

# /r/-sandhi in the speech of Queen Elizabeth II

Jose A. Mompean 

Department of English, University of Murcia

[mompean@um.es](mailto:mompean@um.es)

This paper looks at the use of /r/-sandhi in the speech of Queen Elizabeth II. Potential contexts of /r/-sandhi were identified and analysed for the presence or absence of rhoticity and glottalisation in a corpus of Christmas speeches over a period of seven decades. The results show that the Queen avoids intrusive /r/ altogether but that she uses linking /r/ in most potential cases, that glottalisation is common when /r/-sandhi is not used, and that linking /r/ and glottalisation can also co-occur. A comparison with a longitudinal corpus of speakers also shows that the Queen resembles group-level trends in the case of linking /r/ but differs in the case of intrusive /r/. The results also indicate that a number of phonetic and usage-based variables influence linking /r/ usage, including previously unreported factors such as vowel quality collocation frequency. The findings are discussed in the context of usage-based theory.

## 1 Introduction

Sandhi is a cover term for a wide variety of processes (e.g. assimilation, elision, etc.) that occur at word boundaries in connected speech. One well-known sandhi phenomenon is liaison – or the use of a word-final consonant not present in the canonical form of a lexical item immediately before vowels. In non-rhotic English varieties of the British Isles (Foulkes & Docherty 1999), the United States (e.g. Halle & Idsardi 1997, Nagy & Irwin 2010), or the Southern Hemisphere (e.g. Hay & Sudbury 2005, Hartmann & Zerbian 2009), among others, an r-sound may occur in heterosyllabic  $V_1[-\text{high}] + V_2$  contexts (e.g. *here it is* [ $^h\text{h}\text{e}\text{r}\text{i}\text{t}\ \text{i}\text{z}$ ]; cf. *here* [ $\text{h}\text{i}\text{r}\text{e}$ ]). This type of /r/-sandhi is commonly referred to as LINKING /r/ and coexists with another type, INTRUSIVE /r/ (e.g. *idea of justice* [ $\text{a}\text{i}\text{d}\text{i}\text{a}\text{v}\ \text{d}\text{z}\text{a}\text{s}\text{t}\text{i}\text{s}$ ]; cf. *idea* [ $\text{a}\text{i}\text{d}\text{i}\text{e}$ ]). The main difference between the two types is whether or not an r-sound was pronounced at an earlier stage of the language and whether the conventional spelling still contains an <r>. However, both linking /r/ and intrusive /r/ are sometimes considered to be the same linguistic phenomenon from a synchronic point of view given their similar distribution patterns, phonological context, and hiatus-breaking function (McCarthy 1993). They are, therefore, collectively referred to as /r/-liaison or /r/-sandhi.

The current paper describes Queen Elizabeth II's /r/-sandhi usage over a period of 67 years, compares that use with a reference corpus, and then assesses the influence of phonetic, linguistic, and usage-based variables in its variability. There are various reasons for such a description. One motivation is that this speaker provides a large, stylistically consistent, life-long longitudinal corpus. Other relevant and more theoretical motivations are discussed in Section 3. The specific research questions addressed are the following:

RQ1: How does Queen Elizabeth II resolve hiatus at potential /r/-sandhi sites?

RQ2: Is her use of /r/-sandhi in line with speech community trends?

RQ3: What factors condition /r/-sandhi usage in her speech?

The paper first reviews some variables that can potentially condition variation in /r/-sandhi. It then describes the use of /r/-sandhi by the Queen. Finally, it discusses some implications of the findings obtained that are relevant not only for the speaker and the variety under investigation but for other non-rhotic varieties of English at large where /r/-sandhi is used.

## 2 /r/-sandhi in English

In what follows, the focus is mostly on data-driven studies and the types of variables influencing variability in the use of /r/-sandhi. These studies include both corpus-based and lab-based experimental studies, although the first type predominates. They also focus on various non-rhotic varieties. The focus is on phonological, lexical/grammatical, and usage-based factors, which are relevant for the current study. Sociolinguistic factors (e.g. gender, ethnicity, socioeconomic status, accentual variety, style, etc.) were controlled for in this study, so they are not discussed at this point, although potentially relevant studies are mentioned in the discussion.

### 2.1 Phonological factors

/r/-sandhi seems conditioned by both prosodic and segmental phonological factors. Regarding prosody, /r/-sandhi is believed to be more frequent within than across tone units with no intervening pause (Joseph 1999). Moreover, linking /r/ seems to be disfavoured when the second hiatus vowel is stressed (e.g. Foulkes 1997, Hannisdal 2006, Hay & Maclagan 2010, Mompean & Gómez 2011, Cox et al. 2014, Yuen, Cox & Demuth 2015, Pavlík 2016), which coincides with a left-dominant foot boundary in English (Giegerich 1999).

Concerning segmental factors, one potential predictor is the quality of the vowels flanking the /r/-sandhi site. Some studies describe no effect of the preceding vowel in both linking /r/ and intrusive /r/ (e.g. Mompean & Mompean-Guillamón 2009, Pavlík 2016). However, other studies suggest that at least some back vowels favour /r/-sandhi. Hay & Sudbury (2005) showed that intrusive /r/ was more common in 19<sup>th</sup>-century New Zealand English (NZE) after /ɑ:/ than after schwa and least of all after /ɔ:/. In contrast, Foulkes (1997) for Derby English, Hay & Warren (2002) for contemporary NZE, and Hannisdal (2006) for Received Pronunciation (RP) – or Standard Southern British English (SSBE) – found intrusive /r/ to be more common after /ɔ:/ and /ɑ:/ than after schwa. Finally, the influence of the following vowel has also been considered. Hay & Sudbury (2005) found that back vowels favoured linking /r/. Pavlík (2016) reported that linking /r/ and intrusive /r/ were more common in RP before mid central or close front vowels.

Two further segmental factors potentially affecting variation in /r/-sandhi usage are the presence of another /r/ in the vicinity (before, after or at both sides) of the potential /r/-sandhi site and the length of syllables flanking that site. In this respect, Hannisdal (2006) found that the presence of another /r/ in the vicinity clearly inhibited /r/-sandhi in RP, although no such effect was described in other studies on RP (e.g. Mompean & Mompean-Guillamón 2009, Pavlík 2016) or NZE (Hay & Sudbury 2005). Finally, Hay & Sudbury (2005) found no influence of syllable length in NZE, but Navarro (2013) noted that monosyllabic words triggered higher rates of linking /r/ than polysyllabic words for the syllable before the /r/-sandhi site in a corpus of Lancashire English.

## 2.2 Lexical and grammatical factors

A few studies have looked at lexical factors that might influence variability in /r/-sandhi. Lewis (1975) and Hannisdal (2006), for example, found linking /r/ to be less frequent after honorifics (e.g. *Dr.*, *Mr.*, *Sir*, etc.). Most studies, however, focus on grammatical features. According to McCarthy (1993), /r/-sandhi production should not be particularly sensitive to syntactic constituency. However, /r/-sandhi might be less frequent at clause boundaries because such boundaries mark the division of two separate syntactic planning units (Foulkes 1997). In this respect, Hannisdal (2006) noted that, although the presence of a clause boundary did not prevent the insertion of /r/ significantly, the frequency of /r/ was somewhat lower at clause boundaries than in other environments. She found, in addition, that speakers with lower linking /r/ rates tended to avoid /r/ also before syntactic boundaries, but found no such effect for speakers with higher linking /r/ rates.

Another grammatical variable potentially influencing /r/-sandhi production is the word class (lexical vs. function) of the items flanking the /r/-sandhi site. Hay & Sudbury (2005) found no evidence of this effect. However, Bauer (1984), Foulkes (1997) and Viollain (2014) observed that linking /r/ was more common after monosyllabic function words, and a recent study by Pavlík (2016) also found a small effect. Finally, Hannisdal (2006) determined that a combination of two lexical words disfavoured the use of linking /r/, and found higher frequencies between function words. One possible explanation for this is that function words are normally unstressed and carry little informational value, facilitating linking /r/ (Hannisdal 2006).

/r/-sandhi has also been claimed to be sensitive to morphological structure, such as the type of morpheme boundary (word-external vs. word-internal). As a sandhi phenomenon, /r/-sandhi typically occurs across word boundaries but it can also be found at word-internal morpheme boundaries in affixed words (e.g. *feve*[ɹ]*ish*, *draw*[ɹ]*ing*) and compounds (e.g. *four-yea*[ɹ]*-old*, *salsa*[ɹ]*-evening*). Linking /r/ before suffixes is considered to be categorical (i.e. obligatorily present) and often excluded from data analyses in studies of /r/-sandhi. Intrusive /r/ in affixed words and both linking /r/ and intrusive /r/ in compounds show some degree of variability. However, a few studies suggest that /r/-sandhi becomes lexicalised in word-internal positions. Mompean & Mompean-Guillamón (2009) found that intrusive /r/ was more common word-internally in the noun *withdrawal* than across word boundaries. Similarly, Hannisdal (2006) found linking /r/ to be more common within compounds than across word boundaries. Finally, Hay & Warren (2002) and Hay & Maclagan (2010) found that the rate of intrusive /r/ is a partial function of suffix identity, with intrusive /r/ being increasingly more common after nominal suffixes (e.g. *-ism*), adjectival suffixes (e.g. *-ish*) and verbal suffixes (e.g. *-ing*).

## 2.3 Usage-based factors

The structural factors mentioned above may correlate with – and sometimes be explained by – usage-based factors such as lexical, morphological and vowel type frequency as well as collocability and collocability frequency. The general finding is that higher frequencies favour higher rates of /r/-sandhi (see Hay & Maclagan 2012 for an exemplar-based explanation). Hay & Sudbury (2005), for example, found that a high lexical frequency of the word after the /r/-sandhi site disfavoured the use of linking /r/, and Hay & Maclagan (2012) found that a high lexical frequency of the previous word favoured it. Pavlík (2016), however, found word frequency to have no influence. In contrast, he found that the frequency of /r/-sandhi was predicted by the frequency of the vowel type. For example, /r/ was favoured in sites followed by mid central and close front vowels, the most frequent /rV/ sequences in connected speech. Finally, Hay & Warren (2002) found that affixes with high type frequencies (e.g. *-ing*) were more likely to attract intrusive /r/ than less frequent ones (e.g. *-ism*).

Regarding collocability, linking /r/ seems more common in frequent collocations than in free combinations (Hay & Sudbury 2005, Hannisdal 2006, Mompean & Mompean-Guillamón 2009, Navarro 2013, Sóskuthy 2013, Cox et al. 2014), which suggests that

common phrases containing highly predictable sequences of words are stored and accessed as a whole. Finally, Pavlík (2016) suggests that, although collocability has an effect on /r/-sandhi frequency, it is the frequency of the collocation that has the largest effect.

One interesting finding in some of the studies above is the interaction between frequency of use and degree of constriction used for /r/ as measured by *F3*. The third formant is the most salient characteristic of /r/ (Lisker 1957), often leading to a strong percept of rhoticity (Lawson, Stuart-Smith & Scobbie 2018). Yet English /r/ displays gradient features in its realisation (see Section 4.2.2 below for more details). *F3*, for example, can be higher or lower, with lower values resulting in strong robust /r/. As far as /r/-sandhi is concerned, there is evidence that speakers who use intrusive /r/ (Hay & Maclagan 2010) or linking /r/ (Hay & Maclagan 2012, Cox et al. 2014) more often have stronger, more robust /r/s. The effect also seems to emerge in potential linking /r/ cases with words that tend to occur more often before vowels. These are more likely to elicit an r-sound in linking /r/ contexts than words which tend to occur before consonants (Hay & Maclagan 2012). Finally, lower *F3* values are also more common in frequent collocations (Cox et al. 2014).

### 3 /r/-sandhi in the speech of Queen Elizabeth II

The research reviewed above reveals a picture of /r/-sandhi as a very complex phenomenon shaped by numerous factors. Yet our understanding of the phenomenon is far from complete for various reasons. One motivation for the current study is that, despite the literature reviewed above, there is limited evidence for some factors. Two examples suffice to illustrate the point. Firstly, lexical frequencies of the words at potential /r/-sandhi sites, lexical collocational frequencies, and the frequency of the vowels flanking those sites have been investigated. Yet, vowel quality collocational frequency – or the frequency of the combination of linking and linked vowels – has not yet been assessed; and secondly, the effect of tone unit boundaries does not seem to have been explored. Since pauses prevent a potential case of /r/-sandhi, studies rule out such cases and, consequently, many tone unit boundaries (see also Section 4.2.1 below). Yet prosodic boundaries can occur without intervening pauses. It has been suggested, in this respect, that /r/-sandhi could be more frequent within than across tone units with no intervening pause (Joseph 1999), but this does not seem to have been explored empirically. Finally, it should be pointed out that /r/-sandhi has been studied in various non-rhotic accents of English but, other than their non-rhoticity, those accents are far from homogenous so findings regarding a particular factor may differ across varieties or social groups, as the research reviewed above suggests.

The current study also takes a longitudinal approach to the phenomenon of /r/-sandhi. In this respect, there is some anecdotal (e.g. Lewis 1975) or production data evidence pointing to sound change in /r/-sandhi in Queen Elizabeth II's accentual variety (i.e. RP). In an analysis of recordings made between 1949 and 1966, Bauer (1984) found that younger RP speakers (born after 1940) tended to use intrusive /r/ more than older speakers. As for linking /r/ in RP, there is also some anecdotal evidence pointing to a decline in its use in the course of the twentieth century (Jones 1956, Allerton 2000). Other than anecdotal evidence, almost all studies trying to assess possible changes over time in the use of /r/-sandhi in RP (e.g. Bauer 1984) or other non-rhotic varieties (e.g. Foulkes 1997, Hay & Sudbury 2005, Piercy 2012, Cox et al. 2014) approach the phenomenon from an APPARENT-TIME perspective. These studies collect data from different age groups at one point in time under the assumption that there is little or no change in a person's accent after reaching adulthood. Yet potential diachronic change can also be studied from a REAL-TIME perspective in at least two main ways: (a) by collecting data from comparatively similar groups at different periods (e.g. years or decades); and/or (b) by collecting data longitudinally. Finally, real- and apparent-time approaches need not be mutually exclusive. Used in combination, they can provide enormous insights into language change, although both approaches have strengths and weaknesses (Tillery & Bailey 2003).

Finally, most studies are conducted at the level of the (social) group or community and focus on group-level trends, but studies on idiolectal speech are rare. In the case of /r/-sandhi, individual data are often masked by group-level trends (e.g. males vs. females, older speakers vs. younger speakers, etc.). Moreover, when data are available for specific speakers, these are often limited. In Hannisdal's (2006) study of /r/-sandhi in contemporary TV newsreaders, she provided data on potential linking /r/ data from 15 male and 15 female speakers. These ranged from 162 to 275 cases but they seem limited, however, to explore the numerous variables (e.g. phonological, lexical and grammatical, usage-based, etc.) that seem to condition linking /r/. Studies in which various hundreds of (or a few thousand) potential cases of linking /r/ do exist, but similar numbers are never analysed for specific individuals.

It is necessary at this point to justify why the focus of this study is on idiolectal speech as opposed to the speech of a (large) group of similar speakers (i.e. a sociolect), as is common in other studies. It is well known that linguistic research has progressively shifted from a focus on an idealised monolingual speaker–hearer in the generative tradition to a focus on a socially and culturally situated language user that is a member of a social group, as in much research in traditional sociolinguistics. Yet, much of this second tradition and related approaches are based on the assumption that patterns of variation are homogeneous across individual speakers in a group as well as across groups of speakers. A consequence of this view is that inter-speaker and intra-speaker variability are often factored out, discarding individual profiles, particularly when they deviate from tendencies, as anomalous outliers. Yet, such individualities do have an impact not only on immediate actual language use but also, in the long term, on linguistic change (Stevens & Harrington 2014). Moreover, current THIRD-WAVE approaches to the study of sociolinguistic variation view speakers 'not as passive and stable carriers of dialect, but as stylistic agents, tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation' (Eckert 2012: 97–98). Finally, the grammar of a given language or community variety can be considered as a composite of the individual grammars of its members (Ferguson 1979) and, although individual systems may overlap to a substantial extent, adult monolingual native speakers of the same language do not share exactly the same linguistic knowledge (or mental grammar) – contrary to widespread belief (Dąbrowska 2012). One reason for this is that language users vary in the quality and amount of input they receive throughout their lives.

Given the considerations above, the study of an individual's idiolectal use of /r/-sandhi – or of any speech phenomenon – can be considered worthy of attention for both levels of linguistic analysis (individual and community) to be reconciled in the scientific attempt to understand language as used by socially affiliated individuals that also have idiosyncratic traits, personal histories, roles and agency. Since there is also evidence that individuals' linguistic behaviour changes throughout their lifespan (e.g. Sankoff & Blondeau 2007, Bowie 2015), such type of studies can also shed light not only on whether individuals are in line with or differ from group-level trends but also on whether, in cases of sound change in progress, individuals adjust to those changes. Moreover, they can also shed light on whether certain variables considered to have an effect on linguistic phenomena also have the same effect at the idiolectal level, and whether the contribution of multiple variables to the variation in a given phenomenon as is often gleaned from group-level studies also surfaces at the individual level.

The current study looks at the use of /r/-sandhi in the speech of Queen Elizabeth II (henceforth QEII) over a period of seven decades. As mentioned in Section 1, the choice of this particular speaker was mainly motivated by the existence of a large, stylistically consistent, life-long longitudinal dataset and her unique sociolinguistic status. QEII was also the focus of attention in a series of studies by Harrington and colleagues (e.g. Harrington, Palethorpe & Watson 2000, Harrington 2006, 2007), who found that sound changes undergone by RP such as final prevocalic /ɪ/-tensing, /u/-fronting and /æ/-lowering had also taken place over a fifty-year period in QEII's speech. Under the hypothesis that further changes might have taken place in QEII's speech over the years, a recent preliminary longitudinal

study looked at her use of /r/-sandhi (Mompean 2019). In that study, QEII's use of linking /r/ – in the corpus of Christmas broadcasts analysed in the current study – was compared with the use of linking /r/ in another corpus to determine whether the use of linking /r/ had significantly declined over the decades. The group-level corpus comprised a larger time span (1927–2017) and data from 72 speakers. The results suggested some decrease in the use of linking /r/ across time in both QEII and the group-level corpus. The study is revisited in Section 5 and new evidence is provided there.

The aims of the current study are to provide a more detailed account of QEII's use of /r/-sandhi, to determine to what extent her use is characteristic of the larger linguistic community to which she belongs, and to find out what phonological, linguistic and usage-based variables determine the variation in her use of /r/-sandhi. In doing this, evidence can also be gained – as pointed out above – regarding whether the contribution of multiple variables to /r/-sandhi that the literature has identified through group-level studies also surfaces at the individual level.

The variables that can be studied are numerous and an analysis of them all is beyond the limits of the current study. However, an attempt was made to study a representative number of phonological, grammatical and usage-based variables that apply to the segments, syllables or lexical items that precede and follow potential /r/-sandhi sites, from now on referred to as linking or linked, respectively. These include variables that have not been explored before with others for which there is some evidence. The phonetic variables studied are (a) presence/absence of a tone unit (TU) boundary; (b) rhythmic stress in linking and linked syllables; (c) vowel quality of linking and linked phonemes; and (d) presence/absence of neighbouring /r/ before and after the /r/-sandhi site. The grammatical variables are: (a) presence/absence of a syntactic boundary; and (b) word class of linking and linked words. The usage-based variables are: (a) lexical frequency of linking and linked words; (b) lexical collocation (linking + linked word) frequency; (c) vowel quality frequency of linking and linked vowels; and (d) vowel quality collocation (linking + linked vowel) frequency. Sociolinguistic stylistic variables beyond age cannot be studied given that other variables such as gender, ethnicity, geographical origin, socioeconomic status, and even style remain constant in the corpus.

Given the research questions of this study, the hypotheses entertained were the following:

RQ1:

- i. Linking /r/ will be used more frequently than intrusive /r/.
- ii. Glottalisation will be used more frequently than hiatus when /r/-sandhi is not used.

These hypotheses are based on the fact that all previous studies have shown that linking /r/ is more frequent than intrusive /r/ (e.g. Hannisdal 2006, Mompean & Gómez 2011, Pavlík 2016) and that glottalisation is typical when /r/-sandhi does not apply, as suggested by some anecdotal (e.g. Allerton 2000), production (e.g. Bauer 1984), and corpus-based (e.g. Mompean & Gomez 2011) evidence.

RQ2:

- i. QEII's use of /r/-sandhi will be in line with overall group-level trends in the use of both linking /r/ and intrusive /r/.

This is explored from a longitudinal (real-time) as well as a synchronic (apparent-time) perspective. This hypothesis is based on the fact that QEII has already been shown to follow group-level patterns, as pointed out in the Harrington and colleagues' studies of vowel changes in her speech over the years. This change is also noticeable more impressionistically in her speech, which seems to have shifted from a more aristocratic form of RP towards a more mainstream variety (Robson 2016).

RQ3:

- i. Phonological variables (tone unit boundary, stress, neighbouring /r/, vowel quality) and grammatical variables (syntactic boundaries and word class) will condition

/r/-sandhi usage, but phonological variables will have a greater effect on /r/-sandhi than grammatical variables.

- ii. Usage-based variables (lexical/vowel quality and collocation frequency) will condition /r/-sandhi.

These hypotheses are based on findings and suggestions in the relevant literature, although the specific contribution of each variable is less clear to predict. It was hypothesised, however, that phonological variables would have a greater effect on /r/-sandhi than grammatical variables such as syntactic boundaries or word class, for which there seems to be little evidence (see Section 2.2. above). Finally, usage-based variables were predicted to have a strong effect on /r/-sandhi. Usage-based theory (e.g. Bybee 2001, 2006, von Mengden & Coussé 2014) predicts various effects of frequency of use such as the routinisation and automatism of productions, phonetic reduction effects, decay of previously established patterns, etc. More specifically, and concerning /r/-sandhi, high-frequency sequences tend to favour the entrenchment of linguistic features. Given this, it can be expected that high-frequency patterns will lead to higher rates of /r/-sandhi, pointing to an entrenchment of /r/ in such high-frequency contexts.

## 4 Method

### 4.1 Materials

A corpus containing the annual Christmas broadcasts produced by QEII was compiled for this study (henceforth QXB corpus). The Christmas message is a traditional broadcast made by the British sovereign to the people of the Commonwealth of Nations each Christmas since 1932, when King George V delivered the first of such messages (to the then British Empire). Each message serves the same communicative purpose, that is, an annual season's greetings to the public in which, unlike other official speeches, the monarch does not deal with government policy but offers instead her views on personal milestones and the year's major events. Despite its personal character, the Christmas message has, in fact, a great political importance in the relationship between the monarch – or more generally the monarchy – and the country since it is the only occasion in which she can convey a message of her own without needing the Prime Minister's consent (Catsiapis 2005). The messages have been televised since 1957 and recorded in advance since 1960. The location is usually Buckingham Palace.

At the time of the analysis, the corpus spanned a period of 67 years (December 1952–December 2018) and contained 66 broadcasts, with an annual broadcast except for 1969, when the Christmas message was replaced with a documentary on the Royal Family. Table 1 shows the years ( $n = 63$ ) for which the full or part of the broadcast could be accessed and analysed. This corpus expands substantially the dataset analysed in some studies by Harrington (2006, 2007) and Hinton (2015) on potential diachronic change in QEII's speech, which included only 15 and 13 broadcasts, respectively. The corpus analysed amounts to over six hours of speech and 38,800 words. The audio material was obtained from *YouTube* and *AP Archive* (<http://www.aparchive.com/>). The transcripts were obtained from the official website of the British monarchy (<https://www.royal.uk/history-christmas-broadcast>). The quality was generally considered acceptable for subsequent perceptual and acoustic analyses. Broadcasts of lesser acoustic quality underwent a noise reduction process using *Audacity* (version 2.1.0; Audacity Team 2012), a noise profile filtering which was set at a 12 dB reduction, 6 for sensitivity, and 3 bands of frequency smoothing.

The corpus is stylistically consistent, as mentioned above, given QEII's unchanging role and communicative purpose as well as the broadcasts' format and style (i.e. formal). Thus, there are no shifts in role (she always addresses the audience as the Queen) and the speech sample is genuine (often not the case of actors/actresses). Moreover, the corpus only contains

**Table 1** QEII's corpus of Christmas broadcasts (QXB). Decades 1950s to 2010s and years. Availability for the study: NA (non-available); white cell (no broadcast); <sup>(p)</sup> partial broadcast.

50s			52	53	54	55	56	57	58	59
60s	60	61	NA	63	64	65	66	67 <sup>(p)</sup>	68	
70s	70	71	72	73	74	75	76 <sup>(p)</sup>	77 <sup>(p)</sup>	NA	NA
80s	80 <sup>(p)</sup>	81 <sup>(p)</sup>	82 <sup>(p)</sup>	83	84	85	86	87	88	89
90s	90	91	92	93	94	95	96	97	98	99
00s	00	01	02	03 <sup>(p)</sup>	04	05	06	07	08	09
10s	10	11	12	13	14	15	16	17	18	

adult non-pathological data. The Queen was sufficiently mature (26 years) at the time of the first message to rule out changes in her voice potentially affecting /r/-sandhi; and although aging has an impact on speakers' general voice quality, we do not expect it to have an impact on structural patterns of /r/-sandhi production. Her accent can be described as an upper-crust or MARKED variety of RP as opposed to a more mainstream, neutral or UNMARKED variety (see Wells 1982, Upton 2008, Cruttenden 2014 for discussions on varieties of RP). There are indications, however, that the Queen's accent seems to have shifted from a more aristocratic form of RP towards a more mainstream variety characteristic of the traditional professional middle classes in England (Harrington et al. 2000, Robson 2016).

In order to address RQ2, QEII's production was compared with a reference corpus to find out whether her use of /r/-sandhi was in line with group-level patterns across time. As mentioned above, the group-level data are from a study of 72 RP speakers (36M; 36F), aged 25–55, and a corpus of around 58,000 words that had been compiled from some of their public speeches (Mompean 2019). The speakers were public figures or newsreaders, their date of birth ranged from 1879 to 1990, and the data were produced from 1927 to 2017. They were grouped around the concept of demographic generation/cohort. This can be defined as a group of people born within a given period of time and who experience the same significant social events. The cohorts identified were WW1 (1879–1900), WW2 (1901–1925), Silent (1926–1945), Baby boomers (1946–1960), X (1961–1980), and Millennials (1981–2000), with each cohort comprising six males and six females. The procedure for the identification and analysis of potential /r/-sandhi sites in this corpus was the one described in Section 4.2 below. Regarding linking /r/, the number of potential cases per speaker was 10, so the corpus amounted to 720 potential cases. The number of potential intrusive /r/ cases in the whole corpus was 57.

## 4.2 Procedure

### 4.2.1 Identification of potential /r/-sandhi sites

As a preliminary step, potential spelling-based word-external /r/-sandhi sites were identified with the Find feature in a word processor and checked manually for their suitability. For linking /r/, the searches in the corpus found all cases of <r> in the coda of syllables at the end of words followed by a vowel. These searches identified typical linking /r/ spellings such as <-r/-re> (e.g. *our*, *here*, etc.). Less usual spellings such as <rrh> (e.g. *catarrh*) and even numbers containing potential linking /r/ in their pronunciation (e.g. 4) were also identified. Regarding intrusive /r/, the searches found all vowel-final words before vowel-initial words. These searches identified typical intrusive /r/ spellings ending in <-a> or <-aw> (e.g. *India*, *saw*, etc.) but also marginal ones such as <eu> (e.g. *milieu*). Finally, searches also considered vowel + silent letter combinations, often found in proper nouns such as loanwords of Semitic, Indo-Iranian or Celtic origin ending in <-h, -gh> (e.g. *Abdullah*, *Omagh*), French loanwords ending in <-s> (e.g. *bourgeois*) or Old English place names ending in <-urgh> (e.g. *Edinburgh*).

Next, potential spelling-based /r/-sandhi contexts containing a pause were ruled out, as pauses prevent a potential case of /r/-sandhi. This was the case for contexts where a breath pause was made or where some sort of hesitation on the part of the speaker was perceived. Yet, following Pavlík (2016), a pause was considered to exist in all these cases if characterised by 100 ms of silence or more, as determined acoustically with PRAAT (<http://www.fon.hum.uva.nl/praat/>). Following this process, 241 potential spelling-based cases of linking /r/ and 36 of intrusive /r/ containing a pause were ruled out. The final number of retained cases without such pauses was 760 for linking /r/ and 53 for intrusive /r/ (QXB corpus). As mentioned above, the reference corpus yielded similar figures (720 potential cases of linking /r/ and 57 of intrusive /r/).

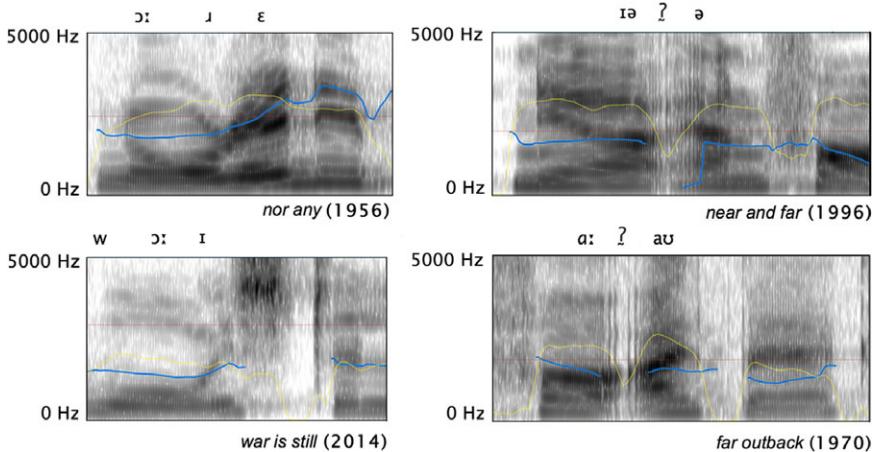
It should be borne in mind that this study is concerned with word-external /r/-sandhi. However, identification of potential cases of /r/-sandhi in word-internal position was also carried out for affixed and compound words, since /r/-sandhi can also happen in that context. In the current study, there were only three potential linking /r/ sites after prefixes (*unde*[ɹ]*estimate*(d), *ove*[ɹ]-*anxious*). In contrast, linking /r/ before suffixes is very common. There were also 34 potential cases of linking /r/ in compounds, including open-form (e.g. *Fa*[ɹ] *East*), closed-form (e.g. *whe*[ɹ]*ever*) and hyphenated (e.g. *fá*[ɹ]-*off*) compounds. As for intrusive /r/, there were no potential cases after a prefix and only one potential site (*draw*[ɹ]*ings*) before a suffix. Moreover, three compounds involving potential intrusive /r/ were identified (*awe*[ɹ]-*inspiring*, *law*[ɹ]-*abiding*, *Teresa*[ɹ] *of Calcutta*). Given the scarce data on /r/-sandhi within word boundaries as well as the impossibility of coding those potential cases for several of the variables, these were excluded from the statistical analyses. Nevertheless, Section 5 contains some discussion of these cases.

#### 4.2.2 Analysis and coding of /r/-sandhi sites

As with most phonetic phenomena, /r/-liaison realisations display gradient features and variability. To start with, there are different realisations of /r/-sandhi depending on the accent and, to some extent, on the historical period under study. The labiodental approximant [v], often stigmatised by prescriptivists, seems to be gaining ground as a feature of some British accents, particularly southeastern England (Foulkes & Docherty 2002). In RP, [v] has been described as ‘an upper-class affectation’ (Wells 1982: 282). A voiced alveolar tap [ɾ] (or tapped /r/) seems to have been common up until the 1950s in RP (Fabricius 2017). Since then, and currently, the most typical realisation in RP is a postalveolar approximant [ɹ] and this is the only realisation found in QEII’s speech. Moreover, there are differences between /r/-sandhi and canonical /r/ in terms of tongue tip height and duration (see Gick 1999, Tuinman, Mitterer & Cutler 2011, cf. Mullooly 2004). Finally, there seems to be variation in *F3* values across /r/ productions, with lower *F3* leading to stronger percepts of rhoticity (Lawson et al. 2018).

Despite this variation, most studies treat /r/-sandhi as a binary variable that opposes the presence of constricted /r/ to no (or vocalised) /r/. Following this approach, a categorical decision was adopted as to the presence or absence of /r/-sandhi. For this, every instance of an /r/-liaison site was inspected acoustically with PRAAT by the author, a long-term observer of the variety under investigation (e.g. Mompean & Hernández-Campoy 2001). Following Pavlík (2016), the criteria used were: lowered *F2* and *F3*, a regular structure of voicing pulses, and relatively stable pitch and intensity contours.

Regarding glottalisation, it should be pointed out that the phenomenon occurs on a continuum from laryngealisation to full glottal closure, including intermediate stages such as creaky-voiced glottal approximants (i.e. [ʔ] or [ʔ̥]), and it may be difficult even for phoneticians to discriminate glottal stops from creaky voice. In fact, the occurrence of true glottal stops in English may be unusual as so-called glottal stops in English are often realised as creaky voice (Ashby & Przedlacka 2014). To avoid these problems, and following previous studies (e.g. Cox et al. 2014), varying degrees of glottalisation were not differentiated in this



**Figure 1** (Colour online) Examples of QEII's productions: linking /r/ in *nor any* (above left); hiatus in *war is still* (below left); glottalisation in *near and far* (above right) and *far outback* (below right). Blue solid lines represent pitch while yellow solid lines represent intensity. Glottalisation is represented with [ʔ] irrespective of its degree.

study. In any case, typical acoustic cues for glottal stops (e.g. stop gap, absence of  $f_0$ , sudden energy change), creaky voice (lowered  $f_0$ , decreased acoustic intensity, irregularly spaced glottal pulses with periods of varying length), or a combination of these (e.g. creaky voice + glottal stop) were considered to determine the presence of glottalisation. Finally, cases involving no glottalisation were treated as hiatus, manifesting itself as two adjacent vowels without a clearly audible transition ( $V\#V > VV$ ) or gemination (e.g. ( $V\#V > V:$ ), often the case of rapid realisations of sequences of two schwas. Figure 1 shows examples of linking /r/, hiatus and glottalisation from the QXB corpus.

Coding reliability was checked by an independent phonetician on two corpus subsets. The inter-rater agreement was high, as measured by Cohen's kappa (linking /r/: 0.9; intrusive /r/: 0.95). Although the coding was based on visual inspection of spectrograms and  $f_0$  contours – and their related values – most potential cases categorised as /r/ were perceived as an r-sound, as /r/-sandhi tends to be a perceptually noticeable phenomenon.

### 4.2.3 Coding of variables

In the rest of this section, a description of how each variable was coded is presented.

#### *Phonological variables*

- (a) *Tone unit boundaries*. Common acoustic criteria were used to determine the presence or absence of a boundary (see Wagner & Watson 2010 for a review). These include duration and, crucially,  $f_0$  changes. As for duration, TU boundaries are often accompanied by final (or pre-boundary) lengthening. More specifically, syllable durations are longer at the end of a TU than in the middle of it, while syllables seem to be lengthened at the beginning of TUs. As for  $f_0$ , pitch tends to decline across the course of a TU and reset to a higher value after the TU boundary. As mentioned in Section 4.2.1 above, the presence of a pause (i.e. 100 ms of silence or more) rules out a potential /r/-sandhi context, so only TUs including no pause were considered.
- (b) *Rhythmic stress*. This was coded as a binary option: stressed/strong (S) vs. unstressed/weak (W), in both the linking and linked syllables. This yields four possible combinations across the /r/-sandhi site: (a) stressed–unstressed (SW), unstressed–stressed (WS), stressed–stressed (SS), and unstressed–unstressed (WW). The WS and

SS contexts represent the foot boundary (Giegerich 1999). The focus in this study was on whether stress in one of the two syllables (linking or linked) has an effect on /r/-sandhi production irrespective of the stress pattern of the other syllable involved at the /r/-sandhi site.

- (c) *Vowel quality*. This was determined given the perceived height and backness of the linking and linked vowels. Since /r/-sandhi sites can be flanked by many different linking and linked vowels, individual vowel qualities were not considered as such. Instead, vowels were coded for the linking vowel based on backness and, following Pavlík (2016), a combination of height and backness. This yielded five categories for the linking vowel. Based on the backness criterion, vowels were classified as central vs. back. The backness category did not include a front distinction since linking vowels cannot be front vowels. In the case of height + backness, three categories were considered: open back (/ɑ:/), close-mid back (/ɔ:/), and mid central (monophthongs /ə ɜ:/; centring diphthongs /ɪə eə ʊə/, and triphthongs /əʊə aʊə aɪə eɪə ɔɪə/).

For the linked vowel, grouping vowels into categories was considered even more appropriate since almost every vowel in the RP inventory can occur word-initially. Pavlík (2016) tested various groupings and arrived at a six-category scheme. This scheme was based on its high explanatory power – as measured by statistically significant differences between the groups in terms of rates of linking /r/ usage as opposed to absence of /r/. However, the scheme also seems to group vowels around various areas in the vowel space based on the height and backness of the vowel at its onset. Pavlík's categories were: close front (/i i: i ɪə/), mid front (/e eɪ(ə) eə/), open front (/æ/), mid central (/ə ɜ: əʊ(ə)/), open central/back (/ʌ ɑ: aʊ(ə) aɪ(ə)/), and close-mid/open back (/ɒ ɔ: ɔɪ(ə)/). Pavlík excluded an additional category of close back vowels (/ʊ u: ʊə/) given their extremely low frequency, which prevented it from being tested alongside the other categories. In the current study, a five-category scheme was used following the categories established by Pavlík for the sake of cross-study comparison. However, the mid front and open front categories were merged, given their phonetic similarity and the infrequent number of linked syllables starting with an open front vowel in the QXB corpus. Thus, the five categories used were: close front (/i i: i ɪə/), mid/open front (/e eɪ(ə) eə æ/), mid central (/ə ɜ: əʊ(ə)/), open central/back (/ʌ ɑ: aʊ(ə) aɪ(ə)/), and close-mid/open back (/ɒ ɔ: ɔɪ(ə)/). It should be pointed out that there are phonetic similarities between vowels across different subgroupings. For example, /ɑ:/ in the open central/back category has some similarities with /ɒ/ in the close-mid/open back category. One difference between the two categories, however, is a certain degree of lip rounding at the vowel's onset.

It is also worth mentioning that previous analyses (e.g. Piercy 2012) determined the quality of the vowel by referring to canonical forms of the words, or did not specify whether canonical or actual qualities were used (e.g. Cox et al. 2014, Pavlík 2016). In this study, vowel qualities of the forms as actually pronounced in their respective contexts were used, which coincided mostly with their canonical ones in the case of lexical words. Special attention was given, however, to the vowel quality of function words whose backness can change from central to back (e.g. *our* /a(u)ə/ → /ɑ:/) or, more commonly, from back to central (e.g. *for* /fɔ:/ → /fə/, *are* /ɑ:/ → /ə/) given vowel reduction.

- (d) *Neighbouring /r/*. The presence of /r/ was checked in the linking and linked syllables as well as in the syllable immediately preceding the linking syllable and in that following the linked syllable. In the case of the linking syllable, the presence of /r/ was checked in the onset of the linking syllable (e.g. *p[ɹ]aye[ɹ] even*), but also in the onset of the previous syllable (e.g. *children*, in *child[ɹ]en a[ɹ]e enjoying*). Concerning the linked syllable, the presence of /r/ was checked in the coda of that syllable – in cases of a second concatenate linking /r/ (e.g. *are*, in *the[ɹ]e a[ɹ]e also*) – and also in the onset of the following syllable (e.g. *real*, in *the[ɹ]e is [ɹ]eal hope*).

### Grammatical variables

- (e) *Syntactic boundaries*. These were considered independently of prosodic ones, given that, although both types of boundaries often coincide, there is no one-to-one correspondence between them. Given that the cases of major (or sentential) syntactic boundaries (e.g. subject/verb predicate, main/subordinate clause) were few, minor (or phrasal) boundaries such as those between object nouns, prepositional phrases, and adverbial phrases were also taken into account. Since minor syntactic boundaries can also have prosodic effects, it was assumed that they could also have an impact on /r/-sandhi usage. Use was made of the *LinGO English Resource Grammar* (Flickinger 2000) in cases of doubt as to the syntactic structure of the context in which the potential /r/-sandhi sites were embedded.
- (f) *Word class*. The lexical vs. function word status of items flanking the /r/-sandhi site was determined given the specific contextual function of each item. As a case in point *are* can function as a lexical item (main verb) or as an auxiliary (function word) and *there* can be a lexical item (adverb) or the existential pronoun as in *there is/are* (function word). In any case, most nouns, verbs, adjectives and adverbs were classified as lexical words, while prepositions, pronouns, auxiliary verbs, conjunctions and determiners were considered as function words.

### Usage-based variables.

- (g) *Lexical frequency*. This was calculated for both the linking and linked word types using the British National Corpus (BNC), which consists of around 100 million words (Davies 2004). The lexical frequency of words was also calculated for the QXB corpus.
- (h) *Lexical collocation frequency*. In this study, the collocation frequency of bigrams (2-grams) including the linking and linked words (e.g. [*our own*]) was measured in the BNC and the QXB corpus. Use was made, respectively, of Fletcher's (2007) BNC-based interface *Phrases in English* and the online concordance *WebCorp* (Renouf, Kehoe & Banerjee 2007). Trigrams (3-grams) or larger n-grams were not analysed, given that relevant bigrams were already included in them (e.g. [*sooner or*] *later*).
- (i) *Vowel quality frequency*. We measured the frequency of each vowel quality category described above at /r/-sandhi sites, namely open back, close-mid back and mid central for the linking vowel (/Vr/) as well as close front, mid/open front, mid central, open central/back and close-mid/open back for the linked vowel (/rV/). Pavlík (2016) compared the frequencies of /rV/ and /Vr/ at /r/-liaison sites in his data against the frequency of /Vr/ and /rV/ sequences in single-word lexemes in a manually coded mini-corpus of written English of approximately 15,000 words. In the current study, no external corpus was used, given the difficulty in obtaining a comparable spoken corpus. General /VrV/ (e.g. *carry*) or /rV/ (e.g. *red*) sequences in the QXB corpus were not counted alongside those at /r/-sandhi sites.
- (j) *Vowel quality collocation frequency*. The collocation frequency of the vowel quality (height + backness) categories described above was also calculated, which amounted to 15 different combinations.

It should be pointed out that frequency is coded differently across studies. Some studies use a binary criterion (common vs. rare/less common) based on a fixed number of occurrences (e.g. Hay & Maclagan 2010). Other studies transform raw word frequencies (i.e. raw number of occurrences) into some other measure (e.g. Piercy 2012). In this study, and for lexical/vowel quality token and collocation frequencies, the median (or middle value) was used to split the data into two: data above the midpoint of the distribution and data below that midpoint. Items above the midpoint were then coded as having a higher frequency (higher than the middle value) items while those below the midpoint were coded as lower-frequency items. The median seems to be a more appropriate measure than the mean (or average) for a strongly skewed distribution. In the current study, all distributions were strongly right-skewed, so the

median seems to provide a better idea of a typical value and, in the current case, of the typical lexical frequency of a given item or collocation.

#### 4.2.4 Statistical analyses

The results presented below required the use of statistical techniques in some cases. In the case of RQ2, these included linear regression, Generalised additive modelling (GAM), and statistical inference by confidence intervals. Linear regression is a basic and commonly used type of predictive analysis used to explain the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). GAM was used to model potential change within the reference group. GAM does not require aggregation or the pre-selection of a fixed time point. Instead, the method is able to identify general patterns over dynamically varying data, while simultaneously accounting for subject and item-related variability (Wieling 2018). Finally, statistical inference by confidence intervals (CIs) was also used to compare QEII with the data in Hannisdal (2006). A CI provides a range of population values with which a sample statistic is consistent at a given level of confidence (usually 95%).

As for RQ3, an attempt was made to find a suitable algorithm for regression and classification. Logistic regression and random forest are two very common and widely used machine learning methods, which have been found to perform similarly across datasets, with performance more influenced by choice of dataset rather than model selection (Kirasich, Smith & Sadler 2018). Initially, both a logistic regression and a random forest followed by a logistic regression were carried out. The results reported below correspond to the initial logistic regression since it had a lower Akaike information criterion (AIC) than the logistic regression preceded by the random forest (525.8 vs. 552.91). AIC is an estimate of the information lost when a given model is used to represent the process that generates the data. Thus, when choosing between two models, a model with less AIC is often preferred. Incidentally, the initial logistic regression not only found more factors to be statistically significant, but also some of those factors had been previously unreported in the literature (e.g. vowel collocation frequency).

The results of the logistic regression are presented in Section 5.3 below. They reveal that 11 variables (out of 28) had a  $p$ -value  $\leq .1$  including eight with a  $p$ -value  $\leq .05$ . The logistic regression after a random forest only identified five variables with a  $p$ -value  $\leq .05$ . Four of these variables were also identified by the logistic regression without the random forest. The 11 variables were: TU boundary; rhythmic stress (linking and linked syllables); backness (linking vowel); neighbouring /r/ (pre- and post-linking site); word class (linking word); lexical frequency of linked word: BNC and QXB corpora; lexical collocation frequency (BNC corpus); and vowel quality collocation frequency (QXB corpus).

Taking these 11 variables into account, and excluding all others, the R package *glmulti* (Calcagno 2013) was used to obtain the best set of variables considering the AIC. Out of 8,450 models tested, the model with the lowest AIC (525.8) was chosen as the best set of variables. To evaluate model fit (or how well the logistic regression model fits the observed data), several assessments were carried out, including a chi-square goodness-of-fit test of the deviance ( $p = 0$ ) as well as a calculation of several coefficients of determination, namely Hosmer-Lemeshow ( $R^2 = .38$ ), Cox & Snell ( $R^2 = .34$ ) and Nagelkerke ( $R^2 = .51$ ). The goodness-of-fit test reveals that the model is better at predicting linking /r/ usage than random. A high  $R^2$  approaching 80% (.8), 90% (.9), or 100% (1) is required for studies in the natural sciences – given that the behaviour of molecules and/or particles, for example, can be reasonably predicted to some degree of accuracy. However,  $R^2$  values as low as 10% (0.1) are often found in social and behavioural research and, in any case, lower than 50% (Hashim et al. 2016). This is because human behaviour is more difficult to predict. Given this, the  $R^2$  values obtained are not particularly low, which indicates good overall model fit. Based on the previous model of logistic regression, we then calculated the odds ratio (OR) together with

their CIs at 95% to quantify the strength of the association between the predictor and the response variables.

## 5 Results

### 5.1 RQ1: How does QEII resolve hiatus at potential /r/-sandhi sites?

RQ1 addressed how QEII resolved hiatus at word boundaries where a potential /r/-sandhi site is found. In the corpus under study, intrusive /r/ was not used at all in any of the 53 potential cases. The limited dataset and the zero occurrences prevent us from carrying out an in-depth analysis of the phenomenon for the variables studied in RQ3. Yet one interesting finding regarding these cases is QEII's avoidance of intrusive /r/ altogether. She did not use it in any of the four word-internal positions described in Section 4.2.1, within affixed words (*draw*[ɹ]ings) or within compounds (*awe*[ɹ]-inspiring, *law*[ɹ]-abiding, *Teresa*[ɹ] of *Calcutta*).

**Table 2** Rates of presence/absence of linking /r/ in potential cases ( $n = 760$ ) with and without concomitant glottalisation (glott.) or hiatus.

	+linking		-linking		
	%	$n$	%	$n$	
-glott.	72.63	552	+glott.	21.3	162
+glott.	3.42	26	Hiatus	2.6	20
All	76.05	578	All	23.9	182

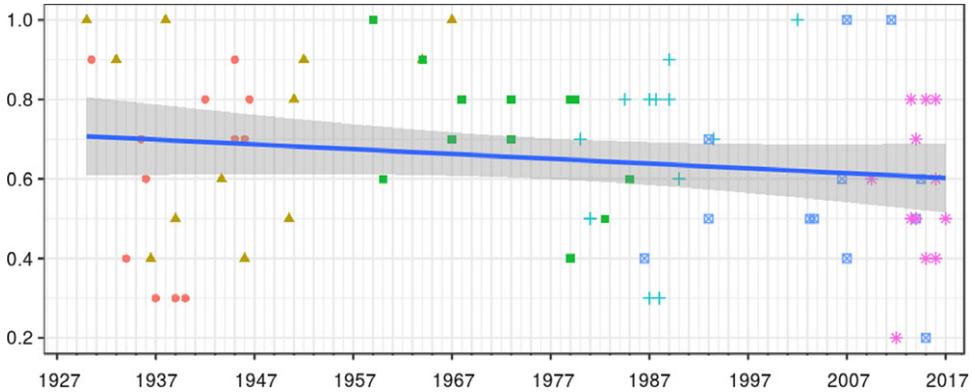
Table 2 summarises the results for linking /r/, which show that out of 760 cases it was used 76.1% of the time. Glottalisation was the most common strategy when /r/-sandhi was not used (21.3%) and pure hiatus was rare (2.6%). Thus, these data allow us to confirm the two hypotheses entertained, that is, that linking /r/ would be more frequent than intrusive /r/ and that glottalisation would be typical when /r/-sandhi does not apply. One interesting result from the current study is the use of both linking /r/ and glottalisation in a few cases (3.4%). This finding shows that, although linking /r/ typically occurs without glottalisation on the one hand, and glottalisation without linking /r/ on the other, both phenomena can co-occur at potential /r/-sandhi sites.

The results reveal that potential word-external cases of linking /r/ are non-categorical. However, at word-internal sites (see Section 4.2.1.), linking /r/ is categorical. An inspection of 100 random cases confirmed this when a potential linking /r/ site was followed by a suffix, as well as in the three cases in which it was preceded by a prefix. As for compounds, it was found in 29 out of the 34 cases (84.3%). Since three tokens were instances of *Siste*[ɹ] *Ethel*, and these contain an honorific (*Sister*), a semantic category claimed to prevent /r/-sandhi (Lewis 1975), linking /r/ can be considered nearly categorical in compounds excluding those three cases, which leaves the rate of word-internal linking /r/ at 93.6%.

### 5.2 RQ2: Is QEII's use of /r/-sandhi in line with speech community trends?

This question was first analysed from a longitudinal perspective. As pointed out in Mompean (2019), an adjusted (multiple) linear regression (ALR) showed that the variable year of production (YoP) was not a good predictor of linking /r/ ( $R^2 = .21$ ) in the group-level corpus, although some decrease over time could be observed (Figure 2).

The same decrease can be observed applying GAM modelling to the data. In this type of model, the rate of linking /r/ production depends linearly on at least two smooth functions of some predictor variables, with each variable being added individually. In this study, the



**Figure 2** (Colour online) Graph of a linear regression line: linking /r/ in a longitudinal corpus of RP speakers (rate per speaker). The line (trend line) represents the fitted values and the shaded bar represents the 95% CI. The small dots are actual data points. Each shape represents an individual in a given cohort: dots (WW1), triangles (WW2), squares (Silent), crosses (Baby boomers), barred squares (X), and asterisks (Millennials).

**Table 3** Analysis of two GAM models applied (linking /r/ production) in a longitudinal corpus of RP speakers.

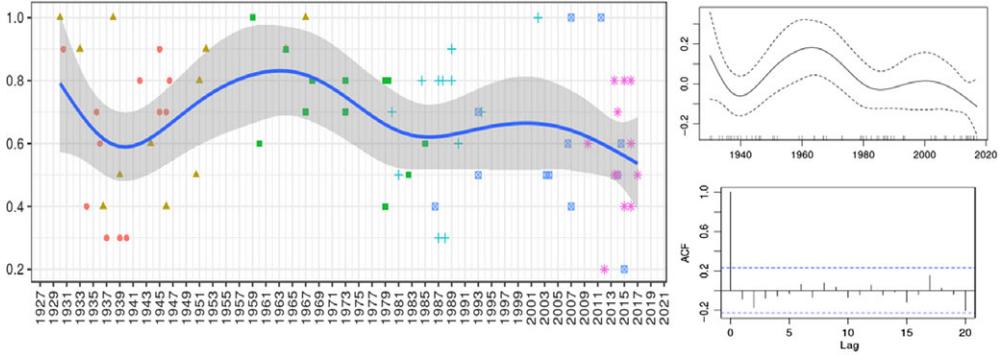
GAM models	Residual df	Residual deviance	df	Deviance	F	Pr(>F)
% linking /r/	71.000	3.4200				
% linking /r/ ~ s(YoP)	64.024	2.7667	6.9757	0.65334	2.2056	0.04539 *
% linking /r/ ~ s(YoP) + s(YoB)	66.165	2.9397	-2.1405	-0.17307	1.9040	0.15453
% linking /r/	71.000	3.4200				
% linking /r/ ~ s(YoB)	67.444	2.9843	3.5557	0.43568	2.7900	0.03878 *
% linking /r/ ~ s(YoB) + s(YoP)	66.165	2.9397	1.2794	0.04459	0.7936	0.40518

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

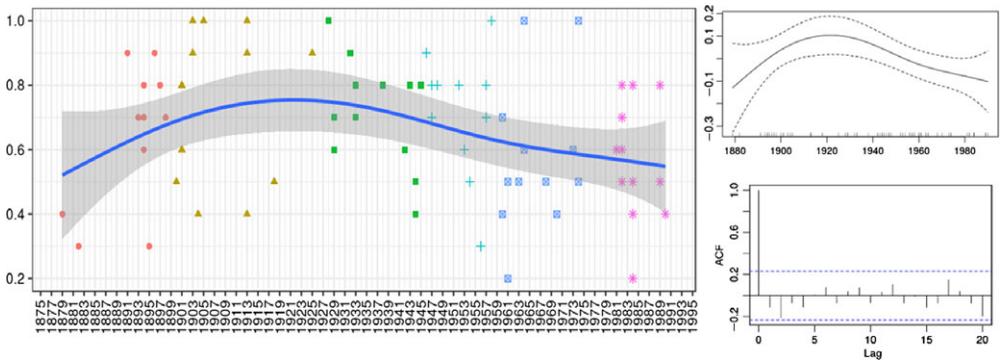
predictor variables were year of production (YoP) mean and year of birth (YoB) mean. These two variables were explored because they are often taken into account in real-time (i.e. YoP) and apparent-time (i.e. YoB) studies. Four GAM models were fitted using a smoothing spline for each of the covariates. The deviance was used to measure the difference between models.

The results are shown in Table 3 as well as in Figure 3a and Figure 3b, which contain three plots each: best fit, spline, and residual autocorrelation function (ACF). Figure 3a relates to YoP while Figure 3b relates to YoB. The figures show no autocorrelation in the residuals for a lag of twenty years, as indicated by the fact that the lines at the different lags are within the area circumscribed by the dashed lines. In the first model, the variable YoP was added first, which produced a significant model fit ( $p = .045$ ) and an undulating, non-linear decrease in the use of linking /r/ across time. Notice, however, that the wiggle around the late 1990s/early 2000s might be due to the lack of data in that part of the curve, so the actual decrease could be even clearer if that part of the curve were not taken into account. Moreover, adding YoB did not result in a significant improvement in model fit ( $p = .154$ ). Similarly, after applying a second GAM model and adding YoB first, a significant model fit was obtained ( $p = .039$ ), again with some non-linear decrease in the use of linking /r/ across time; yet adding YoP the model did not result in a significant improvement in model fit ( $p = .41$ ). The graphs reveal a decreasing trend in the use of linking /r/ over the years. Yet, since the data have high variability, the confidence intervals for the curves are wide, which suggests that further analyses should be conducted with a larger sample size.

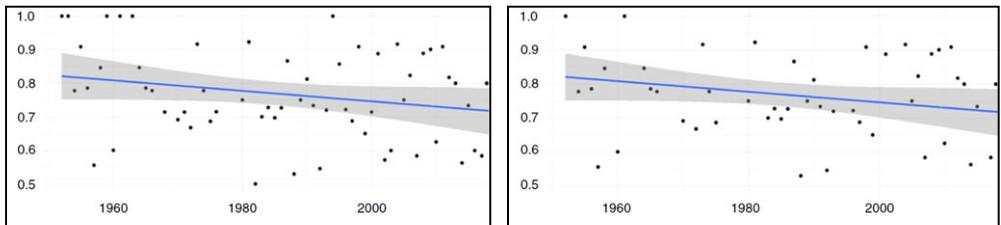
As for QEII's longitudinal pattern of linking /r/ usage, a decrease over the years was also observed. However, a simple linear regression revealed that YoP was not a good predictor of



**Figure 3a** (Colour online) GAM modelling using YoP. Left: GAM plot displaying the best fit (line), 95% CI (shaded bar), and actual data points (dots, as in Figure 2) for YoP. Right (above): spline curve of the relationship between linking /r/ production and YoP as well as the 95% CI (dashed lines). Right (below): residual ACF plot for a lag of twenty years.

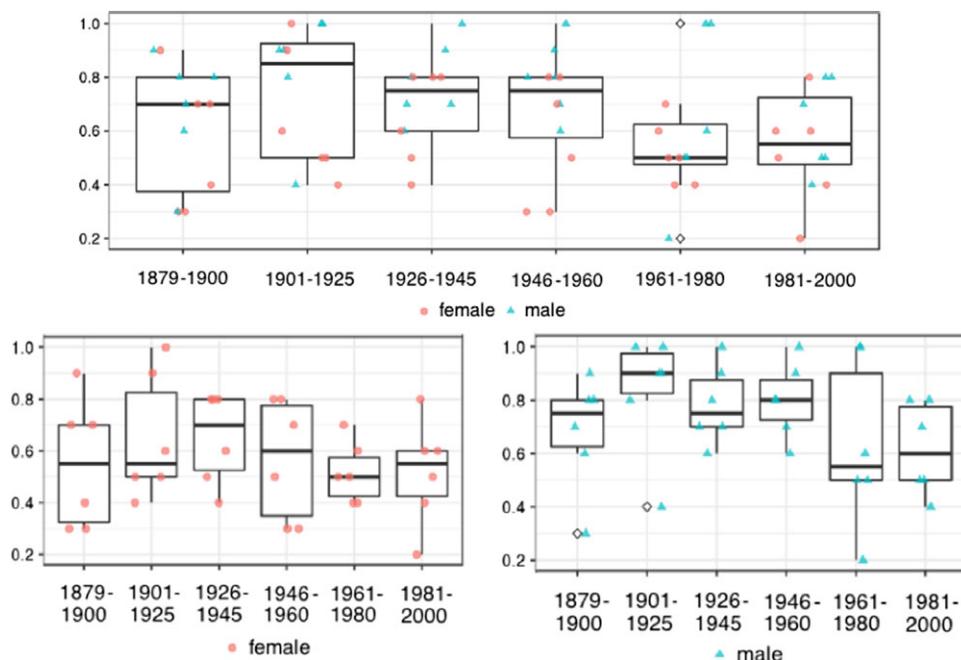


**Figure 3b** (Colour online) GAM modelling using YoB. Left: GAM plot displaying the best fit (line), 95% CI (shaded bar), and actual data points (dots, as in Figure 2) for YoB. Right (above): spline curve of the relationship between linking /r/ production and YoB as well as the 95% CI (dashed lines). Right (below): residual ACF plot for a lag of twenty years.



**Figure 4** (Colour online) Graphs of the rate of linking /r/ per year of production: all years with potential linking /r/ cases (left) and years only with eight or more cases (right).

linking /r/ ( $R^2 = .05$ ), given that  $R^2$  values close to zero suggest that the model explains none of the variability. Since the range of cases per year in the QXB corpus is 2–33, and this may obscure any effect of YoP, a second linear regression was carried out excluding the years with fewer than eight (first quartile) potential cases (Figure 4). The results reveal that YoP is still not a predictor of linking /r/ production ( $R^2 = .01$ ).



**Figure 5** (Colour online) Mean rate of linking /r/ in cohorts including all cohort members (top) or by gender (bottom).

A general comparison between the group-level and QEII's data suggests that both appear to drift downward – i.e. use linking /r/ less over the years (see Figure 5). A closer inspection of QEII's data with those of her cohort and gender group shows that QEII also conforms to their patterns. Kruskal–Wallis tests reveal no statistically significant differences in the mean rate of production across cohorts, including all cohort members or speakers analysed by gender, although females seem to produce less linking /r/ than males.

Given the data above, QEII can be said to belong to the Silent generation, for which demographers and researchers normally use mid-to-late 1920s and early-to-mid 1940s as starting and ending birth years, respectively. As Figure 5 reveals, speakers born in the Silent generation produced linking /r/ at a mean rate of around 75%, while the females of that generation produced it with a mean of above 70%. With a mean rate of 76%, QEII would then be considered to also fall within the normal range of linking /r/ for her generational and gender group.

Regarding intrusive /r/, the data are more difficult to interpret given the limited data from both the general corpus, its cohorts and the QXB corpus, with only 57 and 53 potential cases, respectively. They are also difficult to interpret because the potential cases in the general corpus are mostly limited to one or two cases per speaker and the data in both corpora are also irregularly spread along the time span. Yet the data from the general corpus are clear in showing that intrusive /r/ – produced at a rate of 21% (12 out of 57 cases) – was found in all cohorts except WWI (see Table 4). It is present, for example, in QEII's generational cohort, with four actual cases out of 15 potential ones. With 53 potential cases but no single occurrence of intrusive /r/ in her Christmas messages, QEII clearly differs from what appears to be a general trend in the group-level corpus from the WW2 generation on, that is, a low but constant rate of intrusive /r/ usage.

Another way of assessing whether QEII's use of /r/-sandhi is in line with speech community trends is to compare her data with those of a population at a given point in time. Hannisdal (2006) analyzed the speech of 15 male and 15 female middle-aged RP newscasters in samples produced between 2002 and 2004 (Hannisdal 2006: 246–247 for scores).

**Table 4** Potential and actual occurrences of intrusive /r/ in a longitudinal corpus of RP speakers by generational cohort: WW1, WW2, Silent, Baby Boomers (BB), X, and Millennials (Mill.). Potential/actual cases.

	WW1	WW2	Silent	BB	X	Mill.
M	5/0	6/1	6/3	5/2	6/1	4/0
F	4/0	2/0	9/1	3/1	3/2	4/1

**Table 5** Analysis of CI in a population of 30 RP speakers from Hannisdal (2006) and QEII.

		Point estimate	95% CI		$\chi^2$	df	p-value
			Lower	Upper			
Linking /r/	All	0.598	0.585	0.610	229.560	1	2.2e-16
	M	0.628	0.611	0.646	195.120	1	2.2e-16
	F	0.568	0.550	0.585	56.242	1	6.408e-14
Intrusive /r/	All	0.326	0.288	0.367	66.754	1	3.075e-16
	M	0.320	0.267	0.379	35.587	1	2.439e-09
	F	0.332	0.278	0.391	30.556	1	3.244e-08

Potential cases per speaker ranged from 162 to 275 for linking /r/ ( $M = 201.5$ ) and from 8 to 43 for intrusive /r/ ( $M = 18.6$ ). This dataset was used alongside QEII's scores (as subject 31) to compute the 95% confidence interval for a proportion in one sample with continuity correction. The data from this analysis are shown in Table 5. The data show, for example, that the point estimate – or single value which serves as a best guess or best estimate of actual intrusive /r/ – was .326 (or 33%), and with 95% confidence the true estimate lies between .288 and .367 (or 28.8% and 36.7%). QEII, who produced 76.1% and 0% of linking /r/ and intrusive /r/, respectively, is not within the 95% confidence interval, a range of values that you can be 95% certain contains the mean of the population. Given this, QEII's means cannot be considered to lie within the means of the population.

### 5.3 RQ3: What factors condition /r/-sandhi usage in QEII's speech?

RQ3 addressed the nature of the factors potentially conditioning /r/-sandhi usage in the QXB corpus. Table 6 reports the data from the final (best-fitted) logistic regression model as well as the odds ratio.

The OR data allow for the statement that, for example, the likelihood that linking /r/ will be produced when Stressed linking  $\sigma = 1$  increases by a factor of 3.6 as opposed to when Stressed linking  $\sigma = 0$ . In terms of statistical significance, the results of the logistic regression show that the following variables influence the probability of success or likelihood of occurrence of linking /r/ (i.e. they are predictors):

Phonetic variables regarding:

- Tone units*: QEII is less likely to produce linking /r/ when there is a TU boundary.
- Stress*: QEII is more likely to produce linking /r/ when the linking syllable is stressed and she is less likely to produce linking /r/ when the linked syllable is stressed.
- Vowel quality*: QEII is more likely to produce linking /r/ when the vowel quality of the linking syllable is open back (/ɑ:/).

**Table 6** Best-fitted logistic regression analysis and odds ratio for predictors in the QXB corpus.

Predictor	Estimate	SE	z value	Pr(> z )	OR	95% CI		
						2.5 %	97.5 %	
(Intercept)	2.0659	0.3701	5.581	2.39e-08	***	7.892	3.913	16.75
TU boundary	-4.0353	0.4331	-9.317	< 2e-16	***	0.01768	0.007246	0.03985
Stressed linking $\sigma$	1.2804	0.3578	3.579	0.000346	***	3.598	1.822	7.455
Stressed linked $\sigma$	-2.5662	0.3028	-8.474	< 2e-16	***	0.07683	0.04151	0.1365
VQ: backness (linking $\sigma$ )	1.4947	0.3495	4.277	1.90e-05	***	4.458	2.281	9.012
/r/ before linking site	-0.9455	0.3627	-2.607	0.009136	**	0.3885	0.191	0.7952
/r/ after linking site	-0.5081	0.3384	-1.502	0.133196		0.6016	0.3108	1.176
Word class (linking $\omega$ )	-0.4949	0.2843	-1.741	0.081730		0.6096	0.3474	1.062
Lexical <i>f</i> (linked $\omega$ , BNC)	0.7633	0.4841	1.577	0.114908		2.145	0.8138	5.46
Lexical <i>f</i> (linked $\omega$ , QEII)	-1.2777	0.4881	-2.617	0.008861	**	0.2787	0.1083	0.7372
Lexical colloc. <i>f</i> (BNC)	0.6931	0.2676	2.590	0.009601	**	2	1.187	3.397
VQ colloc. <i>f</i> (QEII)	1.2368	0.2776	4.455	8.41e-06	***	3.444	2.008	5.973

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

(d) *Neighbouring /r/*: QEII is less likely to produce linking /r/ when the linking syllable or the previous one contains /r/.

Usage-based variables regarding:

(e) *Lexical token frequency*: QEII is less likely to produce linking /r/ when the lexical frequency of the linked word is high (QXB).

(f) *Lexical collocation frequency*: QEII is more likely to produce linking /r/ when lexical collocation frequency (linking + linked words) is high (BNC).

(g) *Vowel quality collocation frequency*: QEII is more likely to produce linking /r/ when vowel quality collocation frequency (linking + linked vowels) is high (QXB).

The results obtained indicate that phonological and usage-based variables condition, to various degrees, linking /r/ production. They reveal, however, that the grammatical variables studied have no influence.

## 6 Discussion

### 6.1 QEII's /r/-sandhi resolution and individual vs. group-level trends

The fact that QEII uses a high proportion of linking /r/ is unsurprising given that linking /r/ is supposed to be the rule rather than the exception at potential /r/-sandhi sites. In contrast, although some use of intrusive /r/ was expected to occur in the corpus, even if at very low rates, the Queen's lack of intrusive /r/ is interesting, although perhaps not too surprising as the latter has traditionally been stigmatised in RP.

The current study, as well as the author's previous exploration, suggest some decrease in the use of linking /r/ over the years in the speech of both QEII and a reference group. Statistically, the effect is not very strong, so the question of whether /r/-sandhi is in a state of ongoing sound change or a case of stable variation needs further investigation. As a matter of fact, 'there is lack of consensus on what it means for a variable to be "stable" over time . . . on how to measure and predict stability vs. change' (Gardiner & Nagy 2017: 78–79), and it may be difficult to ascertain whether change is extremely slow or does not occur. Moreover, change may not occur in linear, progressive increments (Dannenberg 2000), as often assumed in apparent-time studies. Figure 3 above, for example, shows a non-linear decrease in the use of linking /r/ over the years and not a steady, linear decrease. Rates seem to rise, for

**Table 7** Rates of linking /r/ in potential cases of linking /r/ and intrusive /r/ in three recent studies.

	Hannisdal (2006)		Mompean & Gómez (2011)		Pavlík (2016)	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Linking /r/	59.8	6045	66.5	1476	62	647
Intrusive /r/	32.6	558	38.1	399	21	613

example, between 1940 and 1960 before they start to decrease again. Finally, as pointed out in Section 5.2, the data have high variability and, in the case of intrusive /r/, are also rather limited. Thus, further real-time analyses should be conducted that include more data points at multiple points across time. In the case of the reference corpus, this could include more potential cases per speaker, more years, more speakers or a combination of these. As for QEII, further material from other formal speeches could be gathered (e.g. State Opening of Parliament, formal reception speeches, etc.).

Irrespective of the issue of stable variation vs. sound change in progress, the differences in rates of production across studies reveal the existence of a wide range of inter-speaker and intra-speaker variation in the use of /r/-sandhi. This is likely to happen in the case of linking /r/ and that of intrusive /r/. Table 7 shows some rates in three recent corpus-based studies five years apart: Hannisdal (2006) on TV newsreaders, Mompean & Gómez (2011) on radio correspondents, and Pavlík (2016) on radio newscasters. These studies provide rates of linking /r/ lower than QEII's in the current study (i.e. 76.3%), which seems to point to the fact that QEII may be more of an /r/-linker than the average RP speaker, at least in the case of linking /r/.

Hannisdal's (2006) study is particularly useful here. In that study, 25 speakers (i.e. 83.3%) had rates below 70%, with 11 speakers in the 50%–59% range (the group's mean was 59.8%). Only three speakers had rates between 70% and 75% and only two between 76% and 80%. As reported in Section 5.2, QEII's rate is not within the 95% confidence interval.

The findings from the three corpus-based studies show that intrusive /r/ is relatively common in contemporary RP, ranging from 20% to 40% depending on the study. Yet QEII does not use it, despite its apparently greater acceptance in contemporary RP (Wells 1999). As in the case of linking /r/, there is also high inter-speaker variation in the use of intrusive /r/. The data from Hannisdal (2006), for example, show rates of 0% (3 speakers), 5%–10% (3 speakers), 10%–30% (11 speakers), 30%–50% (8 speakers), 51%–70% (3 speakers) and 71%–85% (2 speakers). As reported in Section 5.2, QEII's rate of intrusive /r/ is not within the 95% confidence interval.

Similar inter-speaker variation in the use of intrusive /r/ has also been found in NZE (Hay & Warren 2002, Viollain 2014). QEII could then be part of a minority of RP speakers that does not use intrusive /r/ (Wells 1982), at least in formal public speech. Related to this explanation, it has been pointed out that some speakers also often avoid intrusive /r/ altogether in experimental conditions. In their study of intrusive /r/ in New Zealand English, for example, Hay & Warren (2002) pointed out that their 16 participants, who read 48 sentences targeting 48 examples of intrusive /r/, ranged from 0% to 77%, with a median of 21% and a mean of 29%. Interestingly, one subject produced no tokens of intrusive /r/. According to Hay & Warren, intrusive /r/ was 'likely to be much lower than would occur in natural speech' (2002: 49).

As to why QEII avoids intrusive /r/ altogether, the formal register of the broadcasts cannot explain by itself all the facts, as the studies referred to above (i.e. Hannisdal 2006, Mompean & Gómez 2011, Pavlík 2016) have shown that speakers produce intrusive /r/ in formal contexts (e.g. newscasting). In contrast, gender cannot be overlooked as a possible explanation

for some of the inter-speaker variation observed in the studies. Most studies on /r/-sandhi in RP (Hannisdal 2006, Mompean & Mompean-Guillamón 2009, Pavlík 2016) have found no general gender effect. However, the three speakers in Hannisdal's study who avoided intrusive /r/ altogether were females. Related to this, it has been claimed that, in many speech communities, female speakers use a higher proportion of prestige forms than male speakers and tend to use fewer stigmatised forms than men, particularly in formal speech (Coates 2004). It should be borne in mind that intrusive /r/ has traditionally been regarded as a vulgarism (Crystal 1984) and it may still carry some degree of stigmatisation in RP. This stigmatisation may be apparent when attempting to elicit intrusive /r/ in experimental studies, as pointed out above, or when speaking to an audience. In this latter case, the Queen might, therefore, use a high proportion of prestige forms (linking /r/) and no stigmatised forms (intrusive /r/) in her (surely very) speech-conscious delivery of her broadcasts. It could also be mentioned, as possible evidence of the prestige of linking /r/, that the latter has been claimed to be a feature of colloquial style and not so much of scripted speech in non-rhotic English (Hannisdal 2006, Navarro 2013, Viollain 2014). However, the prestige of linking /r/ for the Queen may be such that she uses it even in formal speeches. Alternatively, as pointed out by one reviewer, what might be prestigious could be to know and apply the difference between linking /r/ – possibly neutral – and intrusive /r/ – somewhat stigmatised. This difference might be difficult to maintain for some speakers who, trying to avoid intrusive /r/ in formal contexts, may also tend to suppress linking /r/ as some sort of hypercorrection effect (Wells 1982, Brown 1988).

Another possible explanation of QEII's avoidance of intrusive /r/ could be related to her personal role. She has a unique status as both Queen of the United Kingdom and other Commonwealth states (political status) and a symbol/stereotype of her accentual variety (sociolinguistic status), which is often referred to in connection with herself as head of state (i.e. the Queen's English). During the Christmas broadcasts, then, she projects a public persona (the Queen) who is – as another reviewer points out – professionally careful in what she says, so it would not be surprising if she were always careful in how she says it, particularly when addressing such a very wide audience (the UK and the Commonwealth) in a very formal context. Given the emphasis on stylistic practice in third-wave approaches to sociolinguistic variation, which consider speakers as stylistic agents who place themselves in the social landscape through their stylistic practice (Eckert 2012), QEII's avoidance of intrusive /r/ could be a stylistic resource adding a layer of social meaning to her speech and indexing professional carefulness and attention to detail as well as general standardness.

A comparison of QEII's /r/-sandhi usage patterns with the studies mentioned above as well as with the data from the longitudinal group-level corpus and Hannisdal's (2006) study points to a two-way pattern. On the one hand, QEII seems to be a more-than-average /r/-linker (linking /r/) who shares the apparently decreasing use of linking /r/ on a longitudinal dimension with the community's trend. On the other hand, QEII seems to differ from the relatively low yet constant use of intrusive /r/ on a longitudinal dimension and also at a given point in time, as revealed by various corpus-based studies. This finding suggests that the same individual can be more or less in line and at the same time differ from group-level trends or norms depending on the linguistic feature analysed.

There are many reasons why individuals resemble or differ from such trends or norms in their language choices. In the case of intrusive /r/, as pointed out above, stylistic agency and/or the underlying prescriptivism of a language ideology may explain her avoidance of the phenomenon. Language ideologies – or sets of beliefs about languages as used in social contexts – often play a role in shaping and influencing linguistic structures and speech forms (Silverstein 1979). Prescriptivism can lead to a static situation in which a change in progress does not go to completion, is prevented from developing entirely, or is excluded from the standard forms of the language (Hickey 2012). If language users think, for example, that sounds not represented in spelling should not be pronounced (e.g. an r-sound in *I saw*[ɹ] *it*), their pronunciation is likely to reflect it eventually. Intrusive /r/ in RP and many other non-rhotic varieties seems to be in that situation. It is found across language cohorts and over

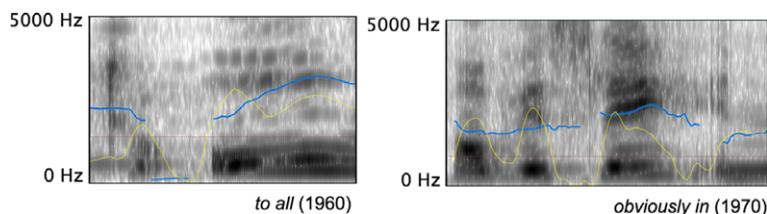
**Table 8** Rates of linking /r/ (with and without concomitant glottalisation) as well as rates of lack of linking /r/ (+glottalisation, hiatus) in potential /r/-sandhi sites: Current study (QXB) and Cox et al. (2014).

	QXB: RP		Cox et al. (2014): AusE	
	%	<i>n</i>	%	<i>n</i>
Linking	76.1	578	74.3	2194
–glott.	72.6	552	70.9	2093
+glott.	3.4	26	3.4	101
No linking	23.9	182	25.7	758
+glott.	21.3	162	22.8	673
hiatus	2.6	20	2.9	85

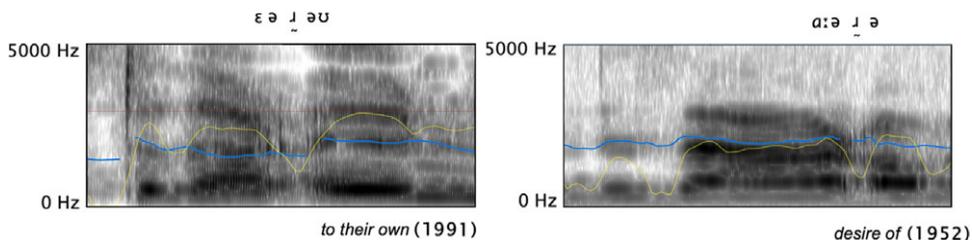
decades, but never quite reaches the rate of linking /r/. At the level of individuals such as QEII, a prescriptivist language ideology strictly adhered to, apparently, suppresses the phenomenon altogether, therefore setting her aside from the general trend. It should be pointed out, at this stage, that language ideology regarding intrusive /r/ may not be the same across all non-rhotic groups of speakers. Some studies have shown that rates of intrusive /r/ are typically lower among speakers with a higher socioeconomic status/education in NZE (Hay & Warren 2002, Hay & Sudbury 2005, Hay & Maclagan 2010) or Singaporean English (Tan 2012). The RP speakers in the general corpus and QEII are middle- to upper-class speakers, and their use of intrusive /r/ seems to fit this pattern. Interestingly, however, Foulkes (1997) found that intrusive /r/ was more frequent in Newcastle among working-class than among middle-class speakers in spontaneous conversation but not in reading passage style. Since intrusive /r/ is not a salient feature of the local Newcastle dialect, Foulkes speculated that intrusive /r/ could be seen as a non-regional feature often heard in the media and perceived as prestigious by middle-class speakers.

A noticeable finding in this study is how glottalisation is used to resolve hiatus at potential /r/-sandhi sites. Glottalisation can be described as a hiatus-breaking strategy when it occurs in isolation, which is attested at a non-negligible rate (over 20%). This finding is in line with a similar tendency in a corpus of modern BBC newscasts (Mompean & Gómez 2011), where glottalisation was used in 31.3% of all potential /r/-sandhi sites. The role of glottalisation as a hiatus-breaker at word boundaries is also confirmed by similar findings from a recent corpus-based study on /r/-sandhi in Australian English (AusE) by Cox et al. (2014) and by another study on a related external sandhi phenomenon known as glide insertion (Davidson & Erker 2014). As for Cox et al.'s (2014) study, it is interesting to note that the rates found in that study and in the current one are remarkably similar (see Table 5), although QEII's variety and Australian English are different – yet apparently similar in the case of /r/-sandhi. As Table 8 shows, glottalisation was used as the single hiatus-breaking strategy in over 20% of the potential cases.

Evidence on glottalisation as a more general hiatus-breaking strategy comes from a study on a sandhi phenomenon often referred to as glide insertion in V<sub>1</sub> [+high] + V<sub>2</sub> contexts, with the glide being [j] if the high vowel is front (e.g. *may I* (e.g. [meɪ<sup>h</sup>ai]) and [w] if the high vowel is back (e.g. *go inside* [gəʊ<sup>w</sup>ɪnsaɪd]). In their lab-based study of glide insertion in American English, Davidson & Erker (2014) found that potential cases of glide insertion across word boundaries were actually resolved with glottalisation. This happened in nearly half of the cases, with a higher rate when the linked vowel was stressed. Although glide insertion was not the object of study in the current research, a preliminary inspection was carried out of 50 potential glide insertion contexts in the QXB corpus. This confirmed that QEII also uses glottalisation in such contexts (around 50% of the time), as shown in Figure 6, both when the second vowel is stressed (e.g. *to all*) or unstressed (e.g. *obviously in*).



**Figure 6** (Colour online) Examples of glottalisation at potential glide insertion contexts when the second vowel is stressed (left) and unstressed (right).



**Figure 7** (Colour online) Simultaneous glottalisation and linking /r/ at sites followed by stressed (left) and unstressed (right) syllables.

As to why glottalisation may override /r/-sandhi, it could be argued that glottalisation is less marked than coronal articulations (Lombardi 2002) and that it tends, therefore, to replace /r/-sandhi in areas with significant language and sociocultural contact (Britain & Fox 2009, Fox 2015) and in first language acquisition, where glottalisation seems to be the default form in children's English until forms are acquired at a relatively later stage (Newton & Wells 2002). Thus, although some have tried to provide theoretical explanations of /r/-sandhi as a natural process and /r/ as a good sandhi segment (e.g. Broadbent 1991, Orgun 2001, Lombardi 2002, Uffmann 2007), /r/-sandhi may, after all, be a historical artefact (Britain & Fox 2009), a synchronically arbitrary process with rare equivalents in the world's languages (e.g. Blevins 2008, Hall 2013). Glottalisation might, therefore, represent a more 'natural' strategy to prevent hiatus. Increasing use of glottalisation could be, then, one reason for the decrease in the use of linking /r/ noted in various studies (Allerton 2000).

Irrespective of the motivation for the use of glottalisation instead of /r/-sandhi, studies such as Cox et al. (2014) or the current study have implications in our understanding hiatus resolution in /r/-sandhi in non-rhotic accents. More specifically, the finding points to a more complex system of hiatus resolution at both speech community and individual levels than that typically modelled by most phonological accounts (see Section 6.2 below). In these models, /r/-sandhi is assumed to be categorical. Faced with the evidence that /r/-sandhi is not a categorical phenomenon, hiatus might seem to be the default value. However, hiatus is rare and glottalisation is the norm when /r/-sandhi is not used. To complicate matters even more, /r/-sandhi and glottalisation can also co-occur, although this co-occurrence is further discussed below.

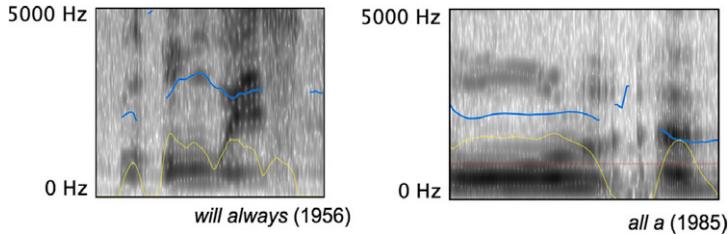
The study has also shown that /r/-sandhi and glottalisation co-occur at external /r/-sandhi site, although at a relatively marginal rate. Figure 7 shows how linking /r/ and glottalisation co-occur at linking /r/ sites followed by stressed (e.g. *to their own*) and unstressed (e.g. *desire of*) syllables. The examples involve creaky voice but not full glottal stops, which are rare in QEII's speech in these contexts.

To the best of our knowledge, this combination of /r/-sandhi and glottalisation only seems to have been found previously by Cox et al. (2014) in their study on contemporary AusE,

which also found the combination to occur at a very low rate (3.4%). An interpretation of such a finding, however, is necessary. The use of linking /r/ alongside glottalisation could be interpreted as a combination of two strategies to resolve hiatus at potential /r/-sandhi sites. Yet the glottalisation observed alongside /r/-sandhi might be a word-boundary phenomenon that has nothing to do with an attempt to break hiatus. To start with, the co-occurrence of /r/-sandhi and glottalisation seems to be rare, as revealed by the very low rates in both the current study and in Cox et al.'s study. If glottalisation reinforced /r/-sandhi in an attempt to avoid hiatus its frequency would have been expected to be much higher. However, the low frequency of the combination may be one reason why it has been unnoticed until now. Some recent work has specifically addressed the possible use of /r/-sandhi, glottalisation, or both in L1 English learners of L2 German. Bissiri & Scobbie (2016, 2017) carried out both acoustic and ultrasound recordings of four such learners producing English and German sentences with potential /r/-sandhi contexts (e.g. *cider apples*; *jeder Abend* 'every evening'). The researchers were interested in the possible role of the presence/absence of a phrase boundary at /r/-sites and the stress/lack of stress in the linked vowel. Moreover, they were also interested in whether there was any transfer of glottalisation and /r/-sandhi from learners' L1 (English) to their L2 (German). Bissiri and Scobbie's preliminary findings revealed that /r/-sandhi was only articulated in L1 English without an intervening phrase boundary and when the linked vowel was deaccented. /r/-sandhi was also absent in L2 German. Interestingly, no overlapping of /r/-sandhi and glottalisation was found in L1 English. Glottalisation was found, however, at phrase boundaries and before accented linked vowels, and more frequently in G2 German than in L1 English.

One manner of checking whether the glottalisation occurring alongside /r/-sandhi does not specifically serve a hiatus-breaking purpose would be to inspect whether glottalisation also occurs in positions that attract creaky voice at phrase-internal word boundaries. Previous studies (e.g. Cox et al. 2014, Yuen et al. 2015, Pavlík 2016) and the current one have found that /r/-sandhi is disfavoured when the second hiatus vowel is stressed. In many such cases, however, glottalisation is used. Yet glottalisation seems a more general phenomenon at phrase-internal word boundaries. A case in point is word-initial glottalisation (Garellek 2012) or the phenomenon whereby word-initial stressed and/or accented vowels are preceded by a glottal stop and/or have creaky voice (e.g. Nakatani & Dukes 1977, Umeda 1978), irrespective of whether the segment before the word boundary is a consonant or a vowel or whether there is a pause. Research has found that word-initial glottalisation occurs with higher frequency at the beginning of tone unit boundaries and to a greater degree if the word is pitch-accented (Dilley, Shattuck-Hufnagel & Ostendorf 1996). Yet, the segmental context also plays a role, with glottalisation occurring significantly more frequently after vowels – particularly back vowels – and liquids at phrase-internal word boundaries (Umeda 1978, Dilley et al. 1996). This finding is consistent with an Electropalatography (EPG) study by Scobbie & Pouplier (2010) on word-final /l/ produced by speakers of SSB (or RP) and Scottish-accented Standard English. The researchers found that speakers from both groups often used glottalisation alongside /l/ in word-final contexts followed by a vowel-initial, lexically stressed syllable in an accented word. In the case of QEII, limitations of space prevent a thorough analysis of her use of glottalisation in the corpus under analysis. However, an inspection of 50 tokens of  $V_1/l/ + V_2$  with an intervening phrase-internal word boundary revealed that glottalisation was common alongside /l/ followed by stressed (e.g. *will always*) or unstressed (e.g. *all a*) vowels. Figure 8 shows some examples with /l/ followed by a stressed or an unstressed vowel.

Since word-final /l/ and /r/-sandhi would already prevent hiatus, the concomitant presence of glottalisation could serve some prosodic function given that glottalisation is used for various prosodic and communicative functions (see e.g. Redi & Shattuck-Hufnagel 2001). Some of these functions could be, depending on the case, cuing open juncture, marking the boundary between two rhythmic feet or between two intonation units, or a combination of



**Figure 8** Glottalisation at  $V_1/V_2$  contexts in QEII: stressed  $V_2$  (left) and unstressed  $V_2$  (right).

these. In other words, glottalisation alongside /r/-sandhi may be acting as an external boundary signal rather than simply as a hiatus-breaking strategy. Glottalisation without /r/-sandhi, however, can be considered as a hiatus-breaking strategy, although it may also be performing other prosodic functions.

## 6.2 QEII and factors conditioning /r/-sandhi usage

The results confirm the hypotheses that phonetic and usage-based variables condition /r/-sandhi usage. The hypothesis that grammatical variables such as syntactic boundaries and word class condition /r/-sandhi usage is, however, not confirmed.

As regards phonetic factors, the study has shown that the absence of a TU boundary predicts the use of linking /r/. This is in line with suggestions that /r/-sandhi is more frequent within phonological phrases than phrase-finally (Joseph 1999) or within clauses than across clause boundaries (Foulkes 1997), which may effectively mark prosodic boundaries. Besides, the results suggest that the presence of a stressed linking syllable and the absence of stress in the linked syllable facilitate linking /r/. These results indicate that a prototypical example of linking /r/ would be in an SW prosodic site and least likely of all in a WS site. This result is in line with Cox et al.'s (2014) finding that the likelihood of occurrence of linking /r/ is higher in SW prosodic sites (86% in her study) as opposed to WS sites (49%). These findings suggest that that prosodic and metrical effects, found to be critical in the use of /r/-sandhi at group level (Cox et al. 2014), also occur at the idiolectal level.

While the influence of prosodic factors is clear in this study, the contribution of segmental ones is more modest. The logistic regression analysis suggests that the presence of /r/ one syllable before the linking syllable prevents the use of linking /r/, pointing to a dissimilation effect. This contrasts with previous studies such as Pavlík's (2016), as well as Mompean & Mompean-Guillamón's (2009), which found little or no effect of a neighbouring /r/ in the production of linking /r/ at all in RP. However, our findings coincide with those reported by Hannisdal (2006), who found that the presence of /r/ in the vicinity inhibited /r/-sandhi. Finally, our analysis suggests that the presence of an open back vowel /ɑ:/ in the previous syllable facilitates the use of linking /r/. However, no other linking or linked vowel quality category had a significant effect. Thus, the findings obtained in this study seem to suggest that prosodic factors are better predictors of /r/-sandhi usage than segmental ones, at least in QEII's sample.

The results also show that grammatical factors seem to play little or no role in the production of linking /r/. The hypothesis formulated in this respect, that is, that phonological factors (TU boundary, stress, neighbouring /r/, vowel quality) would have a greater effect than grammatical factors (syntactic boundaries and word class) on /r/-sandhi is confirmed. In this respect, linking /r/ may be avoided when syntactic boundaries coincide with TU boundaries but less likely when they do not. As to the lexical vs. grammatical nature of the linking and linked words, the results obtained are in line with the absence of an effect described by Hay & Sudbury (2005) or the little effect found by Pavlík (2016). It may be the case that the perception that function words could favour /r/-sandhi can simply be an effect of

prosodic stress and the use of schwa in function words, which are typically unstressed. It seems, therefore, that prosodic structure predicts /r/-sandhi better than syntactic structure or grammatical information.

The fact that prosodic factors are better predictors of /r/-sandhi usage than segmental ones in the QXB corpus underscores the relevance of prosodic phrasal context in /r/-sandhi usage. This context may also explain rate of speech and coarticulation effects. As discussed below regarding the phonological modelling of /r/-sandhi, the latter is often modelled as a segmental process. However, the phenomenon can also be viewed as a type of gradient hyper- and hypo-articulation (H&H) or overshoot/undershoot alternation. It was mentioned above (Sections 2.3, 4.2.2) that /r/ displays gradient features in its realisation. Moreover, speech rate might also determine /r/-sandhi production. It has been pointed out, for example, that linking /r/ is less common at both rather markedly rapid and slow rates (Lewis 1975). Pavlík (2016), for example, found a small effect of speech rate on /r/-sandhi production, with a decrease in linking /r/ as speech rate increased. This could be due to the fact that fast speech is characterised by high articulator velocity and gestural overlap. A reviewer suggests that this might be indicative of lack of (post-lexical) phonetic planning in more citation-like, slower speech with frequent pauses vs. undershoot in faster speech. QEII's Christmas broadcasts seem to be delivered at neither a noticeably slow nor a particularly fast speech rate (i.e. average tempo). Rate of speech was not investigated in the current study. However, future studies should look at the possible influence of rate of speech in the same speaker as well as the possible gradient timing/gestural correlations.

Turning now to the usage-based variables, the results indicate a clear effect of frequency on the use of linking /r/. Yet they reveal that collocational frequency is more relevant than the individual lexical or vowel quality frequencies of the words or vowels, respectively, at either side of /r/-sandhi contexts. The results indicate that QEII is less likely to produce /r/ when the lexical frequency of the linked word is high. This coincides with the results obtained by Hay & Sudbury (2005). However, the fact that the lexical frequency of the linking word did not have an influence on linking /r/, and that the frequency of the linked word had only a minimal effect (only in the first analysis and in the QXB corpus) indicates that individual lexical frequencies are probably not an important predictor of /r/-sandhi usage.

The above-mentioned finding is in line with that of a study on the influence of a range of phonological, morphological and lexical frequency variables on the use of t/d-deletion in Canadian English by Walker (2012). He found that higher lexical frequency correlated with higher deletion rates. Yet, when several factors were taken into account, the effect of lexical frequency was non-significant. Given the existence of a body of work that suggests a major role for frequency in phonology (e.g. Bybee 2001), Walker pointed out that frequency effects might lie beyond the word, suggesting collocations as a possible locus for those effects. The current study confirms this: collocational frequency had the highest predictive power, with higher collocational frequencies favouring linking /r/ in both the BNC and the QXB corpora. This may not be surprising given that, as discussed by Biber and colleagues (Biber, Conrad & Cortes 2004), frequent collocations are often used with some pragmatic or discourse function (e.g. thematisation or construction of referential themes such as the year's events and QEII's wishes for the following year (e.g. *this year I . . . , the year ahead . . .*). Thus, they help speakers (in this case QEII) perform discourse functions. Moreover, they play an important part in the speaker's communicative repertoire and may lexicalise – or entrench – phenomena such as /r/-sandhi, giving rise to higher rates, as is the case, for example, of word-internal linking /r/ in compounds. From the point of view of exemplar theory, frequent collocations with a potential /r/-sandhi context have more chances of being produced with /r/. If /r/ is actually produced relatively frequently, speakers' mental representation will contain many memories of that collocation produced with /r/. When those memories are averaged over for a production target, in turn, a robust /r/ will result (Hay & Maclagan 2010, 2012, Cox et al. 2014). Alternatively, the higher frequency of linking /r/ in frequent collocations can be interpreted 'as a remnant of the historical /r/ which continues to resist the /r/-loss in

British English', which is still in progress (Pavlik 2016: 112). As Bybee (2001, 2006) points out, high frequencies encourage phonetic change but they render items more conservative in the face of grammatical or analogical change. As a case in point, most verbs that are still irregular in contemporary English (e.g. *say*, *go*, *get*, etc.) have a high lexical frequency, the latter strengthening their representation in memory and helping them resist the pressure from analogical change.

In addition to lexical collocation frequency, and to the best of our knowledge, this is the first study that describes the role of vowel quality collocational frequency, showing that higher frequencies had a positive effect on the rate of linking /r/. The most frequent vowel quality collocations were mid central + mid central (e.g. /ə.ə/), mid central + close front (e.g. /ə.i/), mid central + open central/back (e.g. /ə.ɑ:/), mid central + close-mid/open back (e.g. /ə.ɔ:/), open back + mid central (e.g. /ɑ:.ə/), close-mid back + mid central (e.g. /ɔ:.ə/), and mid central + mid/open front (e.g. /ə.e/). As can be seen, mid central vowels (particularly schwa and centring diphthongs) are involved in vowel quality collocations favouring linking /r/. Collocations that did not contain mid central vowels had no favouring effect on linking /r/. Finally, it should be pointed out that vowel type frequency did not have an effect on the use of linking /r/, which is consistent with the hypothesis that collocational frequency is more relevant than individual vowel frequencies.

The findings in relation to usage-based variables provide arguments for the debate regarding the phonological modelling of /r/-sandhi in non-rhotic English. Most traditional analyses assume that no underlying coda /r/ is present word-finally and that /r/-sandhi arises as a function of /r/-insertion (e.g. Vennemann 1972, Johansson 1973). Other approaches assume that underlying coda /r/ is present but that /r/-deletion operates before a consonant or pause (e.g. Mohanan 1986, Harris 1994). Yet other accounts claim that both rules (/r/-insertion and /r/-deletion) coexist in the synchronic phonology of varieties (e.g. Anttila & Cho 1998, Kijak 2009). All these approaches aim at constructing synchronic models of competence and exclude considerations of function and usage, so they cannot predict inter-/intra-speaker variation.

An alternative view is to consider lexical representations as being modelled as exemplars of remembered sequences (e.g. Bybee 2001) and that surface forms are not generated from underlying representations. Within this more functional, usage-based framework, every linguistic item in the lexicon consists of a cloud of phonetic exemplars containing detailed information about the context in which they are used. In this view, /r/-sandhi tokens would not be the result of deletion or insertion rules, but rather the result of how entrenched or lexicalised /r/ is in specific utterances (Hay & Sudbury 2005, Mompeán & Mompeán-Guillamón 2009, Hay & Maclagan 2012, Sóskuthy 2013, Pavlik 2016). The more frequent a collocation including a potential /r/-sandhi is, the more likely it is that /r/ will become lexicalised in it and thus be produced. The usage-based factors described in this study can be interpreted as evidence in favour of a usage-based framework. However, this interpretation does not rule out the possibility of a hybrid account including abstract representations, exemplars and phonetic implementation processes as suggested by some researchers (e.g. Cox et al. 2014).

More generally, a usage-based interpretation of the findings also allows accommodating not only inter-group but also inter-speaker variation within the same group, given that language users differ in the quality and quantity of input they receive, leading to overlapping grammars that nevertheless differ in some respects. A functional, usage-based interpretation can also accommodate differences in the relative weight of some predictors when comparing individuals and group-level trends and, possibly, any effects of the individual's agency on his/her production. One interesting conclusion from the study of the factors conditioning /r/-sandhi usage in QEII is that the multiple variables contributing to variation in a given phenomenon – which studies at the group-level typically show – can also be seen at the individual level. Yet the weight of relevant variables may be different across individuals given various idiosyncrasies as well as their different linguistic experience.

## 7 Conclusion

This paper has presented findings from a longitudinal analysis of /r/-sandhi in QEII's Christmas broadcasts over a period of 67 years. The study has revealed that QEII's speech is in line with speech community-level tendencies (e.g. greater use of linking /r/ than intrusive /r/, use of glottalisation to resolve hiatus when linking /r/ is not used, little use of pure hiatus, decrease use of linking /r/ over the years), while it also differs in some respects from those tendencies (e.g. complete lack of intrusive /r/, slightly higher use of linking /r/ overall). Moreover, the study has shown how QEII's use of /r/-sandhi is influenced by several phonetic (e.g. presence of TU, stressed/unstressed nature of the linking and linked syllables, etc.) and usage-based variables (e.g. vowel quality collocation frequency) as well as some previously unreported findings, including the combined use of linking /r/ and glottalisation to resolve hiatus or the effect of vowel quality collocation frequency.

The current study has limitations that should be acknowledged. One such limitation is that the data analysed represent a corpus over which researchers have less control, as opposed to data elicited in experimental settings, and, in some cases, data were limited. This made it more difficult, for example, to shed more light on issues such as potential sound change in linking /r/ or speaker-group relationships in the case of intrusive /r/, among others. Given this, future studies should include more potential cases per speaker, more years, more speakers or a combination of these. Future studies of /r/-sandhi could look at the influence of different styles (scripted vs. spontaneous), although this may prove difficult in the case of QEII since she has never given an interview. It would also be interesting to carry out comparisons of the dialects of two or more individual speakers and consider other variables not addressed in the study. As pointed out by one reviewer, it might also be interesting to see if /r/-sandhi might be affected by the phonetic type or allophone of /r/ that is used (e.g. the alveolar tap [ɾ] or the labio-dental approximant [v]).

## Acknowledgements

I am very grateful to the journal's anonymous reviewers for their invaluable feedback on earlier versions of this paper. Their comments and recommendations have been extremely helpful. Thanks go also to the SAE staff at the University of Murcia for their advice on statistical analyses, particularly Francisco Javier Ibáñez López, Antonio Maurandi López, María Elvira Ferre Jaén and Álvaro Hernández Vicente. This work was funded by Project PGC2018-095050-B-I00, funded by FEDER/Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación, on the compilation of a diachronic corpus of spoken English, currently under construction.

## References

- Allerton, David J. 2000. Articulatory inertia vs. 'Systemzwang': Changes in liaison phenomena in recent British English. *English Studies* 81(6), 574–581.
- Anttila, Arto & Young-mee Y. Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104(1/2), 31–56.
- Ashby, Michael & Joanna Przedlacka. 2014. Measuring incompleteness: Acoustic correlates of glottal articulations. *Journal of the International Phonetic Association* 44(3), 283–296.
- Audacity Team. 2012. *Audacity* (version 2.1.0). <https://www.audacityteam.org/> (accessed 30 June 2020).
- Bauer, Laurie. 1984. Linking /r/ in RP: Some facts. *Journal of the International Phonetic Association* 14(2), 74–79.
- Biber, Douglas, Susan Conrad & Viviana Cortes. 2004. If you look at . . . : Lexical bundles in university teaching and textbooks. *Applied Linguistics* 25(3), 371–405.
- Bissiri, Maria P. & James M. Scobbie. 2016. Cider apples and *jeder Abend*: The interplay of /r/-sandhi and word-initial glottalisation in English-accented German. Poster presentation at the 15th Conference on Laboratory Phonology (LabPhon 2016), 14 July 2016, Cornell University, Ithaca, NY.

- Bissiri, Maria P. & James M. Scobbie. 2017. Tongue dynamics at word-boundaries in English and English-accented German: The influence of prosodic structure and word-initial glottalisation. In *Proceedings of the 13<sup>e</sup> Convegno Nazionale AISV, Sources and functions of speech variation. Disentangling the role of biological and social factors*, 25–27 January 2017, Scuola Normale Superiore di Pisa, Pisa, Italy. [http://linguistica.sns.it/AISV/abstract/Bissiri\\_Scobbie.pdf](http://linguistica.sns.it/AISV/abstract/Bissiri_Scobbie.pdf) (accessed 30 June 2020).
- Blevins, Juliette. 2008. Consonant epenthesis: Natural and unnatural histories. In Jeff Good (ed.), *Linguistic universals and language change*, 79–107. Oxford: Oxford University Press.
- Bowie, David. 2015. Phonological variation in real time: Patterns of adult linguistic stability and change. In Annette Gerstenberg & Anja Voeste (eds.), *Language development: The lifespan perspective*, 603–618. Amsterdam & Philadelphia: John Benjamins.
- Britain, David & Susan Fox. 2009. The regularisation of the hiatus resolution system in British English: A contact-induced ‘vernacular universal’? In Markku Filppula, Juhani Klemola & Heli Paulasto (eds.), *Vernacular universals and language contacts: Evidence from varieties of English and beyond*, 177–205. London: Routledge.
- Broadbent, Judith. 1991. Linking and intrusive *r* in English. *UCL Working Papers in Linguistics* 3, 281–302.
- Brown, Adam. 1988. Linking, intrusive, and rhotic /r/ in pronunciation models. *Journal of the International Phonetic Association* 18(2), 144–151.
- Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan L. 2006. From usage to grammar: The mind’s response to repetition. *Language* 82(4), 711–733.
- Calcagno, Vincent. 2013. Glmulti: Model selection and multimodel inference made easy. R package version 1.0.7. <https://cran.r-project.org/web/packages/glmulti/index.html> (accessed 30 June 2020).
- Catsiapis, Hélène. 2005. The Queen’s Christmas messages. In Peter Vernon (ed.), *Seeing things: Literature and the visual. Papers from the Fifth International British Council Symposium, September 2001*, 73–88. Tours: Presses universitaires François-Rabelais.
- Coates, Jennifer. 2004. *Women, men and language: A sociolinguistic account of gender differences in language*, 3rd edn. London & New York: Longman.
- Cox, Felicity, Sallyanne Palethorpe, Linda Buckley & Samantha Bentink. 2014. Hiatus resolution and linking ‘r’ in Australian English. *Journal of the International Phonetic Association* 44(2), 155–178.
- Cruttenden, Alan. 2014. *Gimson’s Pronunciation of English*, 8th edn. London: Routledge.
- Crystal, David. 1984. Should intruders keep out? In David Crystal (ed.), *Who cares about English usage?* 36–44. London: Penguin.
- Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2(3), 219–253.
- Dannenberg, Clare J. 2000. Sociolinguistics in real time. *American Speech* 75(3), 254–257.
- Davidson, Lisa & Daniel Erker. 2014. Hiatus resolution in American English: The case against glide insertion. *Language* 90(2), 482–514.
- Davies, Mark. 2004. BYU–BNC: The British National Corpus. <http://corpus.byu.edu/bnc/> (accessed 30 June 2020).
- Dilley, Laura, Stefanie Shattuck-Hufnagel & Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4), 423–444.
- Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41, 87–100.
- Fabricius, Anne H. 2017. Twentieth-century RP: Prevocalic /r/. In Raymond Hickey (ed.), *Listening to the past*, 39–65. Cambridge: Cambridge University Press.
- Ferguson, Charles A. 1979. Phonology as an individual access system: Some data from language acquisition. In Charles J. Fillmore, Daniel Kempler & William S.-Y. Wang (eds.), *Individual differences in language ability and language behavior*, 189–201. New York: Academic Press.
- Fletcher, William H. 2007. Implementing a BNC-comparable web corpus. In Cédrik Fairon, Hubert Naets, Adam Kilgariff & Gilles-Maurice de Schryver (eds.), *Building and exploring web corpora: Proceedings of the WAC3 Conference*, 43–56. Louvain: Presses Universitaires de Louvain.

- Flickinger, Dan. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(1), 15–28.
- Foulkes, Paul. 1997. English [r]-sandhi: A sociolinguistic perspective. *Historie, Epistémologie, Langage* 19(1), 73–96.
- Foulkes, Paul & Gerard J. Docherty (eds.). 1999. *Urban voices: Accent studies in the British Isles*. London: Arnold.
- Foulkes, Paul & Gerard J. Docherty. 2002. Another chapter in the story of /r/: ‘Labiodental’ variants in British English. *Journal of Sociolinguistics* 4(1), 30–59.
- Fox, Susan. 2015. *The new Cockney: New ethnicities and adolescent speech in the traditional East End of London*. Basingstoke: Palgrave Macmillan.
- Gardiner, Shayna & Naomi Nagy. 2017. Stable variation vs. language change and the factors that constrain them. *University of Pennsylvania Working Papers in Linguistics* 23(2), 78–85.
- Garellek, Marc. 2012. Word-initial glottalization and voice quality strengthening. *UCLA Working Papers in Phonetics* 111, 92–122.
- Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16(1), 29–54.
- Gick, Bryan. 2002. The American intrusive /l/. *American Speech* 77(2), 167–183.
- Giegerich, Heinz J. 1999. *Lexical strata in English: Morphological causes, phonological effects*. Cambridge: Cambridge University Press.
- Hall, Tracy A. 2013. How common is r-epenthesis? *Folia Linguistica* 47, 55–87.
- Halle, Morris & William J. Idsardi. 1997. r, hypercorrection, and the elsewhere condition. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 331–348. Oxford: Clarendon Press.
- Hannisdal, Bente R. 2006. *Variability and change in Received Pronunciation: A study of six phonological variables in the speech of television newsreaders*. Ph.D. dissertation, Universitetet i Bergen.
- Harrington, Jonathan. 2006. An acoustic analysis of ‘Happy-tensing’ in the Queen’s Christmas broadcasts. *Journal of Phonetics* 34(4), 439–457.
- Harrington, Jonathan. 2007. Evidence for a relationship between synchronic variability and diachronic change in the Queen’s annual Christmas broadcasts. In Jennifer Cole & Jose I. Hualde (eds.), *Laboratory phonology* 9, 125–143. New York & Berlin: Mouton de Gruyter.
- Harrington, Jonathan, Sallyanne Palethorpe & Catherine Watson. 2000. Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen’s Christmas broadcasts. *Journal of the International Phonetic Association* 30(1/2), 63–78.
- Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
- Hartmann, Dieter & Sabine Zerbian. 2009. Rhoticity in Black South African English: A sociolinguistic study. *Southern African Linguistics and Applied Language Studies* 27(2), 135–148.
- Hashim, Emilda, Norimah Rabli, Maryam Mahdinezhad & Normala Zulkifli. 2016. 4P marketing & housewives’ expenditure: Multiple regression model. *International Journal of Academic Research in Business and Social Sciences* 6(12), 810–820.
- Hay, Jennifer & Margaret Maclagan. 2010. Social and phonetic conditioners on the frequency and degree of ‘intrusive /r/’ in New Zealand English. In Dennis R. Preston & Nancy Niedzielski (eds.), *A reader in sociophonetics*, 41–70. New York & Berlin: Mouton de Gruyter.
- Hay, Jennifer & Margaret Maclagan. 2012. /r/-sandhi in early 20th century New Zealand English. *Linguistics* 50(4), 745–763.
- Hay, Jennifer & Andrea Sudbury. 2005. How rhoticity became /r/-sandhi. *Language* 81(4), 799–823.
- Hay, Jennifer & Paul Warren. 2002. Experiments on /r/-intrusion. *Wellington Working Papers in Linguistics* 14, 47–58.
- Hickey, Raymond. 2012. Internally and externally motivated language change. In Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (eds.), *The handbook of historical sociolinguistics*, 387–407. Chichester & Malden, MA: Wiley-Blackwell.
- Hinton, Martin. 2015. Changes in Received Pronunciation: Diachronic case studies. *Research in Language* 13(1), 21–37.
- Johansson, Stig. 1973. Linking and intrusive /r/ in English: A case for a more concrete phonology. *Studia Linguistica* 27, 53–68.
- Jones, Daniel. 1956. *The pronunciation of English*. London: Cambridge University Press.

- Joseph, Brian D. 1999. Utterance-finality: Framing the issues. In Bohumil Palek, Osamu Fujimura & Brian D. Joseph (eds.), *Proceedings of LP '98: Item order in language and speech*, vol. 2, 3–13. Prague: Charles University Press.
- Kijak, Artur. 2009. Non-rhoticity or where is the historical r lurking? *Linguistica Silesiana* 30, 43–57.
- Kirasich, Kaitlin, Trace Smith & Bivin Sadler. 2018. Random forest vs logistic regression: Binary classification for heterogeneous datasets. *SMU Data Science Review* 1(3), 9. <https://scholar.smu.edu/datasciencereview/vol1/iss3/9>.
- Lawson, Eleanor, Jane Stuart-Smith & James M. Scobbie. 2018. The role of gesture delay in coda /r/ weakening: An articulatory, auditory and acoustic study. *The Journal of the Acoustical Society of America* 143(3), 1646–1657.
- Lewis, Jack W. 1975. Linking /r/ in the General British pronunciation of English. *Journal of the International Phonetic Association* 5(1), 37–42.
- Lisker, Leigh. 1957. Minimal cues for separating /w,r.l.y/ in intervocalic position. *Word* 13(2), 257–267.
- Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19(2), 219–251.
- McCarthy, John J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38, 127–153.
- Mohanan, Karuvannur P. 1986. *The theory of Lexical Phonology*. Dordrecht: Reidel.
- Mompean, Jose A. (2019). Diachronic change in /r/-sandhi? A real-time study at community and individual levels. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS XIX)*, Melbourne, Australasian Speech Science and Technology Association Inc., 770–774.
- Mompean, Jose A. & Alberto Gómez. 2011. Hiatus-resolution strategies in non-rhotic English: The case of /r/-sandhi. In Lee Wai-Sum & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, IPA/City University of Hong Kong, 1414–1417.
- Mompean, Jose A. & Juan M. Hernández-Campoy. 2001. Advantages and disadvantages of RP as an EFL model of pronunciation. In Isabel de La Cruz Cabanillas (ed.), *La Lingüística aplicada a finales del siglo XX: Ensayos y propuestas*, 707–713. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá de Henares.
- Mompean, Jose A. & Pilar Mompean-Guillamón. 2009. /r/-liaison in English: An empirical study. *Cognitive Linguistics* 20(4), 733–776.
- Mullooly, Richard. 2004. *An electromagnetic articulograph study of alternating [r] and the effects of emphatic stress on the articulation of rhotic consonants*. Ph.D. dissertation, Queen Margaret University College.
- Nagy, Naomi & Patricia Irwin. 2010. Boston (r): Neighbo(r)s nea(r) and fa(r). *Language Variation and Change* 22(2), 241–278.
- Nakatani, Lloyd H. & Kathleen D. Dukes. 1977. Locus of segmental cues for word juncture. *The Journal of the Acoustical Society of America* 62(3), 714–719.
- Navarro, Sylvain. 2013. *Rhoticité et 'r' de sandhi en anglais: Du Lancashire À Boston*. Ph.D. dissertation, Université de Toulouse Jean Jaurès.
- Newton, Caroline & Bill Wells. 2002. Between-word junctures in early multi-word speech. *Journal of Child Language* 29(2), 275–299.
- Orgun, Cemil O. 2001. English r-insertion in Optimality Theory. *Natural Language & Linguistic Theory* 19(4), 737–749.
- Pavlik, Radoslav. 2016. A usage-based account of /r/-sandhi in Standard British English. *Journal of Phonetics* 54, 109–122.
- Piercy, Caroline. 2012. A transatlantic cross-dialectal comparison of non-prevocalic /r/. *University of Pennsylvania Working Papers in Linguistics* 18(2), 76–86.
- Redi, Laura & Stefanie Shattuck-Hufnagel. 2001. Variation in the realisation of glottalization in normal speakers. *Journal of Phonetics* 29, 407–429.
- Renouf, Antoinette, Andrew Kehoe & Jay Banerjee. 2007. WebCorp: An integrated system for web text search. In Marianne Hundt, Nadja Nesselhauf & Carolin Biewer (eds.), *Corpus linguistics and the web*, 47–67. Amsterdam & New York: Rodopi.

- Robson, David. 2016. Has the Queen become frightfully common? BBC Future. <http://www.bbc.com/future/story/20160202-has-the-queen-become-frightfully-common/> (accessed 30 June 2020).
- Sankoff, Gillian & H el ene Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83(3), 560–588.
- Scobbie, James M. & Marianne Pouplier. 2010. The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/. *Journal of Phonetics* 38, 240–259.
- Silverstein, Michael. 1979. Language structure and linguistic ideology. In Paul R. Clyne, William F. Hanks & Carol L. Hofbauer (eds.), *Papers from the 15th regional meeting of the Chicago Linguistic Society* (CLS 15). *The elements: A parasession on linguistic units and levels*, vol. 2, 193–247. Chicago, IL: Chicago Linguistic Society.
- S okuthy, M arton. 2013. Analogy in the emergence of intrusive-r in English. *English Language and Linguistics* 17(1), 55–84.
- Stevens, Mary & Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1), e003. doi: <http://dx.doi.org/10.3989/loquens.2014.003>.
- Tan, Ying-Ying. 2012. To r or not to r: Social correlates of /r/ in Singapore English. *International Journal of the Sociology of Language* 218, 1–24.
- Tillery, Jan and Guy Bailey. 2003. Approaches to real time in dialectology and sociolinguistics. *World Englishes* 22(4), 351–365.
- Tuinman, Annelie, Holger Mitterer & Anne Cutler. 2011. Perception of intrusive /r/ in English by native, cross-language and cross-dialect listeners. *The Journal of the Acoustical Society of America* 130(3), 1643–1652.
- Uffmann, Christian 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29(2/3), 451–476.
- Umeda, Noriko. 1978. Occurrence of glottal stops in fluent speech. *The Journal of the Acoustical Society of America* 64(1), 88–94.
- Upton, Clive. 2008. Received Pronunciation. In Clive Upton & Bernd Kortmann (eds.), *Varieties of English: The British Isles*, 237–252. Berlin & New York: Mouton de Gruyter.
- Vennemann, Theo. 1972. Rule inversion. *Lingua* 29, 209–242.
- Viollain, C elina. 2014. *Sociophonologie de l'anglais contemporain en Nouvelle-Z elande: Corpus et dynamique des syst emes*. Ph.D. dissertation, Universit e de Toulouse Jean Jaur es.
- von Mengden, Ferdinand & Evie Couss e. 2014. Introduction: The role of change in usage-based conceptions of language. In Evie Couss e & Ferdinand von Mengden (eds.), *Usage-based approaches to language change*, 1–19. Amsterdam & Philadelphia, PA: John Benjamins.
- Wagner, Michale & Duane G. Watson. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes* 25(7/9), 905–945.
- Walker, James A. 2012. Form, function, and frequency in phonological variation. *Language Variation and Change* 24(3), 397–415.
- Wells, John C. 1982. *Accents of English*, vol. 2. Cambridge: Cambridge University Press.
- Wells, John C. 1999. Whatever happened to Received Pronunciation? In Carmelo Medina-Casado & Concepci n Soto-Palomo (eds.), *II Jornadas de estudios ingleses*, 19–28. Ja en: Universidad de Ja en, Servicio de Publicaciones.
- Wieling, Martijn. (2018). Analyzing dynamic phonetic data using generalised additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70, 86–116.
- Yuen, Ivan, Felicity Cox & Katherine Demuth. 2015. Anticipatory planning of r-insertion in Australian English. In The Scottish Consortium for ICPhS 2015 (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences* (ICPhS XVIII), Glasgow, The University of Glasgow, <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0521.pdf> (accessed 30 June 2020).