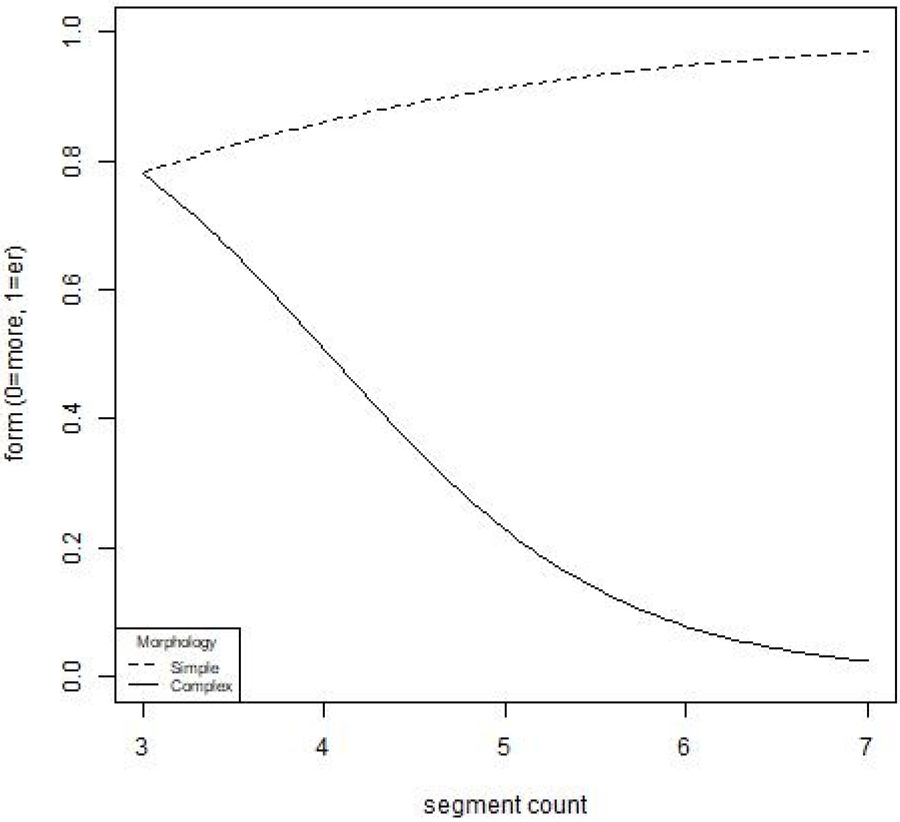
From Table 6, the initial model, Model 1-Wells, finds a significant simple effect of segment count (estimate=-1.273, SE=0.274, z=-4.642, p<.001), and so does Model 2-Wells (estimate=-1.267, SE=0.274, z=-4.622, p<.001), though latter excluded the interaction between period and penultimate weight. Since Models 1-Wells and 2-Wells do not significantly differ (chi-square chi=0.7086, df=1, p>.05), Model 2-Wells, with fewer parameters/factors and hence, relative simplicity, is accepted over Model 1-Wells. To see whether significance might obtain for morphology, the interaction between morphology and penultimate weight was dropped in Model 3-Wells; here morphology and period come close to, but do not reach significance, while segment count remains significant (estimate=-1.261, SE=0.268, z=-4.710, p<.001). Models 2-Wells and 3-Wells do not significantly differ (chi-square chi=0.0126, df=1, p>.05), accepting Model 3-Wells for its relative simplicity. Since effects each of period and morphology were found significant in a prior work (Chua 2016: 115–16; 2018: 482), the interaction between them was dropped in Model 4-Wells to see whether this would yield their significance again; here, period is found significant (estimate=0.227, SE=0.081, z=2.812, p<.01), alongside segment count (estimate=-1.274, SE=0.269, z=-4.745, p<.001). Model 4-Wells does not differ significantly from Model 3-Wells (chi-square chi=0.5219, df=1, p>.05), accepting Model 4-Wells for its relative simplicity. Since penultimate weight effects have not drawn close to significance in the first few modellings of dataset-Wells, its interaction with segment count and as a simple effect were, respectively, dropped in Models 5-Wells and 6-Wells. Both models retain as significant period (Model 5-Wells: estimate=0.218, SE=0.080, z=2.738, p<.01; Model 6-Wells: estimate=0.218, SE=0.079, z=2.767, p<.01) and segment count (Model 5-Wells: estimate=-1.393, SE=0.264, z=-5.272, p<.001; Model 6-Wells: estimate=-1.392, SE=0.252, z=-5.516, p<.001). Given non-significant differences between them, Model 5-Wells is accepted over Model 4-Wells (chi-square chi=1.9472, df=1, p>.05), and Model 6-Wells over Model 5-Wells (chi-square chi=1e-04, df=1, p>.05), with the simpler model accepted in each case. To see whether any significance of morphology as a simple effect might obtain, the interaction between morphology and segment count was dropped in Model 7-Wells, which has period (estimate=0.223, SE=0.082, z=2.720, p<.01) and segment count (estimate=-1.147, SE=0.274, z=-4.190, p<.001) as significant effects. Models 7-Wells and 6-Wells do not significantly differ (chi-square chi=2.6966, df=1, p>.05), accepting Model 7-Wells for its relative simplicity. Morphology was dropped from Model 7-Wells, to see if a better explanation of the data might obtain. The resultant model, however, is non-convergent, i.e. the model struggles to explain the data. This indicates that morphology, though non-significant, must be retained, accepting Model 7-Wells as most explanatory of dataset-Wells.

From the model accepted for dataset-MOP (Model 4-MOP), Figures 1 and 2, plot, respectively, the significant effects of period (estimate=0.232, SE=0.081, z=2.852, p<.01), and the interaction between segment count and morphology (estimate=1.789, SE=0.867, z=2.063, p<.05). From the model accepted for dataset-Wells (Model 7-Wells), Figures 3 and 4, plot, respectively, the significant effects each of period (estimate=0.223, SE=0.082, z=2.720, p<.01) and segment count (estimate=-1.147, SE=0.274, z=-4.190, p<.001).



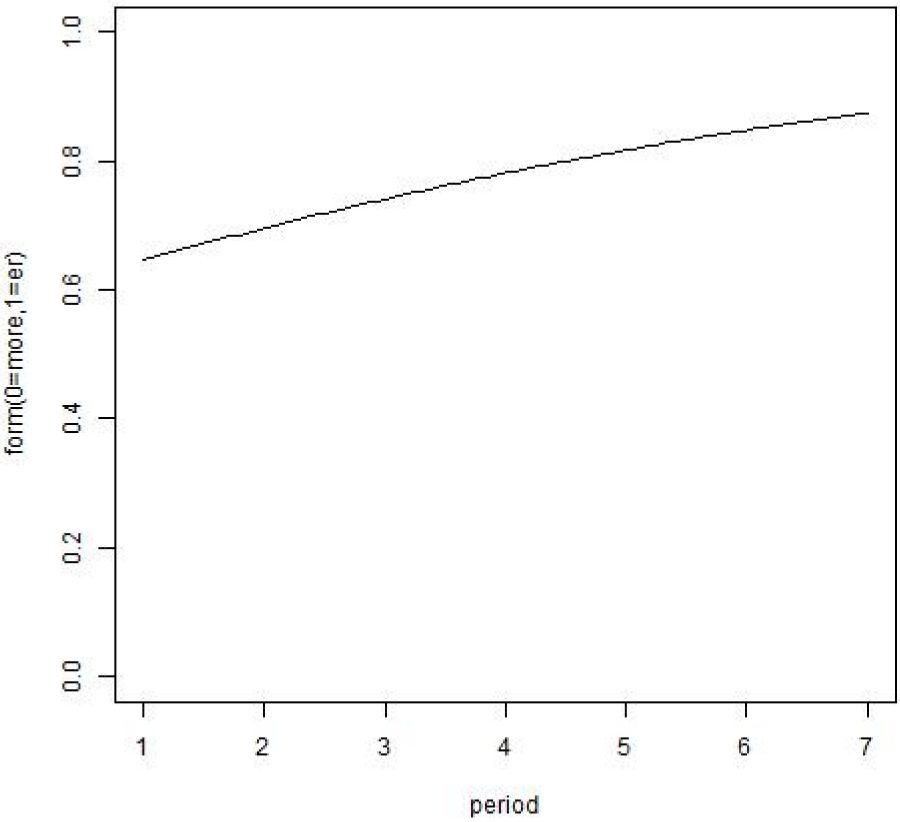
Notes: The span in years represented by each period along the x-axis are as follows – period 1 (1601–1650); period 2 (1651–1700); period 3 (1701–1750); period 4 (1751–1800); period 5 (1801–1850); period 6 (1851–1900); and period 7 (1901–1950). Values along the y-axis approaching 1.0 indicate a comparative *-er* tendency, and those approaching 0.0, a comparative *more* tendency.

Figure 1. Graph of the effect of period in Model 4-MOP, dataset-MOP.



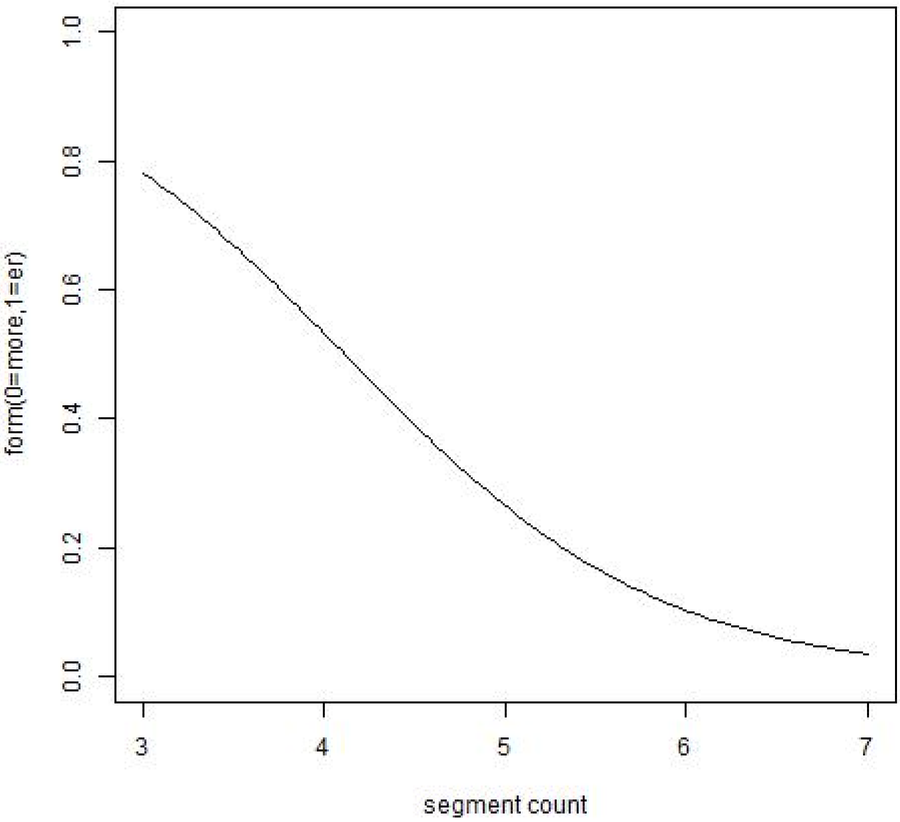
Notes: Segment counts exclude the final /i/ consistently found in *y*-adjectives. An example of a simple adjective is *silly*, and a complex one, *lucky*. Values along the y-axis approaching 1.0 indicate a comparative *-er* tendency, and those approaching 0.0, a comparative *more* tendency.

Figure 2. Graph of the interaction effect between segment count and morphology in Model 4-MOP, dataset-MOP.



Notes: The span in years represented by each period along the x-axis are as follows – period 1 (1601–1650); period 2 (1651–1700); period 3 (1701–1750); period 4 (1751–1800); period 5 (1801–1850); period 6 (1851–1900); and period 7 (1901–1950). Values along the y-axis approaching 1.0 indicate a comparative *-er* tendency, and those approaching 0.0, a comparative *more* tendency.

Figure 3. Graph of the effect of period in Model 7-Wells, dataset-Wells.



Notes: Segment counts exclude the final /i/ consistently found in *y*-adjectives. Values along the y-axis approaching 1.0 indicate a comparative *-er* tendency, and those approaching 0.0, a comparative *more* tendency.

Figure 4. Graph of the effect of segment count in Model 7-Wells, dataset-Wells.

Y-axis values approaching 1.0 in all figures indicate a comparative *-er* tendency, and those approaching 0.0, a comparative *more* tendency. Figures 1 and 3 therefore show a shift towards comparative *-er* from Periods 1 to 7 for both datasets-MOP and -Wells.

For dataset-MOP, Figure 2 shows that while for simple *y*-adjectives, an increase in segment count sees an increased comparative *-er* probability, for complex *y*-adjectives, an increase in segment count sees a decreased *-er* probability. Figure 2 also shows an increase in *-er* probability for simple *y*-adjectives as most apparent with segment count increases from 3 to 4, after which the trend begins to equilibrate with segment count increases from 4 to 7. With complex *y*-adjectives, a decrease in *-er* probability is sharpest with segment count increases from 3 to 5 and the decrease begins to equilibrate as segment count increases from 5 to 7. A gradient difference between the two lines in Figure 2 suggests that for dataset-MOP, with an increase in segment count, the shift away from *-er* (presumably, towards *more*) is more apparent for complex *y*-adjectives than the shift towards *-er* for simple *y*-adjectives. This may in part be because both simple and complex *y*-adjectives with around 3 phonemic segments have, as is, a relatively high probability of an *-er* pairing, close to 0.8 (the point the two lines meet in Figure 2). Any shift towards *-er* is therefore more constrained than a shift away from it.

For dataset-Wells, Figure 4 shows segment count to independently predict comparative form with no implication of morphology. Though it is the case for all *y*-adjectives alike in dataset-Wells, the trend in Figure 4 closely mirrors that of complex *y*-adjectives in dataset-MOP. That is, an increase in segment count sees a decreased comparative *-er* probability, where the decrease is sharpest with segment count increases from 3 to 5 and begins to equilibrate as segment count increases from 5 to 7.

6 DISCUSSION

Datasets-MOP and -Wells differ in having penultimate weight informed by the maximal onset principle (MOP; Carr 1999: 74; Schlüter 2009: 169) in the former, and by Wells's (2002) syllabification conditions in the latter. The analyses of both datasets point to segment count effects on *y*-adjective comparatives. This is regardless of whether the effects are found independently, or in interaction with morphological complexity. Independently, an *-er* bias diminishes with an increased segment count of *y*-adjectives. Where it implicates morphology, an increased segment count sees more of an *-er* bias for simple *y*-adjectives and more of a *more* bias for complex *y*-adjectives. Segment count is studied because it aligns with previously studied factors such as syllable and morpheme count in attending to the volume of linguistic units to explain English comparative form choice (Section 2). Hence, findings that segment count indeed explains this choice buttress the theory that comparative *more* and *-er* alternation rests, in part, on the volume of material constituting an adjective. In my case, this volume takes the form of the number of phonemic segments.

Where segment count consistently predicts *y*-adjective comparative forms across datasets-MOP and -Wells, empirically supported is a redefinition of base adjective length from syllable to segment count to advance an understanding of comparative alternation. Broadly speaking, this implies that whenever adjectival length proves insufficiently robust in predicting comparative forms, we may do well to examine whether our conception of this length has been exhausted before concluding that adjectival length per se is unhelpful for the predictions. The present article shows that adjectival length based on segment count is helpful in predicting *y*-adjective comparatives. On the other hand, the place of penultimate weight in accounts of *y*-adjective comparatives remains difficult to determine. Penultimate weight is not found to be a significant predictor in either the account accepted for dataset-MOP or dataset-Wells. However, to surface the significance of morphological effects, a consideration of penultimate weight had to be retained in the account accepted for dataset-MOP (Table 5); in the account accepted for dataset-Wells (Table 6), though, no consideration of penultimate weight had to be included. To recall, penultimate weight is analysed in the first instance because this weight is informed by, and thus, recognises the syllable unit where segment count does not (Section 4). Where the role of penultimate weight remains difficult to pin down, the question that is then begged is what the role of the syllable unit is in understanding the comparatives of *y*-adjectives. This, together with the consistent finding of segment count in explaining *y*-adjective comparatives, reinforces the important point that where the syllable unit informs a default conception of adjectival length, a rethinking of this length in terms of phonemic segments is crucial to explain *y*-adjective comparatives, if not also, subject to further study, the comparatives of other English adjectival groups.

It is notable that any issue on the syllable unit that refers back to penultimate weight derives primarily from different syllabification principles – the MOP (Carr 1999: 74; Schlüter 2009: 169) versus Wells (2002). This differentiation affects the way penultimate weight is scoped for coding and any consequent data analyses. Therefore, in the non-consistent demands between the accounts accepted for datasets-MOP and -Wells to keep in penultimate weight, what is apparent is the capacity of these accounts, taken together, to make manifest a non-congruence between different syllabification theories. Seeking a resolution to the non-congruence is not the goal here, though the downstream implications of this non-congruence are worthy of note. Datasets-MOP and -Wells differ only in the syllabification principles they draw upon to inform penultimate weight assessments. Hence, any discrepancy between them on the contribution of morphological complexity to explaining *y*-adjective comparatives may reasonably be linked to a differentiation in the syllabification principles applied to each dataset; specifically, the discrepant findings are in the significance of morphological effects (in interaction) in the model accepted for dataset-MOP, with no matching significance of these effects in the model accepted for dataset-Wells. Granted, the model accepted to explain dataset-MOP factors in previous support for morphological complexity in comparative form choice (Section 5.3), without which the model accepted for dataset-MOP might not be unlike the one accepted for dataset-Wells, i.e. with no significant morphological effect. However, we may note that while morphological effects are significant in four models from dataset-MOP (Table 5), in no model from dataset-Wells (Table 6) are they significant. Given this and the fact that morphological boundaries potentially overlap with the syllable boundaries that inform penultimate weight, especially in dataset-Wells, it is fair to propose that the models accepted for datasets-MOP and -Wells may, between them, turn out discrepant findings on morphological effects because of upstream differences in subscribed syllabification theories.

Where they do contribute to data explanation, however, in dataset-MOP, morphological effects found advance prior understandings; the interaction of these effects with segment count in dataset-MOP (Figure 2, Section 5.3) adds granularity to our understanding of how morphological complexity predicts comparative forms. In no other work has segment count been infused into this prediction, such that where discrepancies arose between corpus and experimental data of independent morphological effects on *y*-adjective comparatives (Chua 2016, 2019), a suppression theory was instead proposed. The idea then was that a sufficiently high frequency of pre-experimental, real world-derived, cognitive accumulation of *more* and *-er* patterns of comparative constructions could have suppressed any morphological effects that would otherwise surface (Chua 2016: 177–8, 2019: 397). However, the qualification of morphological effects by segment count in the model accepted for dataset-MOP in this work proposes that morphological effects on *y*-adjective comparatives may also be suppressible by base adjective segment count. Equilibria reached towards an *-er* bias for simple *y*-adjectives and away from this bias for complex *y*-adjectives, respectively, at 4 or more and 5 or more segments (Figure 2, Section 5.3), suggest that for *y*-adjectives, morphological effects on comparatives tend to be obvious only below certain segment count thresholds. Morphological effects are very subtle, in other words, when backgrounded against frequencies of patterned comparative constructions

(Chua 2016: 177–8, 2019: 397), base adjective segment count and subscribed syllabification principles, the last given discrepant findings on the morphological factor between datasets-MOP and -Wells (see previous paragraph). The implication here of the claimed subtlety of morphological effects is that their non-consistent emergence in accounts of *y*-adjective comparatives does not necessarily mean the effects' non-existence.

Indeed, previously clear morphological effects from the same corpus data here used (Chua 2016: 115–16, 2018: 482), when contrasted against their present elusiveness, may be evoked to support morphological subtlety in *y*-adjective comparative form predictions. It is possible that in previous relevant works (Chua 2016, 2018), other considered predictors exert relatively less influence on *y*-adjective comparatives, so that morphological effects, even if subtle, may independently surface. Presently, segment count could be influential in predicting comparative forms to the extent that morphological effects may no longer surface independently (in dataset-MOP) or significantly (in dataset-Wells). What is affirmed, though, is that where morphological effects are found, as in the model accepted for dataset-MOP, the direction of those effects cohere with previous accounts. A claimed bias for comparative -*er* in morphological simplexes (Leech & Culpeper 1997: 355; Mondorf 2003: 283; Hilpert 2008: 407; Chua 2018: 482) is reflected in Figure 2 (Section 5.3), where, for the most part, the dashed line (representing morphological simplexes) is visually closer to the value of 1.0 (representing an -*er* bias) than the solid line (representing morphological complexes).

Moving away from the morphological factor, what remains congruent across the models accepted for datasets-MOP and -Wells is a threshold beyond which segment count effects dilute. For complex *y*-adjectives in dataset-MOP (Figure 2) and for all *y*-adjectives in dataset-Wells (Figure 4), the shift away from a comparative *-er* bias is sharpest as segment count increases from 3 to 5, and this shift becomes more gradual with segment count increases from 5 to 7. Where segment count here is an indicator of adjectival length, indicated then is a quantifiable threshold beyond which adjectival length becomes less apparent in predicting *y*-adjective comparative forms. Lending support to this are reports of needed thresholds to surface the predictive effects of word length on linguistic outcomes (McGinnies *et al*. 1952: 69; New *et al*. 2006: 48). That said, implicit notions of these thresholds might already exist for English comparative alternation, where word length constraints in predicting comparative forms are recognised through a study of several factors (Hilpert 2008) alongside syllable count. What the present work importantly does is to secure a handle on these constraints by indicating, for *y*-adjectives, a quantifiable threshold beyond which word length effects on comparative forms begin to dissipate. The way in which this threshold manifests reinforces the value of phonemic segment count in advancing an understanding of comparative alternation. Especially for *y*-adjectives, where syllable count varies minimally and so, offers little in granting access to word length thresholds that inform comparative form choices, segment count becomes a useful means of granting this access. This usefulness simultaneously highlights an importance of phonetic factors relative to phonological ones in predicting *y*-adjective comparatives. Unlike segment count (predominantly phonetic), penultimate weight (predominantly phonological) remains undetermined in the present work in contributing to an understanding of *y*-adjective comparatives. Indeed, the fuzziness of phonological factors in accounting for comparative forms is previously found as well in such other considerations as stress positioning (Chua 2018: 461–2).

From the present work, the effect of period on *y*-adjective comparatives is also noteworthy. Found in both accepted models for datasets-MOP and -Wells is an independent biasing of *y*-adjectives towards *-er* with the passage of time, i.e. from the earliest Period 1 to the most recent Period 7 (Figures 1 and 3). In data external to the ones here used, comparative *-er* stabilisation is likewise found for *y*-adjectives by the end of the 20th century – see Bauer (1994: 58–9) and Mondorf (2009: 140). Although Bauer (1994: 58–9) notes an exception to the *-er* bias for *y*-adjectives with a 'suffix *-ly*', these constitute less than a quarter of each of datasets-MOP and -Wells (41 of 253 tokens in each dataset), meaning that most of the *y*-adjectives in these datasets are indeed on their way towards *-er* stabilisation as they approach the start of the 20th century in Period 7 (representing the years 1901–1950). Empirical support for this stabilisation in datasets-MOP and -Wells alike suggest that the accepted statistical models from these datasets that obtain period effects bear external validation in Bauer (1994) and Mondorf (2009). Regardless then of any non-agreement between them of whether (and how) morphological complexity and penultimate weight impact *y*-adjective comparatives, the models accepted for datasets-MOP and -Wells may each be taken to reasonably explain comparative *y*-adjectives from a diachronic view.

In view of this, let us recap the account accepted for dataset-MOP, to examine its potential implication for our understanding of MORPHO-PHONOLOGICAL processes. In this account, a statistical significance in morphological predictions of *y*-adjective comparatives requires a simultaneous consideration, within the account, of penultimate weight. What is here suggested seems to be the existence of a morpho-phonological intersection of some sort, where the English comparative form is the output. Conventionally, morpho-phonological processes realise a phonological response that occur within the vicinity of morphological rule applications (Chomsky & Halle 1968; Kiparsky 1982; Mohanan 1986; Inkelas 2011). However, it is some distance from the demarcation between the penultimate and final syllables of *y*-adjectives where comparative forms show up, as periphrastic *more* or suffixal *-er*. Therefore, if there is indeed a morphological and penultimate weight intersection tantamount to a morpho-phonological process reflected in the account accepted for dataset-MOP, what is implied is that, in a diachronic context, the outcomes of morpho-phonological processes might well diffuse beyond the vicinity of these processes. How this diffusion may be operationalised is yet known, remains worth considering in future work and is subject to the syllable unit's stability in deriving comparative form outcomes.

Finally, if, at its core, this article highlights the importance of segment count for understanding *y*-adjective comparatives, then highlighted as well is the way English *y*-adjective comparatives may advance the English alternation scholarship. Coupled with findings that have shown character count to predict genitives (Ehret *et al*. 2014: 276; Rosenbach 2014: 227), segment count predictions of *y*-adjective comparatives here suggest that word length indicated by units smaller than the syllable may be fundamental to understanding English alternations. The notion of shared tenets that hold across English alternations is not at odds with analogies drawn before across these alternations. A study on children's use of English comparatives has presented evidence supporting an analogy of these uses with children's use of English agentives on the common ground of engaging transparency and productivity principles (Chua 2010: 100–2). A contextualisation of genitive alternation within a general pattern (Wolk *et al*. 2013) that applies also to dative alternation has analogised 'animacy effects' across genitive and dative constructions (Rosenbach 2014: 240–1). The proposal that English *y*-adjective comparatives and English genitives alike are predictable by units more fine-grained than the syllable thus adds on to existing studies that have surfaced shared tenets across English alternations. In so doing, the potential place of English *y*-adjective comparatives in advancing an understanding of English alternations should not be underestimated.

7 Conclusion

In conclusion, a key affirmation of this article is that phonemic segment count and the passage of time consistently predict *y*-adjective comparatives. Availed, additionally, is a quantifiable segment count threshold within which adjectival length remains important for the predictions. Though the morphological factor in predicting *y*-adjective comparatives remains less clear, where found, it coheres with prior notions of an -*er* bias with simplexes, and a *more* bias with complexes. The existence of the morphological factor seems, moreover, subject to syllabification principles that inform penultimate weight considerations in the analyses, positioning accounts of *y*-adjective comparatives, when taken together, as potential sites that make manifest non-congruences between syllabification theories. At least one of the accounts presented points, nonetheless, to potentially important insights for our understanding of morpho-phonological realisations. Turning out broad and varied implications, this article has indeed bitten off more than it can chew! At the heart of these implications are calls to consider our conception of adjectival length for accounts of English comparatives, to consider whether the syllable unit itself is sufficiently stable for these accounts, and to consider whether this stability eventually entails coming to terms with downstream implications of engaging different syllabification theories.

Author's address:
Deborah Chua
National Institute of Education
Nanyang Technological University
NIE5-B3-CRPP Office-Mini Lab 70, 1 Nanyang Walk, Singapore 637616
Singapore
Email: debchua2004@gmail.com; deborah.chua@nie.edu.sg

REFERENCES

Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green & John Fox. 2018. lme4: Linear mixed-effects models using 'Eigen' and S4. https://cran.r-project.org/web/packages/lme4/index.html (accessed 1 June 2018).

Bauer, Laurie. 1994. *Watching English change: An introduction to the study of linguistic change in standard Englishes in the twentieth century*. London and New York: Longman.

Bauer, Laurie. 2012. *Beginning linguistics*. UK: Palgrave Macmillan.

Cambridge University Press. 2020. *Cambridge dictionary*. https://dictionary.cambridge.org/ (accessed 30 July 2020).

Carr, Philip. 1999. *English phonetics and phonology: An introduction*. Oxford, UK: Blackwell Publishers.

Carter, Ronald & Michael McCarthy. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Chua, Deborah. 2010. Are three or more languages really too much to handle? Tracing the possibilities of multilingualism from Singaporean children's choice of adjectival comparatives. In Madelena Cruz-Ferreira (ed.), *Multilingual norms*, 95–112. Frankfurt: Peter Lang.

Chua, Deborah. 2018. Understanding comparative alternation in *y*-adjectives: What else might we need? *Journal of Linguistics* 54(3), 459–91.

Chua, Deborah. 2019. Comparative alternation in *y*-adjectives: Insights from self-paced reading. *Language and Cognition* 11(3), 373–402.

Chua, Deborah Fengyi. 2016. Comparative alternation in *y*-adjectives. Ph.D. disssertation, Victoria University of Wellington. http://researcharchive.vuw.ac.nz/handle/10063/5218 (accessed 18 December 2020).

Clark, Megan J. & John A. Randal. 2011. *A first course in applied statistics: With applications in biology, business and the social sciences*, 2nd edn. Auckland: Pearson.

Cruttenden, Alan. 1994. *Gimson's pronunciation of English*, 5th edn. New York: Edward Arnold.

Curme, George O. 1947. *English grammar: The principles and practice of English grammar applied to present-day usage*. New York: Barnes & Noble.

De Rosario-Martinez, Helios. 2015. Analyzing interactions of fitted models, 29 pp. https://cran.r-project.org/web/packages/phia/vignettes/phia.pdf (accessed 14 August 2020).

Dobson, E. J. 1968. *English pronunciation 1500–1700*, vol II: *Phonology*, 2nd edn. New York: Oxford University Press.

Ehret, Katharina, Christoph Wolk & Benedikt Szmrecsanyi. 2014. Quirky quadratures: On rhythm and weight as constraints on genitive variation in an unconventional data set. *English Language and Linguistics* 18(2), 263–303.

Giegerich, Heinz J. 2009. The English compound stress myth. *Word Structure* 2(1), 1–17.

Gordon, Matthew. 2002a. A phonetically driven account of syllable weight. *Language* 78(1), 51–80.

Gordon, Matthew. 2002b. Syllable weight, 41 pp. http://www.linguistics.ucsb.edu/faculty/gordon/syllableweight.pdf (accessed 30 October 2018)

Hammond, Michael. 1997. Vowel quantity and syllabification in English. *Language* 73(1), 1–17.

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1), 1–33.

Hayes, Bruce. 1981. *A metrical theory of stress rules*. New York: Garland.

Hilpert, Martin. 2008. The English comparative – language structure and language use. *English Language and Linguistics* 12(3), 395–417.

Hyman, Larry M. (2003). *A theory of phonological weight*. Stanford, California: CSLI Publications.

Inkelas, Sharon. 2011. The interaction between morphology and phonology. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, 2nd edn., 68–102. Oxford, UK: Wiley-Blackwell.

Jespersen, Otto. 1949. *A modern English of grammar on historical principles*, part 7: *Syntax*. Copenhagen: Ejnar Munksgaard.

Kiparsky, Paul. I982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, vol. 1, 131–75. Dordrecht: Foris.

Kytö, Merja & Suzanne Romaine. 1997. Competing forms of adjective comparison in Modern English: What could be more quicker and easier and more effective. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *Mémoires de la Société Néopilologique de Helsinki LII: To explain the present—Studies in the changing English language in honour of Matti Rissanen*, 329–352. Helsinki: Société Néopilologique.

Kruisinga, Etsko. 1932. *A handbook of present-day English*, part II: *Accidence and syntax*, vol. 3, 5th edn. Groningen: P. Noordhoff.

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

Leech, Geoffrey & Jonathan Culpeper. 1997. The comparison of adjectives in recent British English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 353–73. Helsinki: Société Néopilologique.

Lindquist, Hans. (1998). The comparison of English disyllabic adjectives in -*y* and -*ly* in Presentday British and American English. In Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling (eds.), *The major varieties of English: Papers from MAVEN 97*, 205–12. Växjö: Acta Wexionensia.

Lindquist, Hans. (2000). *Livelier or more lively?* Syntactic and contextual factors influencing the comparison of disyllabic adjectives. In John M. Kirk (ed.), *Corpora galore: Analyses and techniques in describing English*, 125–32. Amsterdam: Rodopi.

MacMahon, Michael K. C. 1998. Phonology. In Suzanne Romaine (ed.), *The Cambridge history of the English language*, vol. IV: *1776–1997*, 373–535. Cambridge: Cambridge University Press.

McGinnies, Elliott, Patrick B. Comer & Oliver L. Lacey. 1952. Visual-recognition thresholds as a function of word length and word frequency. *Journal of Experimental Psychology* 44(2), 65–9.

Mohanan, K. P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.

Mondorf, Britta. 2003. Support for more-support. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of grammatical variation in English*, 251–304. Berlin and New York: Mouton de Gruyter.

Mondorf, Britta. 2009. *More support for more-support: The role of processing constraints on the choice between synthetic and analytic comparative forms*. Amsterdam & Philadelphia: John Benjamins.

New, Boris, Ludovic Ferrand, Christophe Pallier & Marc Brysbaert. 2006. Reexamining the word length effect in visual word recognition: new evidence from the English Lexicon Project. *Psychonomic Bulletin & Review* 13(1), 45–52.

Palmer, Frank, Rodney Huddleston & Geoffrey K. Pullum. 2002. Inflectional morphology and related matters. In Rodney Huddleston & Geoffrey K.

Pullum (eds.), *The Cambridge grammar of the English language*, 1565–1620. Cambridge: Cambridge University Press.

Proquest. 1996–2013. *Literature Online*. http://lion.chadwyck.com/ (accessed 9 September 2013).

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik, J. 1985. *A comprehensive grammar of the English language*. London and New York: Longman.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/ (accessed 1 June 2018).

Rosenbach, Anette. 2014. English genitive variation – the state of the art. *English Language and Linguistics* 18(2), 215–62.

Ryan, Kevin M. (2016). Phonological weight. *Language and Linguistics Compass* 10(12), 720–33.

Schibsbye, Knud. 1965. *A modern English grammar*. London: Oxford University Press.

Schlüter, Julia. 2009. Cosonant of ‘vowel’? A diachronic study of initial <h> from Early Middle English to nineteenth-century English. In Donka Minkova (ed.), *Phonological weakness in English: From Old to Present-day English*, 168–96. Basingstoke, England & New York: Palgrave Macmillan.

Wells, J. C. 2000. *Longman pronunciation dictionary*. Essex, UK: Pearson Education.

Wells, J. C. 2002. Syllabification and allophony. https://www.phon.ucl.ac.uk/home/wells/syllabif.htm (accessed 16 December 2020).

Whitney, William Dwight. 1868. *Language and the study of language: Twelve lectures on the principles of linguistic science*. New York: Charles Scribner & Company.

Wolk, Christoph, Joan Bresnan, Anette Rosenbach & Benedikt Szmrecsanyi. 2013. Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica* 30(3), 382–419.

Yavas, Mehmet S., & Cynthia W. Core. 2001. Phonemic awareness of coda consonants and sonority in bilingual children. *Clinical Linguistics & Phonetics* 15(1–2), 35–39.

Zandvoort, R. W. & J. A. Van Ek. 1977. *A handbook of English grammar*, 7th edn. London: Longman.