ABSTRACT

This paper presents a novel approach to the phenomenon of intrusive-*r* in English based on analogy. The main claim of the paper is that intrusive-*r* in non-rhotic dialects of English is the result of the analogical extension of the *r*∼zero alternation shown by words such as *far*, *more* and *dear*. While this idea has been around for a long time, this is the first paper that explores this type of analysis in detail. Specifically, I provide an overview of the developments that led to the emergence of intrusive-*r* and show that they are fully compatible with an analogical approach. This includes the analysis of frequency data taken from an 18th century corpus of English compiled specifically for the purposes of this paper and the discussion of a related development, namely intrusive-*l*. The paper also presents a review of the evidence about the variability of intrusive-*r*, which serves as the basis of an evaluation of previous approaches. Once the notion of analogy is made formally explicit, the analogical approach becomes capable of providing a unified account of the historical development and the variability of intrusive-*r*. This is demonstrated through a computer simulation of the emergence of the phenomenon based on the 18th century corpus mentioned above. The results of the simulation confirm the predictions of the analogical approach.

1 INTRODUCTION

The phenomenon of intrusive-*r* in various dialects of English has inspired a large number of generative analyses and is surrounded by considerable controversy, mainly because of the theoretical challenges that it poses to Optimality Theory and markedness-based approaches to phonology (e.g. McCarthy 1993; Harris 1994; Halle & Idsardi 1997; Baković 1999; Uffmann 2007). As a consequence, a large proportion of the research on intrusive-*r* focusses on technical details of analysis, and shows little interest in the complex interactions that make this phenomenon particularly intriguing. Two areas that have been particularly neglected in generative discussions of the phenomenon are its historical development and its variability. The general indifference with respect to these areas stems directly from the underlying principles of the generative programme, according to which the primary goal of linguists is to construct *synchronic* models of *competence*, which therefore do not have to deal with issues of diachrony or performance. However, these restrictions have not proven particularly felicitous in the case of intrusive-*r*. First, the apparent unnaturalness of the phenomenon has led many researchers to claim that it is synchronically arbitrary (McCarthy 1991, 1993; Blevins 1997; Halle & Idsardi 1997; McMahon 2000), thereby implicitly acknowledging diachrony as a potential source of explanation for its behaviour in present-day dialects. Second, hardly any of the generative analyses dealing with intrusive-*r* make any attempts at incorporating the rich body of findings concerning its variability. As a result, it is often not clear whether these analyses are compatible with the observed patterns of variation.

This paper attempts to remedy this situation by proposing an analysis that accounts for the diachronic development and the variability of intrusive-*r* in a unified framework. The main claim is that the pattern of intrusion seen in Southern British English and other intruding dialects is the result of a process of analogical extension (cf. Jones 1964; Gimson 1980; Gick 1999, 2002; Bermúdez-Otero 2005). The first part of the paper (Section 2) develops this argument in more detail and relates it to facts about the history of intrusive-*r* and a number of related developments. This includes a detailed overview of the frequency distribution of word classes that played a part in the emergence of intrusive-*r*. The analogical account receives further support from a similar phenomenon occurring in certain North American varieties of English, namely intrusive-*l*. The second part of the paper (Section 3) brings in additional data on the variability of intrusive-*r* and shows that neither generative accounts, nor the analogical account in its simple form can easily accommodate them. Finally, the third part of the paper (Section 4) elaborates on the notion of analogy and develops an analogical model that can incorporate the fine-grained patterns of variation seen in intrusive-*r*. This model is used in a computer simulation of the emergence of the phenomenon, which takes a corpus of 18th century English as its input, and produces a dialect with a variable process of intrusion. The simulation can be seen as a synthesis of the main arguments of the paper: it connects the history of the phenomenon to its present variability by combining analogy with a token-based model of variation.

It will be useful to clarify the use of a number of key terms in this paper. Words which contain an inetymological *r* in intruding dialects of English are referred to as *r*-LESS (because of their lack of an *r* before the appearance of intrusive-*r*), and words with an etymological *r* as *r*-FUL. Historical dialects which had developed the conditions necessary for the emergence of intrusive-*r* are termed PRE-

INTRUSION DIALECTS. Preconsonantal and prepausal instances of *r* are simply referred to as CODA-*r*, although it should be noted that this term is used purely for convenience, since the present paper makes no assumptions about the syllabic status of this consonant in English.[1]

2 Analogy as the source of intrusive-*r*

This section provides a preliminary outline of the analogical approach to intrusive-*r* and present its main predictions regarding the conditions under which intrusion can develop (2.1). These predictions are then set against 18th century Southern British English (SBE), one of the dialects in which intrusive-*r* emerged (2.2). It will be shown that the conditions in pre-intrusion dialects are fully compatible with an analogical account. Moreover, some of the data presented in Section 2.2 remains unexplained, unless one assumes analogy as the source of intrusion.

2.1 Preliminary analysis

The term intrusive-*r* refers to an *r*∼zero alternation at the end of *r*-less words; the variant with a final *r* appears before a vowel and the one without *r* before a consonant or a pause. According to most traditional accounts, intrusive-*r* only involves words with a final [ɑː], [ɔː] or [ə] (e.g. *spa*, *law* and *pizza*, respectively). The epenthetic consonant may occur across words (e.g. *spa*[r] *is*, *law*[r] *and order*, *idea*[r] *of*) and word-internally as well (e.g. *withdraw*[r]*al*, *saw*[r]*ing*). The account presented in this paper focusses on the former case, although it could potentially be extended to the latter one as well. It is important to note that intrusive-*r* only appears in non-rhotic dialects, that is, dialects in which *r*-ful words also show a final *r*∼zero alternation (e.g. *scar*, *lore*, *Peter*). The alternation in these words is the result of a historical process of *r* Dropping before consonants and at the end of the word.

The main argument of this paper is that intrusive-*r* appeared in *r*-less words under the analogical influence of *r*-ful words. To put it slightly differently, the alternating pattern of *r*-ful words was analogically extended to the *r*-less group, resulting in a merger between the two classes, illustrated in (1) below (the shading illustrates the extent of the merger):

(1)	R-FUL		R-LESS		R-FUL	R-LESS
__{C, ‖}	V#		V#	⟹	V#	V#
__V	Vr#	→	V#		Vr#	Vr#

This insight also forms the basis of several previous analyses of the phenomenon, among them Jones (1964), Gimson (1980), Gick (1999, 2002), and Bermúdez-Otero (2005). While this approach is intuitively appealing, analogy has little explanatory power unless one specifies the exact conditions under which a pattern can be extended and demonstrates that these conditions are present in the language where the extension is suggested to occur. In the present case, this means (i) identifying the situations in which extension is likely to occur; (ii) giving a precise description of what qualifies as a potential analogical source in such a situation; and (iii) showing that such a situation arguably held in pre-intrusion dialects of English with the *r*-ful class being a suitable analogical source.

As for (i), most contemporary approaches to analogical extension assign a crucial role to similarity (Skousen 1989; Albright & Hayes 2003; Albright 2009): the likelihood of the extension of a pattern is a function of the similarity between the analogical source and the analogical target; the more similar they are, the more likely it is that the extension will occur. Turning now to (ii), the likelihood of a pattern to serve as the source of the extension is usually claimed to be proportional to its frequency (Bybee 2001), which means that the direction of the extension is determined by the relative frequencies of the two patterns: the analogical source will normally be of higher frequency than the analogical target. This means that the analogical approach makes two crucial predictions about *r*-ful and *r*-less words in pre-intrusion dialects:

Prediction 1 *r-ful and r-less words are similar.*

Prediction 2 *r-ful words are more frequent than r-less words.*

It should be noted that the exact roles of similarity and frequency in analogical models are left unspecified for the moment – it is simply assumed that the consensus of the recent analogical literature on the importance of these concepts is sufficient to treat them as essential components of an analogical account. This vagueness is remedied in Section 4, where these notions are substantiated and formalised within a computationally explicit framework.

2.2 Analogy and the history of intrusive-r

This section provides an overview of the historical developments related to intrusive-*r* and shows that the predictions of the analogical approach are borne out by the data. It is reasonable to assume a similar set of conditions to have held in all of the dialects where intrusive-*r* emerged (at least with respect to intrusion); therefore, this paper focusses on a single dialect, Southern British English (SBE), assuming that the same points could also be made for other dialects with intrusive-*r*. Since the first evidence of intrusive-*r* in SBE comes from Sheridan's *A Course of Lectures on Elocution* from 1762 (Sheridan 1762/1803), the standard dialect spoken in the south of England in the mid-18th century can be treated as a pre-intrusion dialect.

Let us first take a look at Prediction 1, which concerns the similarity between *r*-ful and *r*-less words. In present-day SBE *r*-ful and *r*-less words share an important structural feature: the set of final vowels appearing in preconsonantal and prepausal allomorphs of *r*-less words (i.e. [ə, ɔː, ɑː]) is a subset of the set of final vowels appearing in *r*-ful words in the same environment (i.e. [ə, ɔː, ɑː, ɜː]). The analogical approach to intrusive-*r* requires this structural similarity to be present in pre-intrusion dialects as well. Therefore, it needs to be shown that SBE acquired this particular distribution of final vowels no later than the middle of the 18th century. There are several pieces of evidence that suggest that this might well have been the case. The single most important factor in the emergence of the partial overlap between the two relevant classes of words is the loss of *r* in preconsonantal and prepausal position, which created the word-final *r*∼zero alternations exhibited by *r*-ful words in present-day English:

(2) *r* Dropping: r → ∅ / __{C, ‖}

__ ‖		__C		__V	
[wɔː‖]	'war'	[wɔːwəz]	'war was'	[wɔːrɪz]	'war is'
[stɑː‖]	'star'	[stɑːlaɪt]	'starlight'	[stɑːrɒn]	'star on'
[bɛtə‖]	'better'	[bɛtəðən]	'better than'	[bɛtərɪn]	'better in'

While Wells (1982) dates this change after 1750, Lass (2000) and McMahon (2000) argue that the decline of coda-*r* started much earlier, perhaps in Early

Modern English, with the weakening of preconsonantal and prepausal *r*, and was already ‘under way, producing variants in the speech community, before 1700’ (McMahon 2000: 234). For a detailed overview of the orthoepic evidence the reader is referred to McMahon (2000: 237-241). It is important to note that the historical sources do not point to a complete disappearance of coda-*r* in SBE: there is a marked lack of agreement among 18th century authors as to whether coda-*r* is pronounced or not, which suggests that *r* Dropping was variable at this stage. However, it is fair to assume that a considerable proportion of coda-*r*’s was now being dropped, creating a sufficient amount of overlap between the *r*-ful and the *r*-less classes to serve as the basis of analogical extension.

There are two further changes that played an important role in shaping the distribution of vowels before *r* termed Pre-*r* Lengthening and Pre-*r* Breaking by Wells (1982). These are illustrated in (3) and (4) below:

(3) Pre-*r* Breaking: ∅ → ə / {iː, eː, oː, uː, aɪ, aʊ}__ r

[biːr]	>	[bɪər]	‘beer’
[tʃeːr]	>	[tʃɛər]	‘chair’
[moːr]	>	[mɔːər]	‘more’
[ʃuːr]	>	[ʃʊər]	‘sure’
[faɪr]	>	[faɪər]	‘fire’
[taʊr]	>	[taʊər]	‘tower’

(4) Pre-*r* Lengthening: {ɑ, ɔ, ɜ} → {ɑː, ɔː, ɜː} / __ r{C, #}

[bɑrd]	>	[bɑːrd]	‘bard’
[hɔrn]	>	[hɔːrn]	‘horn’
[bɜrd]	>	[bɜːrd]	‘bird’

Since both of these changes were conditioned by the following *r*, it is clear that they had to predate the emergence of complete non-rhoticity. This argument is also supported by the historical record: Pre-*r* Breaking seems to have been a long and gradual process, starting as early as the 16th century (see Jespersen 1909; Jones 1989), and Pre-*r* Lengthening was also underway from at least the beginning of the 18th century (see McMahon 2000: 235-236).

It is remarkable that all the dialects where intrusive-*r* has emerged share these features with SBE: all intruding dialects are non-rhotic, and they all show the effects of Pre-*r* Lengthening and Pre-*r* Breaking. This can be interpreted as further evidence for the analogical approach: intrusive-*r* only emerges in dialects where there is a phonetic overlap between the *r*-ful and the *r*-less classes (i.e. where they have identical final vowels). Even more interestingly, the number of non-rhotic dialects without intrusion is conspicuously low. The only dialects where non-rhoticity does not entail intrusion are Southern American English (McDavid 1958: 322) and South African English (Wells 1982: 618). Incidentally, these dialects also share another important feature, namely that etymologically *r*-ful words are more or less consistently realised without a final *r* even in prevocalic position (McDavid 1958; Wells 1982). Once again, these observations receive a straightforward interpretation if we take analogy to be the source of intrusive-*r*: in these dialects, *r*-ful words have a non-alternating pattern, which cannot yield an *r*-zero pattern in *r*-less words through analogical extension.

Let us now turn to Prediction 2, which is about the frequency distribution of *r*-ful and *r*-less words. It was suggested that analogical extension only occurs if the source of the pattern is of higher frequency than the target of the extension. To test whether this relationship held between the *r*-ful and the *r*-less classes in pre-intrusion SBE, I compiled a 2 million word phonetically annotated corpus of early and mid 18th century English (henceforth CE18). The corpus consists of several 18th century novels (among them Samuel Richardson's *Clarissa* and Daniel De-

	R-LESS	R-FUL	RATIO
ə#	1,553	99,979	1:64.38
ɔː#	1,487	51,871	1:34.88
ɑː#	112	9,397	1:83.90
SUM	3,152	161,149	1:51.13

Table 1: *The token frequencies of* r-*ful and* r-*less words in the CE18 corpus*

foe's *Robinson Crusoe*) and all issues of *The Spectator* between 1711 and 1714. The automatic phonetic annotation of the text was based on the transcriptions of the CELEX database (Baayen et al. 1995). While this corpus provides us with more accurate data about the frequency distributions of the relevant word classes in 18th century English than any present-day corpus, it does have a number of drawbacks. The most problematic of these is that the transcriptions – being based on CELEX – reflect present-day pronunciations rather than 18th century ones. However, this may not be such a serious disadvantage, given that the phonological differences between 18th century English and Present-day English do not involve the main characteristics of the lexical classes that this analysis is based on.

The token frequencies of *r*-ful and *r*-less words are presented in Table 1. The reason for choosing token frequencies over type frequencies will become clear from the discussion in Section 4. The size of the *r*-ful class is two orders of magnitude greater than that of the *r*-less class, which confirms Prediction 2: the proposed analogical source is of considerably higher frequency than the analogical target. This is an important finding that has not been reported elsewhere in the literature.[2] Moreover, while the original merger of the *r*-ful and the *r*-less classes is an important part of previous analyses of the phenomenon, the frequency distributions found in SBE do not receive a straightforward interpretation in generative

	L-LESS	L-FUL	RATIO
ə#	45,932	41,282	1:0.90
ɔː#	12,219	110,874	1:9.07
ɑː#	4,237	704	1:0.16

Table 2: *The token frequencies of* l-*ful and* l-*less words in the CELEX corpus*

accounts based on rule inversion (cf. Section 3.2). This speaks strongly for an analogy-based account.

One final piece of evidence in support of the analogical approach comes from a related development in certain Mid-Atlantic dialects of American English, namely intrusive-*l* (see Gick 1999, 2002). Intrusive-*l* shows a very similar distribution and development to intrusive-*r*: it appears in words with a final [ɔː] (and to a much lesser extent with final [ə] and [ɑː]; cf. Gick 2002: 172) when the following word is vowel-initial (e.g. *draw it* [drɔːlɪt] vs. *draw them* [drɔːðəm]), and is only found in dialects where *l* has been lost in preconsonantal and prepausal position. This suggests that intrusive-*l* might also be a case of analogical extension based on the partial merger of previously *l*-ful and *l*-less forms (e.g. *drawl* and *draw*). If this is the case, we expect to find the same asymmetric frequency distribution for *l*-ful and *l*-less forms as for *r*-ful and *r*-less forms. This prediction is partially borne out by the data, as can be seen in Table 2 (the frequency counts are taken from the CELEX corpus). What is particularly striking here is that the frequency distribution necessary for the extension of the *l*-ful pattern (i.e. the analogical source is of higher-frequency than the analogical target) only holds for words ending in [ɔː]. If we adopt an analogical approach, the fact that the required frequency distributions are not found for words ending in [ə] and [ɑː] can serve as an explanation for the resistance of these words to intrusive-*l* in the dialect described by Gick (2002).

Note that the optional extension of intrusive-*l* to words ending in [ɑː] (and potentially even [ə]) is also not inconsistent with the present approach. [ə], [ɔː] and [ɑː] are all non-high vowels from the central and back region of the vowel space, which means that they are highly similar to each other and highly dissimilar to all other vowels appearing in word-final position. Since analogy is based on similarity, some leakage is expected across phonetically similar vowel categories. Table 2 also shows that the overall frequency of [ɔː]-final words is more than twice as high as that of the other two groups. This frequency difference ensures that [ɔː]-final words will play a dominant role in determining the effects of the leakage. Thus, it is even expected that some words ending in [ə] and [ɑː] should follow the lead of [ɔː]-final words. While this evidence is clearly circumstantial with respect to intrusive-*r*, the fact that analogy provides a unified explanation for two independent processes of intrusion in English and makes valid predictions for both is a strong argument for adopting an analogical approach.

To conclude this section, let us sum up its main points. It has been shown that intrusive-*r* conforms to the two main predictions of the analysis presented in Section 2.1: the *r*-ful class and the *r*-less class share essentially the same set of final vowels and the *r*-ful class has considerably higher token frequency than the *r*-less class. Moreover, we have also seen that the analogical approach can provide a straightforward explanation for a number of related issues: the absence of intrusive-*r* in Southern American English and South African English and the development of intrusive-*l* in Mid-Atlantic varieties of English. While these findings argue strongly for an analogical approach, they do not necessarily exclude alternative explanations along generative lines. Therefore, the next section introduces further data that can help us decide between the competing models.

3 The variability of intrusive-*r*

In the preceding section, a number of simplifying assumptions were made to allow for a more streamlined presentation of the issues relevant to the development of intrusive-*r*. Specifically, the relevant word classes and the phenomenon of intrusive-*r* itself were presented as if they behaved categorically and variation was referred to only occasionally. However, it appears that the actual situation is somewhat more complicated. The available evidence suggests that *r* Dropping and the emergence of intrusive-*r* were gradual processes creating a considerable amount of variation in both *r*-ful and *r*-less words, with the latter class still showing a great deal of variability. This section explores this variation in more detail (3.1), and discusses its consequences for previous accounts and the present analysis (3.2).

3.1 Empirical research on intrusive-r

One result that seems to emerge consistently in empirical studies of intrusive-*r* and other related phenomena is that the incidence of final *r* in prevocalic position is highly variable in both the *r*-ful and the *r*-less classes. Table 3 presents the overall proportion of rhotic realisations in prevocalic position as reported in different studies of the behaviour of *r*. While these figures mask a great amount of interpersonal and sociolinguistic variation, they clearly demonstrate a high degree of variability within the *r*-ful and *r*-less classes (which also exists at the level of individuals, as is shown in Mompeán-Gonzalez & Mompeán-Guillamón 2009

	R-FUL	R-LESS
Bauer (1984)	80%	28%
Foulkes (1998) (Derby)	88%	57%
Foulkes (1998) (Newcastle)	58%	9%
Mompéan-G. & Mompéan-G. (2009)	58%	32%
Sóskuthy (2009)	75%	58%

Table 3: *The percentages of rhotic realisations in prevocalic position in the* r*-ful and the* r*-less classes reported in different empirical studies of* r*-liaison.*

and Sóskuthy 2009). The results of these studies also reveal a marked difference between the two classes: *r*-ful words are more likely to be realised with a final *r* in prevocalic position than *r*-less words (or, in traditional terms: linking-*r* is more likely to occur than intrusive-*r*).

It is also possible to isolate more fine-grained patterns of variation in the realisation of intrusive-*r*. Here are some of the main factors that have been suggested to influence the incidence of intrusive-*r*: social class (Foulkes 1998; Hay & MacLagan 2010), gender (Bauer 1984; Foulkes 1998; Hay & MacLagan 2010), age (Foulkes 1998; Hay & MacLagan 2010), the lexical identity of the target word (Sóskuthy 2009; Hay & MacLagan 2010), the quality of the preceding vowel (Jones 1964; Gimson 1980; Hay & MacLagan 2010; Bauer 1984) and the presence of an *r* in the onset of the final syllable (Jones 1964; Wells 1982). Since the analogical account developed in the present paper does not address any of the sociolinguistic aspects of intrusive-*r*, the discussion below focusses on the last three factors.

Let us first look at the lexical identity of the target word. Hay & MacLagan (2010) and Sóskuthy (2009) find that the probability of intrusive-*r* can be significantly different across words within the *r*-less class, suggesting that there are word-specific patterns in the realisation of intrusive-*r*. To give an example, Hay & MacLagan (2010) show that the word *sofa* is more likely to take intrusive-*r* than the word *bra*, which in turn is more likely to take intrusive-*r* than the word *ma*. The presence of such tendencies suggests that the emergence of intrusive-*r* is a lexically diffuse process (cf. Chen & Wang 1975; Labov 1994), which is potentially still ongoing. However, neither Hay & MacLagan (2010) nor Sóskuthy (2009) control for phonetic factors, which means that the lexical conditioning of intrusive-*r* could be an artefact due to the different phonetic environments embodied in individual lexical items. For instance, the difference between *sofa* and *bra* might simply follow from the fact that they have different final vowels ([ə] and [ɑː], respectively). Since the discussion of the two remaining factors below suggests that the phonetic makeup of the target word has little influence on the probability of intrusive-*r*, this does not seriously weaken the evidence for word-specific tendencies.

There is little agreement in the literature on the effect the preceding vowel has on the likelihood of intrusive-*r*. Traditional descriptive works on the phonetics and the phonology of SBE such as Jones (1964) and Gimson (1980) claim that *r* is more likely to be inserted after [ə] than it is after [ɑː] and [ɔː]. Uffmann (2007: 470) presents a slightly different grouping of final vowels: he claims that there exists a class of speakers who only produce intrusive-*r* after [ə] and [ɑː], while other speakers intrude after all non-high vowels (based on Wells 1982). Yet another alleged pattern involves higher rates of intrusive-*r* after [ɔː] than after [ə] and [ɑː] (see Hay & Warren 2002 and Hay & Sudbury 2005 for New Zealand English). Finally, Foulkes (1998), Sóskuthy (2009) and Mompeán-Gonzalez & Mompeán-Guillamón (2009) find no correlation between the quality of the final

	Quantitative	Dialect	Pattern
Jones (1964)	no	SBE	ə > ɑː, ɔː
Gimson (1980)	no	SBE	ə > ɑː, ɔː
Uffmann (2007)	no	SBE	ə, ɑː > ɔː
Hay & Sudbury (2005)	yes	early NZE	ə, ɑː > ɔː
Hay & Warren (2002)	yes	NZE	ɔː > ɑː, əː
Foulkes (1998)	yes	Derby, Newcastle	–
Mompéan-G. & Mompéan-G. (2009)	yes	SBE	–
Sóskuthy (2009)	yes	SBE	–

Table 4: *A summary of previous findings on the influence of the preceding vowel on intrusive-r. The second column indicates whether the findings are based on quantitative analyses or informal observations, the third column indicates the dialect for which the observations were made and the last column indicates the ranking of different environments with respect to their likelihood of taking intrusive-r (or nothing when no differences are reported).*

vowel and the likelihood of intrusive-*r*. Table 4 presents a summary of these results. Interestingly, it appears that most of the tendencies reported for intrusive-*r* in studies based on impressionistic observations (Jones 1964; Gimson 1980; Wells 1982; Uffmann 2007) disappear in statistically more reliable quantitative studies of the phenomenon (Foulkes 1998; Mompeán-Gonzalez & Mompeán-Guillamón 2009; Sóskuthy 2009). The only exception is New Zealand English, where statistically more robust patterns have been found (Hay & Warren 2002; Hay & Sudbury 2005). However, the difference between early NZE and present day NZE suggests that even these patterns are rather unstable. Therefore, the results reported in the literature do not seem to support any substantive hypothesis about the influence of the final vowel on the development of intrusive-*r*.

The status of the second phonetic factor, namely the presence or absence of a tautosyllabic *r* also seems somewhat questionable. Jones (1964) and Wells (1982) suggest that intrusive-*r* might be less likely when there is an *r* in the onset of the final syllable, as in *zebra* and *draw*. However, this claim does not seem to be supported by empirical studies of intrusive-*r*: none of the studies that investigate the role of tautosyllabic *r* find any significant effects associated with it (Hay & Sudbury 2005; Mompeán-Gonzalez & Mompeán-Guillamón 2009). The only study that finds a limited amount of support for such an effect is Foulkes (1998), which does indeed report a lower proportion of *r*-ful realisations in words with a tautosyllabic *r*. However, even Foulkes himself urges caution over the interpretation of his results given the small number of relevant tokens and the absence of statistical significance. In sum, the claim that a tautosyllabic *r* might influence the realisation of intrusive-*r* has not been borne out by empirical studies of the phenomenon.

Although all of the results reported so far are based on studies of present-day dialects of English, they have some bearing on historical accounts of intrusive-*r* as well. First of all, the variability of intrusive-*r* in all present-day dialects that have been investigated quantitatively suggests that the emergence of the phenomenon is unlikely to have been abrupt. Second, the existence of word-specific patterns can be taken as evidence for lexical diffusion: if intrusive-*r* had emerged at the same rate across the lexicon, such effects would not be expected. Finally, the inconsistency of the reports on the phonetic conditioning of intrusive-*r* argues against historical and synchronic accounts that rely strongly on the existence of such tendencies.

Before presenting a summary of the findings on the variability of intrusive-*r*, there is one final study that should be mentioned: Hay & Sudbury (2005). The authors of this paper take a diachronic approach to the questions related to intrusive-*r*, which makes their findings particularly relevant to the present investigation. Hay & Sudbury (2005) examine the incidence of linking and intrusive-*r* in the speech of several generations of New Zealanders born between 1850 and 1930, based on a collection of audio recordings. They find a high degree of variability for intrusive-*r* across all age groups, which lends support to the assumption that intrusive-*r* has been a variable phenomenon from the beginning. Moreover, their study shows that even partially rhotic speakers can exhibit various degrees of intrusive-*r*, and that the incidence of intrusive-*r* is significantly correlated with the speakers' degree of rhoticity: '[i]ntrusive /r/ increases as rhoticity declines' (Hay & Sudbury 2005: 813).

The detailed overview above suggests that accounts of the emergence of intrusive-*r* should be capable of capturing at least the following tendencies with regard to intrusion:

(5) a. gradual emergence
    b. interaction between rhoticity and intrusive-*r*
    c. lexical diffusion

The simple analogical account sketched in the previous sections does not make any particular predictions with respect to these observations. In fact, the diagram in (1) and the short summary of the analogical account presented in section 2.1 may seem to suggest that analogy predicts a categorical pattern with no variation at all. This is certainly not the case: once we make the notion of analogical extension more explicit, it becomes possible to account for the observed patterns of variation. This task is taken up in Section 4. Before that, however, it will be useful to see how previous accounts of intrusive-*r* fare with respect to the observations presented above.

3.2 Previous accounts of intrusive-r

This section presents a brief overview of previous accounts of intrusive-*r*. While there is certainly much to say about the wide range of theoretical devices that have been deployed in previous analyses, this review focusses exclusively on the ability of such accounts to make accurate empirical predictions about the variability and the history of intrusive-*r*. This might be seen as an unfair treatment of these analyses, most of which take an explicitly synchronic and categorical approach to intrusive-*r*. However, the choice to focus on linguistic competence does not render arguments based on diachrony and variation irrelevant: any competence-based account should at least be able to represent intermediate stages in the development of the pattern. As it will be shown, most existing analyses of intrusive-*r* cannot do that, and they also do not provide a motivation for the observed changes.

It is possible to group all existing accounts into three larger classes, based on their underlying structure and their general predictions with respect to variation: deletion-based, insertion-based and analogy-based (this division is based partly on McMahon et al. 1994). Most analyses belonging to a given group can be treated together, since their predictions usually only differ in ways that are not relevant to the present discussion. The rest of this section looks at each of these groups in more detail. Note that the discussion below does not distinguish between rule-based and constraint-based analyses; this is because the crucial differences among the analyses lie in the choice of underlying forms and the number of mechanisms (i.e. rules or constraints) used in deriving the surface forms, but not in the way these mechanisms are implemented. For simplicity's sake, rule-based terminology is used throughout the discussion.

Deletion-based accounts (Donegan 1993; Harris 1994; Gick 1999, Gick 2002; Bermúdez-Otero 2005) are based on the assumption that *r* is present in the underlying representation of *r*-ful and *r*-less forms alike, and the *r*-zero alternations observed in both sets are the result of a rule that deletes *r* preconsonantally and prepausally. Since deletion-based accounts assume that the grammar of pre-intrusion dialects is identical to that of intruding dialects, the locus of the change can only be the lexicon: forms that end in a non-high vowel in pre-intrusion dialects acquire a final *r* in their underlying representation (e.g. 'spa' /spɑː/ becomes /spɑːr/). This analysis is fully compatible with a diachronic process of lexical diffusion, since it allows different words to adopt *r*-ful underlying representations at different times. However, it also predicts that there will be no intra-speaker variability in the rate of intrusive-*r* in a given word: if the underlying representation of the word contains a final *r*, it will always be produced with *r* in prevocalic position; if its underlying representation contains no final *r*, it will never be produced with *r*. This is true even if one assumes variable rules (cf. Labov 1972), since *r*-deletion can only apply in preconsonantal and prepausal position (i.e. 'spa is' /spɑːr əz/ will be invariably realised as [spɑːr əz] even if *r*-deletion only applies 70 per cent of the time). Moreover, deletion-based accounts also do not provide a satisfactory explanation for why novel underlying forms should be adopted at all.

Conversely, insertion-based accounts assume that the *r*-zero alternations in *r*-less forms can be accounted for by a rule of r-insertion, which either exists alongside the original rule of *r*-deletion (McCarthy 1991, 1993; Blevins 1997; Halle

& Idsardi 1997; Antilla & Cho 1998; Baković 1999; Uffmann 2007) or replaces it entirely (Vennemann 1972; Kahn 1976; McMahon et al. 1994; McMahon & Foulkes 1995; McMahon 2000). It is clear that both types of models represent an improvement over deletion-based accounts inasmuch as they are compatible with a variable insertion rule that can account for the variation in intrusive-*r*. However, there are a number of slight differences in their predictions, which makes it necessary to treat them separately.

Insertion-only analyses need to assume that (i) the original deletion rule is replaced by its inverse (i.e. a rule of insertion) and (ii) etymologically *r*-ful forms are reanalysed as underlyingly *r*-less (e.g. 'spar' /spɑːr/ becomes /spɑː/). Although the change that gives rise to intrusive-*r* takes place in the grammar (since the structural description of the inverted rule also covers *r*-less forms), intrusion-only analyses can also incorporate lexical diffusion by adopting Kiparsky's 'lexical diffusion as analogy' approach (1995). This could be achieved by assuming that *r*-less forms are initially lexically specified in a way that the inverted rule does not apply to them (e.g. they are marked as arbitrary exceptions) and that these lexical specifications are removed on an item-by-item basis through a process of lexical simplification. While the insertion-only account might be successful at capturing lexical diffusion, it runs into serious trouble when it comes to the interaction between rhoticity and intrusive-*r*. Essentially the same problem arises as in the case of deletion-only analyses: the restricted scope of the insertion rule does not allow it to control variability outside the prevocalic environment. Preconsonantal and prepausal forms are therefore predicted to be either fully rhotic or fully non-rhotic depending on their lexical specification (e.g. 'spar with' /spɑː wɪð/ will always surface as [spɑːwɪð] and /spɑːr wɪð/ as [spɑːrwɪð]). This goes against the findings reported in Hay & Sudbury (2005), which make it clear that variable rhoticity is possible in pre-intrusion dialects.

This problem does not arise in insertion-plus-deletion accounts, since the co-existence of the two rules in the grammar makes it possible to control the variation in prevocalic vs. preconsonantal and prepausal position separately. However, there are two important objections against insertion-plus-deletion analyses, one of them empirical and the other theoretical. First, even Kiparsky's lexical specification approach cannot yield word-specific rates of intrusion: a word specified as exempt from *r*-intrusion will never surface with a final *r* in prevocalic position, and the rate of intrusion will be uniform across all *r*-less words without such a specification (depending solely on the variability of the *r*-intrusion rule). This is problematic inasmuch as it has already been established in Section 3 that present-day varieties of English do, in fact, show word-specific tendencies with respect to intrusive-*r*. It should be noted that this problem is not specific to insertion-plus-deletion accounts: generative models of phonological competence cannot comfortably accommodate such word-specific effects.

The second problem is related to the motivation for establishing an inverted rule. While the insertion-plus-deletion approach does not explicitly specify the conditions under which rule inversion can take place, it is reasonable to assume that one such condition is the lack of robust evidence against the inverted rule. This condition does seem to hold in the case of intrusive-*r*, given the extremely low frequency of *r*-less forms in prevocalic position. However, it is not clear why a learner would want to add an inverted rule to a grammar that already contains a rule of deletion: the new rule does not simplify the grammar or the lexicon in any sense and all the forms that match its structural description (i.e. *r*-less forms in prevocalic position) show a behaviour (i.e. full non-rhoticity) that constitutes strong counterevidence against it. Proponents of insertion-plus-deletion analyses have argued that the emergence of the *r*-insertion rule is motivated by more general constraints on hiatus (Antilla & Cho 1998; Baković 1999; Uffmann 2007) or vowel-final words (McCarthy 1991, 1993; Blevins 1997). Although this might be taken as an explanation for the establishment of a rule that contradicts the data available to the speakers, it is still not clear how exactly the already existing *r*-deletion rule could condition such a restructuring of the grammar. One question that arises is why it is not possible for speakers without a rule of deletion to establish a similar rule. As it has been shown above, the analogical account has a straightforward answer to this question: the alternation can only be extended to the *r*-less class if it already exists in *r*-ful words.

Finally, a few authors have suggested – similarly to the present paper – that the source of intrusive-*r* is analogy. Some of the earliest 20th century descriptions of intrusive-*r* propose word-based analogy as a potential mechanism that might have led to the emergence of intrusion (Jones 1964; Gimson 1980). Unfortunately, these accounts do not provide any further details that might help us understand how exactly analogy could yield intrusive-*r* and (as it has already been pointed out in Section 2.1) also do not specify the conditions under which analogical extension can occur. The same criticism applies to Gick (1999, 2002) and Bermúdez-Otero (2005), who also suggest that the diachronic source of intrusive-*r* is analogy (although they do propose a synchronic account for the phenomenon).

In sum, no existing account of intrusive-*r* presents a fully satisfactory explanation for its diachrony and variability. Deletion-based analyses cannot explain the variation in *r*-less forms, while insertion-only analyses have the same problem with *r*-ful forms. Insertion-plus-deletion analyses avoid these difficulties, but they run into trouble when it comes to the motivation for the *r*-insertion rule and the link between deletion and insertion. Furthermore, none of these models are capable of incorporating gradient word-specific effects. Finally, even existing analogical accounts fail to elaborate on the details of the diachronic mechanisms that resulted in intrusive-*r* in present-day intruding dialects. These problems do not arise in the token-based analogical model presented below, which addresses both the issues of diachrony and variation in a unified framework.

4 Analogy and variation

In the preceding sections, two different sets of evidence were reviewed, leading to the following conclusions: (i) intrusive-*r* in SBE is readily explained by models based on analogical extension and (ii) it shows certain patterns of variation that are problematic for previous accounts of the phenomenon. This section takes these two seemingly unrelated observations and suggests a model that integrates them in a single framework. To achieve this goal, it will be necessary to provide a more explicit definition of analogy itself and briefly review previous approaches to analogical extension (4.1). Since the structure of existing analogical frameworks makes them incapable of handling variation at the level of individual words, a different approach is introduced (4.2). This model is then tested through a computer simulation based on real data from the 18th century corpus described above (4.3).

4.1 Previous approaches to analogy

While the term analogy is used in a variety of ways in the literature (see Hock 2003 for an overview), this paper focusses on one particular mechanism which seems to serve as the basis of most computationally implemented models of analogy, namely FOUR-PART ANALOGY. Four-part analogy consists in the extension of a certain relationship between a pair of forms to another pair of forms, where the members of the two pairs bear the same structural or semantic relationship to each other. An example is given in (6) below:

(6)		[singular]		[plural]	
	BOW	[baʊ]	∼	[baʊz]	
			↓		
	COW	[kaʊ]	∼	?	(= [kaʊz] < [kaɪn])

The four edges of the analogical rectangle will be referred to as follows: the SOURCE (BOW), the TARGET (COW), the KNOWN ENVIRONMENT ([singular]) and the GIVEN ENVIRONMENT ([plural]). The corners of the rectangles can be identified by referring to the two edges that meet there: for instance, [baʊ] is the source in the known environment and [kaʊz] (the form that we obtain through analogical extension) is the target in the given environment (this will also be referred to as the GIVEN FORM). The particular relationship that is extended in (6) can be described as {$x \sim x$ + [z]}. This relationship clearly yields [kaʊz] when applied to the target in the known environment, that is, [kaʊ].

This type of analogy can also be used to model the extension of the *r*-ful pattern to an *r*-less word:

(7)

	__C		__V	
DEAR	[dɪə]	∼	[dɪər]	
		↓		
IDEA	[aɪdɪə]	∼	?	(= [aɪdɪər] < [aɪdɪə])

As it is pointed out by Albright (2009), this type of formalism does not impose any restrictions on the choice of the analogical source: in the example in (7), the lexeme DEAR is used, but other lexemes, such as MARIA, CAT or SMURF could equally well have been used, in which case no change would have taken place (as these lexemes do not show an alternating *r*∼zero pattern). This is clearly problematic: the transition from the analogical target to the analogical source is arguably guided by frequency and similarity, as it has been noted in Section 2.2.

Most computationally implemented models of analogy take a somewhat simplified version of the four-part analogical mechanism as their starting point and use a number of extra mechanisms to ensure that both similarity and frequency have an effect on the choice of the analogical source. It will be useful to take a brief look at a particular class of such models, namely INSTANCE-BASED LEARNERS, some examples of which are the GENERALIZED CONTEXT MODEL (GCM; Nosofsky 1986, 1988), ANALOGICAL MODELING (AM; Skousen 1989; Skousen et al. 2002) and the TILBURG MEMORY-BASED LEARNER (TiMBL; Daelemans et al. 2007).

Instance-based learners are based on the assumption that the behaviour of a given item can be determined by comparing it to similar items within the dataset. The dataset for an instance-based learner could consist of a list of phonetically transcribed types from the lexicon of English, where each type is associated with a particular behaviour in prevocalic position, as exemplified in Table 5. The types

LEXEME	VARIABLES	BEHAVIOUR
bread	=, b , r , ɛ, d	{+∅}
spin	=, s , p , ɪ , n	{+∅}
city	=, s , ɪ , t , i	{+∅}
idea	a , ɪ , d , ɪ , ə	{+∅}
law	=, =, =, l , ɔː	{+∅}
four	=, =, =, f , ɔː	{+r}
better	=, b , ɛ , t , ə	{+r}
star	=, =, s , t , ɑː	{+r}

Table 5: *Dataset for selection of patterns of alternation in English*

are represented as a set of variables, which, in this case, are the last five sounds of each occurrence (‘=’ means non-specification for a given feature). Instance-based learners can use this dataset to predict the behaviour of any item that is specified using the same variables. This could be a new item, which is not present in the original dataset (this would be similar to a learner trying to establish a certain pattern for a nonce-form or a loanword) or an item from the dataset itself (as in the case of analogical extension, where an existing pattern is replaced by a new one).

The model’s prediction is based on the behaviour of items that are similar to the given form. The precise calculation of similarity values differs from model to model, but in most cases it is a function of the number of overlapping variables, where certain variables can have a greater influence than others. Thus, IDOL [a, ɪ, d, ə, l] and DEAR [=, =, d, ɪ, ə] both share three variables with IDEA [a, ɪ, d, ɪ, ə], but the last variable can be given a greater weight in determining similarity values, as it is more relevant to the task at hand than, say, the first variable.[3]

Frequency influences the predictions of instance-based learners in a less direct way. The likelihood of any individual form to serve as the analogical source or

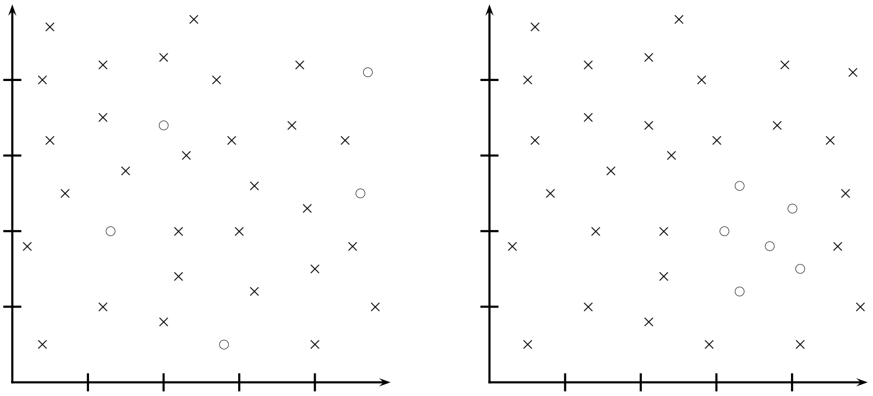


Figure 1: *Left panel: items with the two behavioural patterns are evenly distributed in the feature space; Right panel: items with the low-frequency pattern form a tight group.*

be included in the analogical set is solely determined by its similarity to the given form. However, since a high-frequency behavioural pattern is necessarily better represented in the dataset than a low-frequency one, it will have a greater chance of influencing the outcome of the prediction, provided that the items are relatively evenly distributed in the feature space defined by the variables. For instance, if there are 90 items with behaviour *A* and only 10 items with behaviour *B*, any random point in the feature space will be likely to be surrounded by a majority of items with behaviour *A*. The only scenario in which behaviour *B* can have any significant effect on the outcome of the prediction is when the items with behaviour *B* form a tight group (sometimes referred to as a 'gang'; cf. Bybee 2001) within the feature space, that is, when they are consistently more similar to each other than to items with behaviour *A*. Figure 1 provides an illustration of a dataset where the distribution of the items in the feature space is independent of their behaviour (left) and another dataset where items with a low-frequency pattern form a tight group (right).

This might also help us understand how an instance-based learner could be used to model the emergence of intrusive-*r*. Pre-intrusion dialects of English exemplify the evenly distributed scenario (i.e. *r*-less words are randomly dispersed among members of the *r*-ful class), whereas dialects in which the partial merger between the *r*-ful and the *r*-less classes did not take place exemplify the second scenario, with *r*-less words forming a tight group. Therefore, in a pre-intrusion dialect the outcome of the prediction will be more strongly influenced by the *r*-ful pattern than by the *r*-less one owing to the higher frequency of the former; *r*-ful forms will simply have a greater chance of being included in the analogical set or being chosen as the analogical source. The model will tend to predict an *r*-ful pattern of behaviour even for *r*-less words, that is, analogical extension will take place. However, in dialects where the *r*-ful and the *r*-less classes are fully distinguishable, words within the *r*-less class will be more similar to each other than to words within the *r*-ful class. This similarity will counterbalance the higher frequency of the *r*-ful pattern and result in the retainment of the distinction between the two classes.

While instance-based learners can capture some crucial aspects of the analogical extensions that led to the emergence of intrusive-*r*, their success hinges on a considerable simplification: they do not make a distinction between the known environment and the given environment, and they assign a single pattern of behaviour to each item. By doing so, they essentially reduce the problem of analogical extension to a simple categorisation task: a stimulus represented by a feature vector has to be assigned a category label, which is a certain pattern of behaviour in this case. This simplification comes at a price: we are forced to discard all information about variation below the word level. Each type is assigned a single feature vector and a single pattern of behaviour ({+∅} or {+r}). As a result, a number of arbitrary decisions have to be made, which lead to considerable conceptual and empirical difficulties.

First of all, as types are abstractions over a set of tokens, they often cannot be associated with a unique representation. Choosing the citation forms of the types in Table 5 was a completely arbitrary decision; if the data set was composed of prevocalic forms, there would be no analogical extension at all (as *r*-ful and *r*-less words are distinct in prevocalic position in pre-intrusion dialects). In fact, it might be just simply impossible to assign any phonetic representation to types which have several alternants. If a type is a collection of properties shared by a number of tokens, the precise phonetic forms of the individual tokens are arguably not part of it when they differ from token to token.

Another related problem is that types often cannot be associated with a unique behaviour in a natural linguistic setting, having variable outcomes instead. For instance, a given *r*-less word might follow an alternating pattern 40 per cent of the time and a non-alternating pattern 60 per cent of the time. This is certainly the case for intrusive-*r*, where individual words are often realised variably, and the exact proportions of the variants might differ across words. In a strict type-based approach, there is no straightforward way of representing this variation, which is clearly a problem.

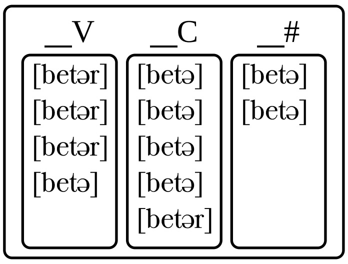
Finally, instance-based learners raise an important question: it is not clear how the extensions described above lead to intrusive-*r*. Let us assume that any production of intrusive-*r* is the result of an active process of analogical extension. For instance, when a speaker of present-day SBE utters the phrase ‘saw it’ [sɔːrɪt], the non-etymological *r* appears due to active analogical influence from *r*-ful words.

If this was indeed the case, the rate of intrusive-*r* would be a function of (i) the probability of using analogy to predict the production of a given form and (ii) the probability of choosing a word with an alternating pattern as the analogical source. Since the first factor is unlikely to be very strong (i.e. it is unreasonable to assume that speakers rely on analogy for the majority of their productions), analogical extension should not be able to produce more than a sporadic pattern of intrusion. However, intrusive-*r* is a robust pattern that can potentially affect the production of *r*-less forms more than 50 per cent of the time (cf. Table 3). The question, then, is how these sporadic extensions could give rise to a consistent pattern of intrusion. The analogical account presented above does not provide an answer to this question.

In sum, instance-based learners suffer from problems that are highly reminiscent of the shortcomings of the accounts discussed in Section 3.2. Although analogy provides an intuitively appealing explanation for why intrusive-*r* emerged in the first place, existing analogical models cannot capture the variability of the phenomenon, and they also do not make the role of analogy entirely explicit. Therefore, the next section proposes a new model, which completely eschews the type-based view.

4.2 Token-based analogy

The model proposed in the present section combines analogy with token-based lexical storage. The main idea is that the sporadic changes produced by analogical extension are recorded in the lexicon, which means that they can accrue over several generations and lead to more robust patterns (Wedel 2004, 2007; Oudeyer 2006). The model of lexical-storage used in the present paper is exemplar the-



__V	__C	__#
[betər]	[betə]	[betə]
[betər]	[betə]	[betə]
[betər]	[betə]	
[betə]	[betə]	
	[betər]	

Figure 2: *Exemplar cloud of* r-*ful word*

ory (Bybee 2001; Pierrehumbert 2001), although it should be pointed out that any model that can represent the same amount of information about word-specific patterns of variation should be able to produce similar results. The basic idea in exemplar theory is that linguistic categories such as words and sounds are represented directly by detailed memory traces of actual utterances. This means that all tokens of use are stored in the lexicon linked to the specific context (semantic, phonological, social, etc.) in which they are used. These tokens then serve as the basis of both production and perception. A model of this type represents variation directly in so-called clouds of exemplars, as in Figure 2.

Whenever a new form is produced by analogical extension, it is stored in the lexicon of both the speaker and the listener. This has clear implications for analogical change. Since production is based on the proportions of previously heard variants, every case of analogical extension increases the overall rate at which the new variant is produced. This leads to a steady increase of analogically remodelled forms even if analogical extension only takes place sporadically. The present paper exploits this property of exemplar-based storage in the following way: an analogy-based production mechanism is used to predict the pronunciation of each item, and the output of this model is used as the input for the next generation. This procedure is repeated several times (this is called ITERATED

LEARNING; cf. Brighton 2003; Kirby et al. 2007).

The rest of this section explains how these insights can be used to build a computer simulation of the emergence of intrusive-*r*. The initial input of the simulation is a list of tokens from the CE18 corpus, represented as ordered triplets consisting of the phonetic form of the item, the lexeme the item belongs to and the phonetic environment it appears in (which can be __C, __V or __#, depending on the first sound of the following word). Thus, a preconsonantal production of IDEA will be encoded as follows: {[aɪdɪə], IDEA, __C}. The model goes through each item in the list and predicts a pronunciation for it using the analogical production mechanism described below (the predicted pronunciation may or may not be identical to the stored one). These pronunciations are stored in the lexicon of the next generation, which starts its own round once the first generation has produced all the items in the dataset. This process can be repeated indefinitely, but we will see that 50 rounds are sufficient for our purposes.

The crucial step in this process is, of course, the prediction of pronunciations for the items in the dataset. These predictions are based on a four-part analogical mechanism, as shown in Figure 3. Here is a step-by-step description of this process. The input of the analogical prediction is an ordered pair consisting of the lexeme the item belongs to and its environment – in Figure 3, this is {IDEA, __V}. This determines the analogical target (IDEA) and the given environment (__V). To complete the analogical rectangle, we first have to find another environment (the known environment) with at least one token of the target. In our example, the known environment is __C. Now, a random token of the target lexeme is chosen in the known environment (step 1 in Figure 3), which will serve as the basis of our choice of the analogical source in the known environment (step 2). The tran-

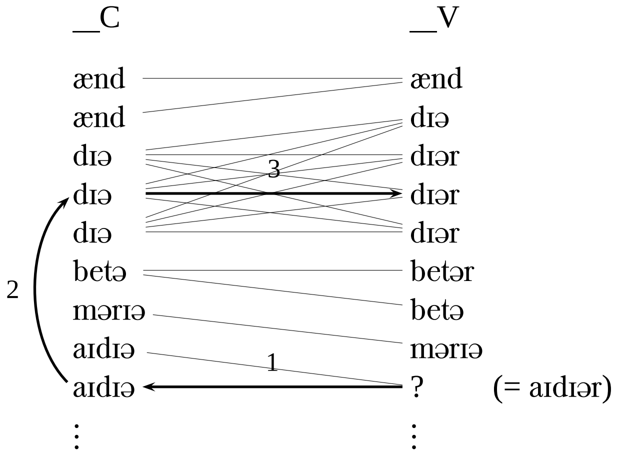


Figure 3: *Constructing a plausible output form for* IDEA *in* __V*; (1): choosing a token of the target in the known environment; (2): choosing an analogical source in the known environment; (3): choosing a token of the analogical source in the given environment.*

sition from the analogical target to the analogical source is determined by three factors: (i) similarity to the target in the known environment, (ii) token frequency and (iii) the availability of at least one form belonging to the same lexeme in the given environment. The third factor is crucial, as the next step consists in randomly choosing another token of the analogical source in the given environment (step 3). After this, the two tokens of the analogical source are compared and their difference is applied to the analogical target in the known environment. The result of this operation is the output of the prediction, which, in this case, is [aɪdɪər].

Although each of these steps is described in more detail in the Appendix, it will be useful to provide a brief outline of step 2. The probability of a form S_i being chosen as the analogical source given the analogical target in the known environment, T_j, is shown below:

(8) $$P(S_i|T_j) = \frac{f_i \eta_{ij}}{\sum\limits_{k \in K} f_k \eta_{kj}},$$

where f_i is the number of tokens the form is exemplified by in the known environment, η is a quantitative measure of similarity, and K is the set of all tokens in the known environment. Since the divisor is constant for all S_i given T_j, the relative probabilities for two different forms are solely determined by their frequency and similarity to T_j. The formal details of the similarity metric η are described in the Appendix.

Note that the model proposed above seems to contradict one of the claims made in the previous section: it uses analogy to predict the pronunciations of all the forms in the data set, although it was claimed that speakers are unlikely to rely on analogy all the time. However, this contradiction is only apparent. The analogical mechanism proposed above does not exclude the possibility of choosing a form of the analogical target itself as the analogical source in the known environment (in Figure 3, this would mean choosing one of the forms representing IDEA in step 2). In these cases, the analogical mechanism essentially reduces to simple lexical access, producing the same results as if one simply sampled the distribution of the target in the given environment. In fact, this happens in the majority of the cases, since forms that are identical to the analogical target are assigned a higher similarity value than any other form, and therefore have a high probability of being chosen as the analogical source. The probability of choosing a non-identical analogical source for a given target is proportionate to its neighbourhood density and inversely proportionate to its frequency (this combination of frequency and neighbourhood density is referred to as EFFECTIVE CONTRAST in Ussishkin & Wedel 2009). Thus, analogical extension is especially likely to take place in low-frequency words, and words surrounded by many lexical neighbours (especially if those are of high frequency). This corresponds well with the observation that lexical access is more difficult in words with lower effective contrast (Ussishkin & Wedel 2009).

4.3 Simulating the emergence of intrusive-r

The input dataset for the simulation was a set of 1 million tokens randomly chosen from the CE18 corpus, each of them stored in the form presented above (e.g. {[aɪdɪə], IDEA, __C}). The transcriptions were modified to reflect a fully rhotic dialect, such as the one spoken in the South of England before the 18th century. To create the conditions for the analogical extensions described above, an additional bias to delete coda-*r* was introduced into the model (the probability of deletion was 0.2 throughout the simulation). The simulation consisted of 50 rounds. The following discussion evaluates the results with respect to the three main empirical observations listed in (5), and briefly addresses a number of further predictions that seem to emerge from the simulation.

Let us first take a look at how well the results of the simulation match the observations in (5a) and (5b), namely that the emergence of intrusive-*r* is gradual and linked to the decline of rhoticity. Figure 4 provides a summary of the changes in the dataset. Each line shows the proportion of rhotic productions in a specific word class in a given environment plotted against the number of iterations. A full line indicates *r*-ful words and a dashed line *r*-less words; black is used to mark words in prevocalic position, dark grey words in preconsonantal position and light grey words in prepausal position. Focussing only on *r*-less forms in prevocalic position (the black dashed line) and *r*-ful forms in preconsonantal/prepausal position (the full grey lines), it seems clear that the model is capable of simulating the analogical extensions that led to the emergence of intrusive-*r*.

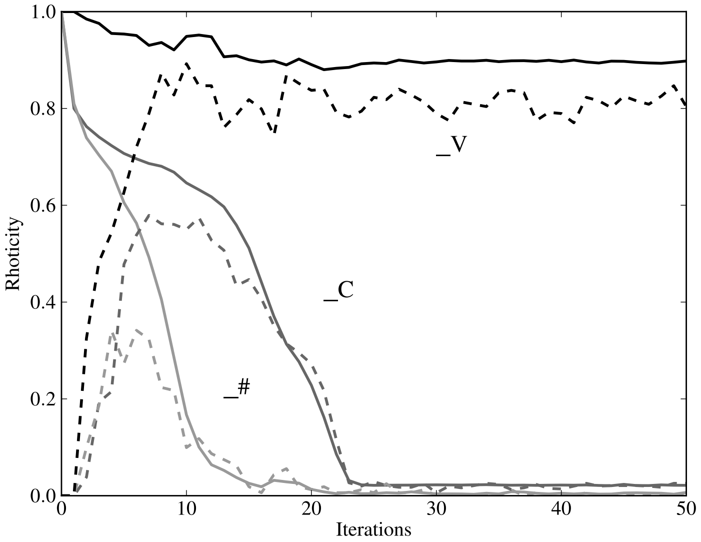


Figure 4: *The loss of rhoticity and the emergence of intrusive-*r*. The full lines indicate* r*-ful forms and the dashed lines* r*-less forms; black stands for prevocalic position, dark grey for preconsonantal position and light grey for prevocalic position.*

The incidence of *r*-ful productions in prevocalic position rises steadily as the degree of rhoticity decreases. Crucially, the model produces variable results in each round, and the proportion of forms with intrusive-*r* increases gradually over the course of the simulation. This is perfectly in line with observation (5a). Moreover, the interaction between rhoticity and intrusive-*r* mentioned in (5b) is present in the simulation as well: there is a negative correlation between the incidence of intrusion and the degree of rhoticity in preconsonantal/prepausal position, which is strong and significant by a parametric correlation ($r = -0.60$, $p < 0.001$). It should also be noted that the extension of the *r*-ful pattern begins well before the model approaches categorical non-rhoticity. This, again, corresponds well with the fact that the dialects in which intrusive-*r* first appeared are unlikely to have been fully non-rhotic.

Before moving on to the third observation in (5), there are a few further trends in the results that deserve some discussion. First, there is a slight decrease in the rhoticity level of *r*-ful words in prevocalic position: the full black line falls to 90 per cent after a few iterations. The reason for this change is as follows. The analogical source for *r*-ful words will sometimes be one that does not show an *r*-zero alternation (i.e. any word that ends in a consonant other than *r* or a vowel other than [ɑː], [ɔː] or schwa). Furthermore, as the level of rhoticity declines, the analogical target in the known environment (i.e. the form that mediates the pattern shown by the analogical source) will become more and more likely to be a form without an *r* when the given form is in prevocalic environment. Therefore, prevocalic tokens of *r*-ful words will sometimes be constructed by applying a non-alternating pattern to a form without *r*, which leads to a certain amount of non-rhoticity in prevocalic position.[4] This is not an unwelcome result, given that the evidence from quantitative studies of *r*-liaison suggests that *r*-ful words are indeed subject to a certain amount of variation in prevocalic position (cf. Table 3).

Second, *r*-less forms in preconsonantal and prepausal position seem to undergo a short phase of partial rhoticity at the beginning of the simulation. This is because the analogical mechanism promotes a merger of *r*-ful and *r*-less forms in all environments if there is a sufficient amount of variation in the data set. Although this tendency might seem somewhat counterintuitive, there is some reason to assume that such ‘hyper-rhotic’ productions were, in fact, present in pre-intrusion dialects. Britton (2007) argues that much of the evidence for intrusive-*r* from 18th century English could also be interpreted as evidence for hyper-rhoticity (given that most authors do not specifically mention prevocalic environment as the only site where non-etymological *r* is found), and he finds hyper-rhotic productions in present-day (partially) rhotic dialects of English as well. I would add to this the informal observation that the variety of Scottish Standard English spoken in Edinburgh also seems to exhibit a certain amount of hyper-rhoticity, and is also usually described as only partially rhotic.[5] Although more research is needed on this phenomenon, the available evidence suggests that the predictions of the model might be borne out by pre-intrusion dialects of English.

In order to evaluate the results of the simulation with respect to the observation about lexical diffusion in (5c), it is necessary to take a more fine-grained look at the class of *r*-less words in prevocalic position. The top panel in Figure 5 shows the relative frequencies of *r*-less words at different levels of rhoticity in prevocalic position plotted against the number of iterations. One way to interpret this graph is to imagine it as a series of histograms lined up in a row shown from a bird's-eye view (with darker colours indicating higher peaks). The distribution of rhoticity values reveals a considerable amount of variation across words: almost none of the distributions are unimodal, and the range of rhoticity levels spans almost a third of all the possible values at any given point in time. The bottom panel takes an even more detailed look at the evolution of intrusive-*r*, showing the changes in the levels of rhoticity for the five most frequent words (*saw*, *idea*, *draw*, *law* and *Clarissa*[6] in order of decreasing frequency). While the words follow roughly similar trajectories, there are some obvious differences in their development, especially in the initial phase of the simulation: certain items change slower than others, which is particularly clear in the case of the word *Clarissa* (indicated by the line that only climbs above 20 per cent after the 15th iteration). The validity of such lexeme-specific predictions is difficult to confirm in the absence of data regarding the likelihood of insertion in specific lexical items. However, the ana-

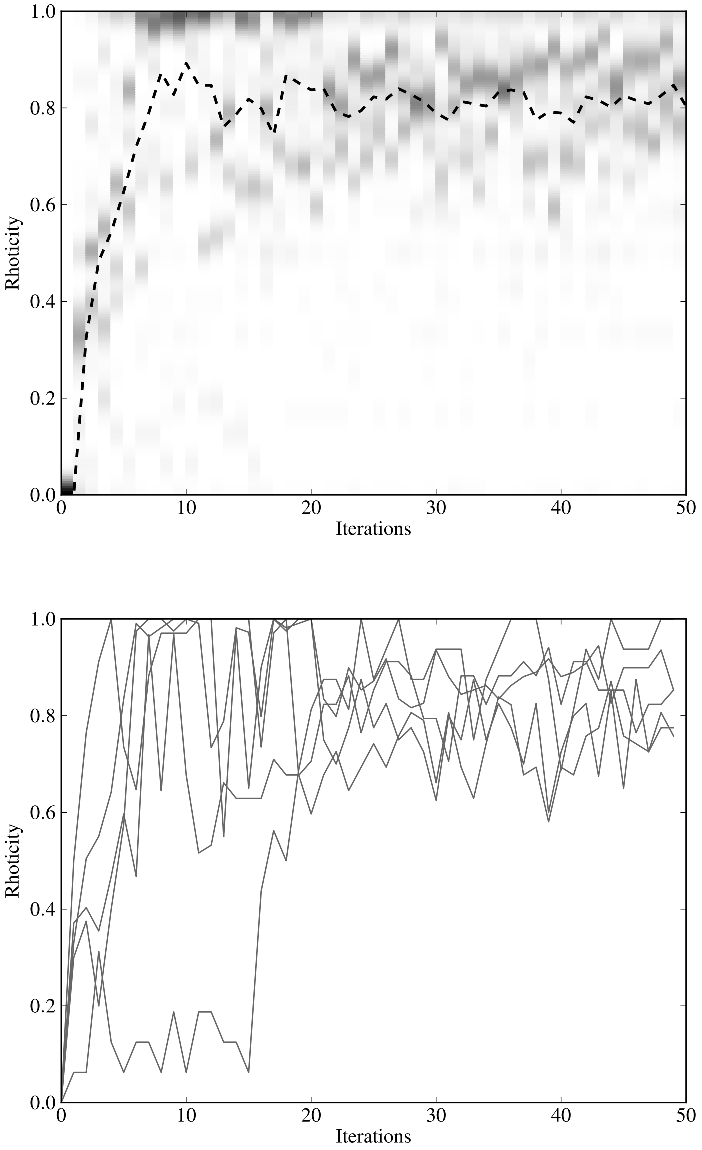


Figure 5: *Top panel: the evolution of the distribution of rhoticity levels in prevocalic* r*-less words; the dashed line shows the mean rhoticity level; Bottom panel: the evolution of intrusive-*r *in the five most frequent forms in the* r*-less group.*

logical model does perform better than previous models of intrusive-*r* inasmuch as it predicts that such tendencies should exist.

Although a detailed discussion of the consistency of the results across different simulations goes beyond the scope of the present paper, it is worth mentioning that the results described above are fairly characteristic of the simulations that have been run. All simulations with similar parameter settings have produced essentially the same results (see the Appendix for a discussion of what these parameter settings are): a gradual increase in the rate of rhotic productions in the *r*-less class in prevocalic position. In other words, the emergence of intrusive-*r* appears practically inevitable in the simulations. This fits in well with the observation that all dialects of English that show *r*-zero alternations in the *r*-ful class seem to have extended them to *r*-ful words as well – there are no dialects with linking-*r* that do not also have intrusive-*r* (cf. Section 2.2).

In sum, the predictions of the token-based approach to analogy match the observations about the history and the variability of intrusive-*r* presented in Sections 2 and 3. Before concluding this section, one final note should be made about the simulation results outlined above. The careful reader will have noticed that the exact rates of rhoticity produced by the simulation do not correspond to any of the dialects described in Section 3. Although this seems to cast doubt on the validity of the present approach, I do not see it as a major argument against token-based analogy as an explanation for intrusive-*r*. The diachronic development and the synchronic behaviour of intrusion is likely to have been influenced by numerous factors, including syntax, prosody, sociolinguistic factors and perhaps even spelling. Therefore, it would be overly optimistic to expect that a simple model relying solely on phonological factors should capture all the subtle patterns of variation related to the phenomenon. The goal of the present account was not to construct a model which fits the data perfectly, but to find a convincing explanation for the emergence of intrusive-*r*. It is all the more surprising to see how closely the predictions of this model correspond to what can be inferred about the development of the phenomenon.

5 Discussion and conclusions

The preceding sections have demonstrated that the emergence of intrusive-*r* is highly compatible with an analogical approach and that a simulation based on this approach can be used to reconstruct the development of the phenomenon in SBE (and possibly other dialects as well). It has also been shown that it is possible to account for certain general patterns of variation in the development of intrusive-*r* in the same analogical framework if the units over which analogy operates are tokens rather than types. The empirical coverage of this approach and the validity of its predictions suggest that the historical source of intrusive-*r* is analogy.

The implications of these results for synchronic analyses of intrusive-*r* are clear: as it is possible to account for both the history and the present behaviour of intrusive-*r* within a diachronic framework, there is no need for an explanation in purely synchronic terms. Of course, this does not mean that we can dispense with synchronic models altogether: we still have to account for speakers' detailed knowledge of their phonology. As a matter of fact, the token-based analogical model described in Section 4.2 makes a number of important assumptions about the nature of this knowledge. Token-based analogy is based on a view of the grammar in which speakers have access to individual instances or exemplars of words and have the ability to make analogical inferences on the basis of these instances. I do not intend to claim that all of phonology boils down to exemplars and analogy; however, it is clear that the present account requires at least these two concepts to be part of the synchronic apparatus of a speaker and can go far in accounting for the phonological facts related to intrusive-*r* without using any additional theoretical machinery.

On a more general note, this account of intrusive-*r* shows that an exemplar-based approach is not necessarily restricted to accounts of phonetic variation, although this is the area where such models have been applied most successfully (cf. Johnson 1997; Pierrehumbert 2001, 2002, 2003; see Wedel 2004, 2007 for other applications of exemplar theory to phonology). By introducing a simple analogical mechanism and re-storing innovative productions, it becomes possible for weak and variable tendencies to give rise to robust patterns. The extensions produced within a single generation in the simulation are sporadic and irregular, but the final pattern is highly systematic. Thus, while variability is an important component of the model, it can also deal with systematic alternations that are traditionally considered part of phonology and have not previously been successfully accounted for in exemplar-based approaches.

APPENDIX

The following paragraphs present a technical description of the token-based analogical model presented in this paper. This essentially consists in specifying the mathematical formulae used in the transitions illustrated in Figure 3 and briefly discussing the effects of different parameter settings.

The transition between the given environment and the known environment (steps 1 and 3 in Figure 3) was described as random in Section 4.2. While this is true, the choice of the output form was not entirely unbiassed. The probability of choosing variant *i* given a specific lexical item and environment was calculated as follows:

$$(9) \qquad P(i) = \frac{f_i^{\gamma}}{\sum\limits_{v \in V} f_v^{\gamma}},$$

where f_i is the number of tokens the variant is exemplified by in the environment, V is the set of all variants associated with the word in the environment and γ is a response-scaling parameter (cf. Nosofsky & Zaki 2002). Parameter γ has an important influence on the behaviour of the model. If $\gamma = 1$, the model shows perfect probability-matching. If $\gamma > 1$, the differences in the probabilities become exaggerated, resulting in increasingly deterministic choices as $\gamma \to \infty$. In the simulation, a value of 1.5 was used, which introduced a weak bias against variation within specific environments. The outcome of the simulation did not change drastically as long as γ was kept in the range $[1,3]$.

The transition from the target to the source in the known environment (step 2 in Figure 3) is slightly more complicated. The general formula for choosing a given form has already been described in (8), although a few details were left unspecified. Below is the formula used for calculating similarity values (based on Nosofsky 1986):

(10) $$\eta_{ij} = e^{-d_{ij}^2},$$

where d_{ij} is a distance measure. The exponential decay function leads to a reduction in the relative influence of forms at a greater distance from the target. The distance measure is specified as follows:

(11) $$d_{ij} = a\sqrt{\sum_{k=1}^{N} w_k |m(x_{ik}, x_{jk})|^2},$$

where N is the number of features used to represent a given form, w_k is the weight assigned to the kth feature, x_{ik} is the value of the kth feature of form i and a is the maximum value of the distance measure (provided that the weights sum to 1). The weights serve to set the relative importance of each sound in choosing the analogical source. The value of a determines the relative importance of similarity vs. frequency: as $a \longrightarrow \infty$, the choice described in (8) becomes entirely dependent on similarity. $m(x_{ik}, x_{jk})$ is defined below:

(12) $$m(x_{ik}, x_{jk}) = \begin{cases} 0 & \text{if} \quad x_{ik} = x_{jk} \\ 1 & \text{if} \quad x_{ik} \neq x_{jk} \end{cases}$$

The parameters were set as follows: the features were the last seven sounds of each form (e.g. $\{=, =,$ a, ɪ, d, ɪ, ə$\}$ for [aɪdɪə]); the last feature (i.e. the last sound of the form) had a higher weight associated with it than the rest of the features; and a was set to 10. It should be noted that the simulation produced similar results when a was varied as long as it remained in the range [10, 30].

References

Albright, Adam. 2009. Modeling analogy as probabilistic grammar. In James P. Blevins & Juliette Blevins (eds.), *Analogy in grammar: Form and acquisition*, 185–213. Oxford: Oxford University Press.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90(2). 119–161.

Antilla, Arto & Young-mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104. 31–56.

Baayen, R. Harald, Richard Piepenbrock & Léon Gulikers. 1995. The CELEX Lexical Database (release 2). University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Baković, Eric. 1999. Deletion, insertion and symmetrical identity. Ms. Harvard University.

Bauer, Laurie. 1984. Linking /r/ in RP: some facts. *Journal of the International Phonetic Association* 14. 74–79.

Bermúdez-Otero, Ricardo. 2005. The history of English intrusive liquids: using the present to ascertain the past. Handout of paper presented to the Department of Linguistics and English Language, University of Manchester, 24 May 2005. URL: www.bermudez-otero.com/intrusion.pdf.

Blevins, Juliette. 1997. Rules in Optimality Theory: Two Case Studies. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 227–260. Oxford: Oxford University Press.

Brighton, Henry. 2003. Simplicity as a driving force in linguistic evolution. Ph.D. dissertation, University of Edinburgh.

Britton, Derek. 2007. A history of hyper-rhoticity in english. *English Language and Linguistics* 11(3). 525–536.

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

Chen, Matthew Y & William S-Y Wang. 1975. Sound change: Actuation and implementation. *Language* 51(2).

Daelemans, Walter, Jakub Zavrel, Ko van der Sloot & Antal van den Bosch. 2007. TiMBL: Tilburg Memory Based Learner, version 6.1, Reference Guide. ILK Research Group Technical Report Series no. 07-07.

Donegan, P. 1993. On the phonetic basis of phonological change. In Charles Jones (ed.), *Historical linguistics: problems and perspectives*, 98–130. London: Longman.

Foulkes, Paul. 1998. English [r]-sandhi: a sociolinguistic perspective. *Leeds Working Papers in Linguistics & Phonetics* 6. 18–39.

Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16. 24–54.

Gick, Bryan. 2002. The American intrusive L. *American Speech* 77. 167–183.

Gimson, Alfred Charles. 1980. *An introduction to the pronunciation of English*. London: Arnold.

Halle, Morris & William J. Idsardi. 1997. r, hypercorrection and the elsewhere condition. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 331–348. Oxford: Oxford University Pres.

Harris, John. 1994. *English sound structure*. Oxford & Cambridge, Mass.: Blackwell.

Hay, Jennifer & Margaret MacLagan. 2010. Social and phonetic conditioners on the frequency and degree of 'intrusive /r/' in New Zealand English. In Dennis Preston & Nancy Niedzielski (eds.), *A reader in sociophonetics*, Berlin: Mouton de Gruyter.

Hay, Jennifer & Andrea Sudbury. 2005. How rhoticity became /r/-sandhi? *Language* 81. 799–823.

Hay, Jennifer & Paul Warren. 2002. Experiments on /r/-intrusion. *Wellington Working Papers in Linguistics* 14. 47–58.

Hock, Hans Henrich. 2003. Analogical change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 441–460. Oxford: Blackwell.

Jespersen, Otto. 1909. *A modern English grammar on historical principles*. London: George Allen & Unwin.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 145–165. San Diego: Academic Press.

Jones, Charles. 1989. *A history of English phonology*. London: Longman.

Jones, Daniel. 1964. *An outline of English phonetics*. Cambridge: Heffer and Sons Ltd 4th edn.

Kahn, Daniel. 1976. *Syllable-based generalisations in english phonology*: Massachusetts Institute of Technology dissertation.

Kiparsky, Paul. 1995. The phonological basis of sound change. In John A. Goldsmith (ed.), *Handbook of phonological theory,* 640–670. Oxford: Blackwell.

Kirby, Simon, Mike Dowman & Thomas L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences* 104(12). 5241–5245.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1994. *Principles of linguistic change. vol. 1: Internal factors*. Oxford: Blackwell.

Lass, Roger. 2000. Phonology and morphology. In Norman Blake (ed.), *The Cambridge history of the English language, Vol. III 1476-1776*, 23–155. Cambridge: Cambridge University Press.

McCarthy, John J. 1991. Synchronic rule inversion. In L. Sutton, C. Johnson & R. Shields (eds.), *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 192–207. Berkeley, CA.: Berkeley Linguistics Society.

McCarthy, John J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38. 169–195.

McDavid, Raven I. 1958. The dialects of American English. In W. Nelson Francis (ed.), *The structure of American English*, New York: The Ronald Press Company.

McMahon, April. 2000. *Lexical phonology and the history of English*. Cambridge, UK: Cambridge University Press.

McMahon, April, Paul Folkes & Laura Tollfree. 1994. Gestural representation and Lexical Phonology. *Phonology* 11. 277–316.

McMahon, April & Paul Foulkes. 1995. Sound change, phonological rules and articulatory phonology. *Belgian Journal of Linguistics* 9. 1–20.

Mompeán-Gonzalez, Jose & Pilar Mompeán-Guillamón. 2009. /r/-liaison in english: An empirical study. *Cognitive Linguistics* 20(4). 733–776.

Nosofsky, Robert M. 1986. Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115. 39–57.

Nosofsky, Robert M. 1988. Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning Memory and Cognition* 14. 54–65.

Nosofsky, Robert M & Safa R Zaki. 2002. Exemplar and prototype models revisited: response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(5). 924–40.

Oudeyer, Pierre-Yves. 2006. *Self-organization in the evolution of speech*. Oxford: Oxford University Press.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Joan L. Bybee & Paul Hopper (eds.), *Frequency effects and the emergence of lexical structure*, 137–157. Amsterdam: John Benjamins.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & N. Warner (eds.), *Laboratory phonology, Vol. VII*, Berlin: Mouton de Gruyter.

Pierrehumbert, Janet B. 2003. Phonetic diversity, statistical learning and acquisition of phonology. *Language and Speech* 46. 115–154.

Sheridan, Thomas. 1762/1803. *A course of lectures on elocution. 2nd American edition*. London: Strahan.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic Publisher.

Skousen, Royal, Deryl Lonsdale & Dilworth B. Parkinson. 2002. *Analogical modeling*. Amsterdam: John Benjamins.

Sóskuthy, Márton. 2009. Why *r*? an alternative look at intrusive-*r* in English. M.A. thesis, Eötvös Loránd University, Budapest.

Uffmann, Christian. 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29. 451–476.

Ussishkin, Adam & Andrew B Wedel. 2009. Lexical access, effective contrast and patterns in the lexicon. In Paul Boersma & Silke Hamann (eds.), *Phonology in perception*, 267–292. Berlin: Mouton de Gruyter.

Vennemann, Theo. 1972. Rule inversion. *Lingua 29* 29. 209–242.

Wedel, Andrew B. 2004. Self-organization and categorical behavior in phonology. Ph.D. dissertation, University of California at Santa Cruz.

Wedel, Andrew B. 2007. Feedback and regularity in the lexicon. *Phonology* 24. 147–185.

Wells, John. 1982. *Accents of English. 3 Volumes*. Cambridge: Cambridge University Press.

NOTES

[1] Some authors prefer the term postvocalic *r*; I will, however, avoid this term as it is misleading and inaccurate: the word-medial *r* in words such as *ferry* and *zero* is also postvocalic, but it is never involved in *r*∼zero alternations.

[2] It should be noted that this finding is foreshadowed in Bermúdez-Otero 2005, where [ə]-final words are claimed to have ‘very low type-frequency’ (ibid. 5).

[3] This is because only types ending in [ə, ɔː, ɑː, ɜː] ever follow an *r*∼zero pattern, which makes the last sound of the word a good predictor of *r*-fulness.

[4] Although this mechanism could also introduce rhotic productions in preconsonantal and prepausal forms, such productions are mostly suppressed by the *r*-deletion bias described at the beginning of 4.3.

[5] As evinced by forms like [aɪdɪər‖] ‘idea’, and [lɛprɪkɔrn] ‘leprechaun’, recorded by the author of this article.

[6] The high frequency of *Clarissa* is due to the inclusion of Samuel Richardson’s eponymous novel in the corpus.