

# **/r/-liaison in English: An empirical study\***

JOSE MOMPEÁN-GONZALEZ and PILAR MOMPEÁN-GUILLAMÓN

## *Abstract*

*This article presents the results of an empirical study on the phenomenon of /r/-liaison (i.e., linking /r/ and intrusive /r/) in non-rhotic English from the perspective of usage-based Cognitive Linguistics. The study looks into sociolinguistic, phonetic and usage-based factors that condition variability in /r/-liaison through the analysis of news archives from the BBC World Service website (years 2004 and 2005). The paper argues that a thorough understanding of the phenomenon of /r/-liaison requires an analysis of the different aspects that condition its use and the use of empirical methods to study it.*

*Keywords:* /r/-liaison; intrusive /r/; linking /r/; usage-based phonology.

## **1. Introduction**

English accents are traditionally divided into two groups according to the distribution of the phoneme /r/: the rhotic and the non-rhotic. Rhotic accents are characterized by the pronunciation of the letter ⟨r⟩ as an r-sound in all positions in a morpheme. In non-rhotic accents, the situation is more complex. While the letter ⟨r⟩ is always pronounced before the

---

\* Correspondence address (both authors): Dept. of English Philology, University of Murcia, Santo Cristo, 1. CP: 30071. Murcia, Spain. E-mails: <mompean@um.es>; <p.mompean@um.es>. Acknowledgements: This paper has benefited from the invaluable feedback and assistance of Jill House (University College London), Dirk Geeraerts (University of Leuven), Ewa Dąbrowska (University of Sheffield), Antonio Maurandi (SAI, University of Murcia), John Meagher (University of Illinois Urbana-Champaign and University of Murcia) as well as two anonymous reviewers. The paper has also benefited from the financial support of Project 05817/PHCS/07, financed by the Fundación Séneca (Región de Murcia, Spain).

nucleus of the syllable (e.g., *read* /ri:d/, *dread* /dred/) in these accents, post-nuclearly an r-sound can only be found when the letter ⟨r⟩ is immediately followed by a vowel sound across morpheme boundaries, a phenomenon known as linking /r/. Accordingly, /r/ is not found (in non-rhotic English) in words such as *store* /stɔ:/, *stores* /stɔ:z/, or *stork* /stɔ:k/ since ⟨r⟩ is followed by a pause (*store*) or consonant sound (*stores*, *stork*). A ‘linking’ r-sound is pronounced, however, when the ⟨r⟩ is followed by a vowel sound across internal (e.g., *storing* /'stɔ:riŋ/) or external morpheme boundaries (e.g., *store it* /stɔ:r it/).

A related phenomenon to linking /r/ in non-rhotic accents is that of intrusive /r/, i.e., an epenthetic r-sound in intervocalic positions where, historically, there has never been an /r/ in the pronunciation of the word and present-day spelling does not contain the letter ⟨r⟩ (e.g., *the idea*[r]<sup>1</sup> of /ði aɪ'diə ɒv/, *I saw*[r] it /aɪ sɔ:r it/, etc.). Since intrusive /r/ is not justified by the spelling, this unetymological /r/ has traditionally been regarded as a vulgarism; therefore its use (but not that of linking /r/) is somehow stigmatized (Crystal 1984: 36; Jones 1956: 114; Knowles 1987: 134; Wells 1982: 224) and has been the focus of prescriptivist thought (at least in England) since the nineteenth century (see e.g., Muggleston 2003).

Despite their different degree of prestige as well as historical linguistic and orthographic differences, linking /r/ and intrusive /r/ are often considered to be essentially the same synchronic phenomenon (Heselwood 2006: 78; Wells and Colson 1971: 95). Both phenomena have the same distribution patterns, i.e., word-internally or across word boundaries, and only after certain non-high back monophthongs such as /ɑ:/ or /ɔ:/, central monophthongs such as /ə, ɜ:/ or centring diphthongs such as /ɪə, eə, uə/ (Collins and Mees 2003: 105; Lewis 1975: 37; Wells 1982: 226; Wells and Colson 1971: 94). In addition, both phenomena (often jointly referred to as /r/-liaison or /r/-sandhi) seem to have the same linguistic function: namely the avoidance of hiatus or lack of a consonant separating two vowels in separate syllables (Knowles 1987: 132).<sup>1</sup> Finally, both linking /r/ and intrusive /r/ are often given the same synchronic phonological interpretation. The most common analysis assumes that no underlying coda /r/s are present, and that /r/-alternations arise as a function of /r/-insertion following non-high vowels by analogy (e.g., Johansson 1973; McMahan 2000; Nespor and Vogel 1986; Vennemann 1972; Wells 1982). An alternative approach is to assume that underlying coda /r/s are

1. Here and henceforth the use of the symbol r between phonetic brackets (i.e., [r]) is for convenience and does not imply a voiced alveolar trill (its IPA value).

present for all words that show an /r/-alternation (see e.g., Donegan 1993; Gick 1999; Giegerich 1999; Harris 1994; Mohanan 1986).

Different descriptive studies have looked at /r/-liaison in many varieties of non-rhotic English. These studies have made it clear that although /r/-liaison is very common, its use is by no means universal or categorical in these accents. For instance, according to Brown (1988: 147), neither linking /r/ nor intrusive /r/ are generally found in South-East United States or in South-Africa. In addition, in some accents that exhibit linking /r/ the use of intrusive /r/ seems to be categorical, such as Norwich (see e.g., Trudgill 1974) or Yorkshire (e.g., Broadbent 1991). In contrast, variability in the use of /r/-liaison has been documented for Tyneside (Watt and Milroy 1999), New Zealand English (Hay and Sudbury 2005; Hay and Warren 2002) and Newcastle (Foulkes 1998). /r/-liaison has also been described as a variable phenomenon in the accent of England traditionally known as Received Pronunciation—henceforth RP—(e.g., Bauer 1984; Gimson 1980; Jones 1956; Lewis 1975; Wells 1982).

## **2. /r/-liaison in RP English: An empirical study**

Despite the existence of a great deal of theoretical work on /r/-liaison, few empirical studies so far have investigated the former's usage patterns in non-rhotic accents. An exception is, for instance, the study by Hay and Sudbury (2005), who analysed /r/-sandhi in the speech of New Zealanders born between 1860 and 1925 (as found in two different speech corpora recorded in the 1940s and between 1989 and 1995 respectively). Another exception is the earlier study by Hay and Warren (2002), in which reading data containing potential cases of intrusive /r/ were elicited from sixteen New Zealanders. For Newcastle English, Foulkes (1998) also analysed data from natural conversation and elicited reading passages. In the case of RP, the only studies available are those by Lewis (1975, 1977) and Bauer (1984). Lewis provided anecdotal comments on data collected by the author in the 1970s from BBC World Service new-readers. Bauer (1984) looked at the recordings of a story by 37 RP speakers (linguistics academics and students) between the years of 1949 and 1966, but his corpus only included eight potential cases of linking /r/ and two of intrusive /r/.

Given the few empirical studies available on /r/-liaison usage, the aim of the present study is to provide more empirical evidence and gain a better understanding of /r/-liaison in non-rhotic English, more specifically in the accentual variety known as RP. Cognitive Linguistics constitutes our theoretical standpoint insofar as Cognitive Linguistics is a usage-based approach to language (e.g., Geeraerts in preparation; Langacker 1999;

Tummers et al. 2005) with the implication that any language system is not simply a knowledge archive to be employed in language use, but rather is itself the product of actual language use. This standpoint has the consequence that if a usage-based model of grammar is taken seriously, one will have to study actual language use or usage events—the actual instantiations of the language system (Geeraerts in preparation: 17–18). The empirical study of actual language use requires, in turn, appropriate methodological tools and analytic methods that can tackle the phenomena investigated.

An empirical approach to /r/-liaison (corpus-based or experimental) is even more necessary, in our opinion, since variability in /r/-liaison usage raises the question of what factors condition such variability and to what extent. Unfortunately, the empirical evidence on the role of different variables affecting /r/-liaison usage is also scarce. In this respect, we believe that the factors that influence such variability can be grouped under three broad categories: sociolinguistic, phonetic, and usage-based.

The sociolinguistic aspect of /r/-liaison relates to usage patterns by specific groups of speakers given variables long studied by traditional sociolinguistics such as age, social class, gender, level of instruction, etc. Providing an account of the sociolinguistic factors that influence /r/-liaison usage and therefore language-internal variation is not only of interest to traditional sociolinguistics but also to Cognitive Linguistics, where recent discussions advocate that a genuinely cognitive approach should take into consideration cultural and social aspects of language and cognition (see e.g., Croft 2005; Geeraerts 2003, 2005; Hougaard 2005; Kristiansen 2003; Kristiansen and Dirven 2006). These discussions claim that as long as Cognitive Linguistics takes the claim that it is a usage-based approach to language, it should take into account the rich and complex patterns of intralingual variation. These patterns are far from descriptive studies carried out at the level of ‘a language’ that provide a picture of a supposedly homogeneous and idealized speech community.

As a case in point, it has been suggested that there might be gender differences in /r/-liaison usage in the sense that females would tend to use intrusive /r/ less than males given that intrusive /r/ is contrary to the overtly prestigious usage in the community and that females tend to use prestigious variants (Bauer 1984: 76; Coates 1993: 183; Dubois and Horvath 1999: 299; Labov 1990: 213). Social class can also be another variable affecting /r/-liaison usage patterns. Foulkes (1998), for instance, found in a study of Newcastle English that linking /r/ was used more by middle class speakers than by working class speakers in spontaneous, natural conversation, while intrusive /r/ was mainly used by working class

speakers. In reading passage style, however, the use of intrusive /r/ increased and was used significantly more by middle class than working class speakers.

A further sociolinguistic factor that may cause variability in /r/-liaison usage is speakers' level of instruction, particularly in relation to the formers' degree of literacy and awareness of spelling. In this respect, it is well-known that speakers who have a higher level of instruction tend to use more prestigious forms and adjust more to linguistic norms than those with a lower level of instruction (Moreno-Fernández 1998: 55). As far as /r/-liaison is concerned, intrusive /r/, but not linking /r/, has traditionally been regarded as a vulgarity by many speakers, its use carrying some degree of stigmatization (Crystal 1984: 36; Jones 1956: 114; Knowles 1987: 134; Wells 1982: 224). This view derives from speakers' knowledge of regular correspondences between spelling and pronunciation and how spelling should capture such associations. These views sometimes make speakers believe that silent letters found in the spelling of words should be pronounced. This explains why some speakers insert sounds in words such as *of*[t]en or *fa*[l]con, or why sounds not justified by the spelling are occasionally avoided (e.g., *the idea*[r] of). Clearly, the more familiar and aware of spelling a speaker is, the more s/he might try to avoid sounds that are not justified by the spelling system (e.g., intrusive /r/s) but make no such attempt with linking /r/s.

A second group of factors affecting variability in /r/-liaison usage is phonetic, i.e., articulatory, acoustic and auditory. Providing an account of such factors is essential in any phonological study since it is at present widely acknowledged that phonetics not only can but should provide explanations of the processes and phenomena that phonology deals with (see e.g., Blumstein 1991; Keating 1991; Kohler 1995; Ohala 1987, 1990; Pierrehumbert 2000, for related ideas). This is similar to ideas about experientialism and embodiment discussed for categories in the Cognitive Linguistics literature (see e.g., Lakoff 1987; Rohrer 2005) in which a central aspect is how the bodily apparatus shapes our linguistic categorization and conceptualization. In the same way, phonological categories and processes are also shaped by articulatory, acoustic and auditory factors (as well as by sociolinguistic and usage-based ones).

As a case in point, it has been claimed that the type of vowel phoneme at the end of the syllable that would make the link may have an influence on the use of intrusive /r/. More specifically, it has been claimed (e.g., Broadbent 1991: 301; Hay and Sudbury 2005; Hay and Warren 2002) that intrusive /r/ could be more frequent after lexical items ending in back vowels (e.g., /ɑ:/ as in *spa* /spɑ:/; /ɔ:/ as in *saw* /sɔ:/) than after those ending in a central monophthong (e.g., /ə/ as in *Emma* /'emə/) or

centring diphthong (e.g., /ɪə/, as in *idea* /aɪ'diə/) given that there are important acoustic similarities between rhotic approximants and back vowels such as the presence of a low third formant—henceforth F3—(see e.g., Ladefoged 2001; Ladefoged and Maddieson 1996; Peterson and Barney 1952; Stevens 1998: 288, 545).

Another phonetic factor affecting variability in the use of /r/-liaison could be the presence of /r/ at the beginning of the syllable that would make the link, as in *a* [r]oə[r] *of laughter* or *Victo*[r]iə[r] *and Albert*. In these cases, according to Jones (1956: 112–113), /r/ is not inserted as a rule. Windsor Lewis claims that the link is generally made (Lewis 1975: 38); however Brown (1988) says that /r/-liaison seems to occur “less readily” (p. 145). Irrespective of the frequency of /r/-liaison in these cases, the reason for avoiding the former when the syllable that would make the link begins with /r/ could be motivated by the lack of preference for similar or identical sounds in the same environment. In the history of a language, this often leads to dissimilation or the process by means of which nearby similar or identical consonant sounds become different over time. In fact, avoidance of two adjacent r-sounds has often been mentioned as a typical case of dissimilation (e.g., Ashby and Maidment 2005: 142) as in Latin *marmor* ['marmor] ('marble') developing into Spanish ['marmol] or rhotic English ['mɑ:ɹbɪl].

Finally, usage-based factors may also determine variability in /r/-liaison usage. These factors relate, for instance, to the type and token frequency of use of expressions, types of units, distributional patterns, etc., leading to the lexicalization or lexical entrenchment (Langacker 1987: 59) of the /r/ in certain groups of words, constructions, etc. Entrenchment is an important psychological phenomenon that refers to the cognitive routinization of linguistic units and structures on grounds of repetitive events in language use. This routinization makes any particular linguistic event turn into “a well-rehearsed routine that is easily elicited and reliably executed” (Langacker 1999: 93) that can be regarded in its own right.

As a case in point, it might be argued that /r/-liaison could be more frequent in expressions or constructions that have a high degree of entrenchment on the assumption that /r/-liaison is a hiatus-breaking strategy and that it will tend to become entrenched under one or more favourable circumstances. These circumstances could be the absence of a pause between the vowels involved, the presence of a single rhythmic and/or intonation unit pattern, a morphological connection between the two morphemes at the boundary of which the potential context is found, the high frequency of occurrence of the expression, etc. (for related ideas see Bybee 2001). If this is so, /r/-liaison should be found to be significantly more

frequent in the case of internal morpheme boundaries that have a bound morpheme or affix appearing before (prefix) or after (suffix) a free morpheme to form a lexeme (e.g., *inte[r]action*, *colou[r]ing*, *withdraw[r]al*, etc.) than at external morpheme boundaries across word boundaries where */r/-liaison* happens as the result of syntactic operations of two morphologically unrelated words (e.g., *the pape[r] under the table*) and/or with a pause across the potential */r/-liaison* context (e.g., *pape[r] . . . under the folder*). */r/-liaison* could also be more frequent in the case of internal morpheme boundaries when two (or more) morphemes are found as the constituents of a compound (e.g., *Fa[r] East*, *law[r]-and-order*, etc.). The underlying assumption here would be that, although the constituents of a compound can stand freely, these are strongly ‘glued’ together with the result that the compound is an independent symbolic unit with its own rhythmic structure and non-compositional meaning. Finally, */r/-liaison* could also be more frequent in the case of collocations (e.g., *fo[r] example*, *the idea[r] of*, etc.) on the assumption that it is the high frequency of such co-occurrence that may lead to entrenchment of the */r/*.

Given the many factors that seem to influence the existence of variability in */r/-liaison* usage, an empirical study was conducted in order to address some of the sociolinguistic, phonetic and usage-based factors that may influence that variability in non-rhotic English. More specifically, the research questions were:

- 1) are there significant differences in the use of linking */r/* and intrusive */r/*?
- 2) is avoidance of */r/-liaison* more common in female than male speakers?
- 3) is intrusive */r/* more common after central or after back vowels?
- 4) is */r/-liaison* generally avoided when the linking syllable begins with */r/*?
- 5) is */r/-liaison* more frequent in words with bound morphemes, compounds and collocations than in expressions with morphologically unrelated morpheme boundaries and no particularly high frequency of occurrence?

Based on the previous discussion, the hypotheses entertained in this study are that: a) there will be significant differences in the rate of use of linking */r/* and intrusive */r/* since the latter is somehow stigmatized; b) female speakers will use intrusive */r/* less often than male speakers since the former tend to use more prestigious forms (and intrusive */r/* is not prestigious); c) intrusive */r/* will be more common after back vowels than after central vowels since back vowels and post-alveolar approximants are phonetically similar in that they share a low F3; d) */r/-liaison* will be

generally avoided when the syllable that would make the link begins with /r/ due to dissimilation; e) /r/-liaison will be more frequent in words with bound morphemes, compounds and collocations than in expressions with morphologically unrelated word boundaries and no particularly high frequency of occurrence.

## 2.1. *Method*

2.1.1. *Data.* 307 texts from the news archives of the BBC | World Service | Learning English | Words in the News (see URL 1)—henceforth BBC WN—were used in order to gain the necessary evidence to answer the research questions formulated above. The full archive contains texts arranged by year starting from 1999 to the present moment. However, for this study, only the news archives corresponding to the years 2004 and 2005 were investigated. A full list of the 2004 and 2005 news archives can be seen at URL 2 and URL 3 respectively.

Among the characteristics of the texts analysed for this study are that they are free of charge and are readily available on the BBC WN website both as audio files and as written passages. This facilitates transcription, comparisons between spoken and written versions and subsequent analysis. Another characteristic of the texts is that most audio files have a relatively good (even studio-like) sound quality. A further feature of the corpus is that, almost without exception, individual texts are read by a single speaker, although texts occasionally include brief excerpts spoken by interviewees or sound (music, noises, etc.). Finally, most texts are preceded by a written introduction with no spoken counterpart and they are followed by specific words/expressions that the text features in bold (since it is a learning resource) with an explanation of their meaning and an additional audio file (typically recorded by a speaker other than the main newsreader). The written introduction, the featured words and the additional audio file were not analysed in this study.

The criteria for a given text to be analysed were that: a) the text should be read by an RP speaker; b) the name of the speaker should be identified; c) the text should be available as an audio file at the time the study was conducted; and d) the text should be read by a professional newsreader/correspondent (on a few occasions the texts are live recordings by the protagonists of news, interviewees, etc.). These requirements ruled out 38 texts read by non-RP speakers, three texts read by individuals other than newsreaders/correspondents such as businessmen, spokesmen, etc., two texts for which the identity of the RP speaker was not known, and one text technically unavailable at the time the study was conducted. Thus, out of the 307 texts of the 2004–2005 news archives,

only 263 were analysed. This represents 86% of the entire 2004–2005 corpus.

Using word-processing and audio software programs the total number of words and duration of texts were measured for each individual text and for the whole corpus analysed. As far as the number of words is concerned, the texts ranged from 98 to 288 words. The whole corpus under analysis contained 50,188 words. Word counts excluded the paragraphs corresponding to speech other than the newsreaders' (e.g., interviewees' appearances) as well as comparing the written and spoken versions of the same text for differences between both which might affect word counts and analyses. This comparison involved adding to or deleting words from the written version if discrepancies were found between the written and spoken versions, contracting words in the written version (e.g., *there is* → *there's*) if a contracted form was found in the spoken version, etc. In any case, a prominent feature of the texts investigated is that they were relatively short, the mean number of words per text being 163. Time was also measured. According to the authors' analysis, the whole corpus of relevant, identified RP speech lasts for around four hours and 20 minutes.

2.1.2. *Speakers.* 153 newsreaders employed by the BBC produced the texts from the 2004–2005 BBC WN archives investigated. As a preliminary step in the analysis, these speakers were identified as male or female by the quality of their voice and by their name, provided next to the written version of the text. In addition, each newsreader was identified as an RP/non-RP speaker based on a number of features such as the presence or absence of rhoticity and typical segmental inventories described for RP not only in classic descriptions of this accent (e.g., Gimson 1980) but also in recent descriptive updates of it (e.g., Collins and Mees 2003).<sup>2</sup>

It is important to bear in mind that no accent is a homogeneous invariant monolith and, consequently, for RP (like for most English accents), specific subvarieties can be (and have often been) identified (see e.g., Wells 1982: 279–280; Gimson 1980: 91, for discussions). However, no distinction is made in this paper between subtypes of RP in relation to the occurrence of */r/-liaison*, which anyway none of the discussions on

---

2. Regarding the non-RP speakers, most of them were rhotic. In addition, a few non-rhotic near-RP speakers were excluded from the study because their pronunciation had some conspicuously regional features such as the use of labiodental approximants for */r/* (e.g., Louisa Lim), typical of South-eastern England (see e.g., Foulkes and Docherty 2000), northern English characteristics such as the use of */ʊ/* instead of */ʌ/* in an otherwise RP-like accent (e.g. Adam Easton), southern-hemisphere features like open start of */er/* (e.g. Kylie Morris) or American English influence in flapped */t, d/* and retroflex */r/*'s (e.g., Gina Wilkinson).

subvarieties of RP mentions as a distinguishing factor across those subvarieties. In any case, all speakers described as RP speakers in this study have in common that, despite minor phonetic and/or phonological differences in their segmental inventories, their accent is one which lacks obvious local or regional features (excluding the fact that they can be typically associated with England). The ‘non-localizability’ of RP speakers has in fact been mentioned as a sociolinguistic feature of RP. In this respect, the present-day RP speakers’ accent has also been referred to with the name “non-regional pronunciation (abbreviated to NRP)” (Collins and Mees 2003: 4). This accent allows the mainstream “present-day variation to be heard from educated middle and younger generation speakers in England who have a pronunciation which cannot be pinned down to a specific area” (Collins and Mees 2003: 4). Moreover, non-localizability is what apparently differentiates RP from accents with similar phonological/phonetic features such as the recently popularized Estuary English, typical of the south-east of England (see e.g., Wells 1994a).

Collins and Mees’ quotation above also suggests a further sociolinguistic criterion of RP: it is an accent spoken by educated people, mostly of affluent status in the social scale. In this respect, the group of speakers analysed for this study can be considered as a relatively homogeneous group as far as social class is concerned (if the latter is defined with reference to the speakers’ professional activity) with most of them explicitly identified as ‘correspondents’ on the BBC WN website.

Finally, it should be mentioned that the accent under investigation in this study can also be referred to as BBC English, a term used in the *Cambridge English Pronouncing dictionary* (Roach et al. 2006: v) to replace what the authors consider “the archaic name *Received Pronunciation*”. BBC English is defined as “the pronunciation of professional speakers employed by the BBC as newsreaders and announcers on BBC1 and BBC2 television, the World Service and BBC Radio 3 and 4 ...” (Roach et al. 2006: v). The editors of the dictionary also acknowledge that there are individual differences between speakers employed by the BBC and that a number of broadcasters have Scottish, Welsh or Irish accents, but they describe the accent that is “typical of broadcasters with an English accent” (Roach et al. 2006: v).

Table 1 shows the number of identified and unidentified female and male speakers (RP and non-RP) from the 2004–2005 BBC WN archives.

As table 1 reveals, the identity of 129 RP speakers out of a potential figure of 131 RP speakers was verified. The figures also reveal a disproportion in the gender groups since female speakers are less numerous than male ones (29% vs. 71% respectively). These figures are similar if identified non-RP speakers are also taken into account (32% vs. 68%). It

Table 1. *Number of identified and unidentified female and male speakers (RP and non-RP) in the BBC WN 2004–2005 news archives*

|        |        | <i>Females</i> |              | <i>Males</i> |              | <i>Females and Males</i> |              |
|--------|--------|----------------|--------------|--------------|--------------|--------------------------|--------------|
|        |        | Identified     | Unidentified | Identified   | Unidentified | Identified               | Unidentified |
| Accent | RP     | 38             | 1            | 91           | 1            | 129                      | 2            |
|        | Non-RP | 10             | 0            | 12           | 0            | 22                       | 0            |

Table 2. *No. of texts read by the same newsreader and no. of speakers (males, females, males and females combined) in the BBC WN 2004–2005 news corpus analysed who read that number of texts*

|                  |         | <i>No. of texts read by the same newsreader</i> |     |    |    |    |    |    |    |      |    |    |
|------------------|---------|---|-----|----|----|----|----|----|----|------|----|----|
|                  |         | 1   | 2   | 3  | 4  | 5  | 6  | 7  | 8  | 9–11 | 12 | 13 |
| No. of speakers  | Males   | 51  | 16  | 7  | 6  | 2  | 5  | 1  | 1  | —    | 1  | 1  |
|                  | Females | 27  | 8   | 2  | —  | 1  | —  | —  | —  | —    | —  | —  |
|                  | Total   | 78  | 24  | 9  | 6  | 3  | 5  | 1  | 1  | —    | 1  | 1  |
| Rate of speakers | Total   | 60%   | 19% | 7% | 5% | 2% | 4% | 1% | 1% | —    | 1% | 1% |

should also be noted that the number of texts analysed (263) is not the same as the overall number of RP newsreaders (129). This is due to the fact that some newsreaders broadcast more than one text. However, Table 2 below shows that, although there are a few speakers who produce three or more texts, 60% of newsreaders broadcast only one text, the percentage of speakers who produce one or two texts is 79%, and one to three texts 86%. In sum, most speakers produce very few texts, the mean in the corpus being 2 texts per speaker.

2.1.3. *Procedure.* Only the texts read by identified RP speakers and available at the time this study was conducted (263) were analysed. For each text, its written version was copied from the website and pasted onto a Word document where the potential /r/-liaison contexts were identified. The identification process involved reading the texts for potential environments and marking them.

Regarding the issue of what a potential context of /r/-liaison in our corpus could be, the following criteria were followed. First, as Wells remarks (1982: 224; 1994b: 198), the weakening to schwa of the final RP diphthong in words such as *window*, *pillow*, etc. (i.e., /əʊ/→/ə/), typical of some regional accents, is resisted by RP, where an unstressed diphthong is the norm. Therefore, words ending in /əʊ/ were not considered as examples of potential intrusive /r/.

Second, words ending in a non-high vowel phoneme with orthographic <r> (for linking /r/) or without orthographic <r> (for intrusive /r/) and followed by the personal pronouns *he* and *him*, the possessive adjectives/pronouns *his* and *her*, the reflexive pronouns *himself* and *herself* and three forms of the verb *to have* (i.e., *have*, *has*, and *had*) were considered as potential contexts only if elision of /h/ had previously occurred. As Knowles notes (Knowles 1987: 133–134), /h/-dropping in these cases produces a context of adjoining vowels and thus the necessary conditions for potential /r/-liaison.

Third, it has often been pointed out that linking /r/ is categorical in word-internal position in polymorphemic words containing orthographic *r* when one of the morphemes is a prefix or a suffix (e.g., *hyper+inflation* or *ignor+ing*). For example, García-Lecumberri and Maidment (2000: 34) claim that the word *bearing* is to be transcribed /'beərɪŋ/, “*NEVER* /'beərɪŋ/” (capitals and italics in the original). Moreover, no author in the specialized literature has ever suggested that /r/ might not be categorical in these cases. Despite this, cases of so-called linking /r/ in polymorphemic words containing orthographic *r* in word-internal position were identified and listened to with the result that no variability whatsoever was found in the corpus studied for such items. Consequently such cases

were not considered as potential contexts of linking /r/ in this study and are excluded in the figures referring to variability in linking /r/ usage in this paper (but see results and discussion of the fifth research question below for more information). In contrast, compounds (e.g., *Far East*) were considered as potential cases of linking /r/ in this study since it has sometimes been claimed that word-internal linking /r/ is not categorical in them (Jones 1960: 196).

With these considerations in mind, searches for typical orthographic contexts, and specific lexical items mentioned in the literature or found through the sound search tool of EPD (Roach et al. 2006) were performed using the word search tools available in Microsoft Word. These searches were carried out to minimize the possibility of overlooking or missing any potential contexts. In the case of linking /r/ across word boundaries, the combinations <r> and <re> were checked with a space after them, which leads to word boundaries, or with a punctuation mark after <r> or <re> (e.g., . , : ; - ? ! ' ' ' ' ) ] \* / #). Other marginal spellings such as <rh> (e.g., *catarrh*) were also checked in a same way.

Regarding intrusive /r/, both word-internal and word-boundary contexts were considered. In the case of word-internal positions, combinations of the letter <a> plus another vowel letter, with or without a hyphen in-between, were searched for (e.g., <ai, a-i> as in *concertinaing*, *magentaish*; <ae, a-e> as in *Kafkaesque*, *salsa-evening*, etc.) as well as the digraph <aw> followed by another vowel letter (e.g., <awa> as in *withdrawal*; <awi> as in *drawing*; <awe> as in *awe-inspiring*) or <awy> (e.g., *strawy*). Across word-boundaries, the spellings <a> and <aw> were checked with a space after them or a punctuation mark. In addition, place names from Old English ending in orthographic <urgh> (e.g., *Oxburgh* (/ˈɒksbrə/)), were search for as well as words ending in <ah>, <eh>, and <agh> since many words of non-Saxon origin (especially Arabic, Hebrew, Celtic and Persian) are often spelled in this way (e.g., *Abdullah*, *Jehovah*, *Methuselah*, *Rafah*, *Omagh*, *Nineveh*, *Ayatollah*, *shah*, etc.) as well as various words such as *yeah* or *pariah*. Other marginal spellings, mainly of French origin, were also inspected such as <oi>, as in *moi*, <ois> as in *François*, *bourgeois*, <eu> as in *milieu* or *Fontainebleu*, <as> as in *Degas*, or <at> as in *nougat*.

For both linking /r/ and intrusive /r/, the search options <r->, <re->, <aw-> or <awe-> were sufficient to detect hyphenated compounds (e.g., *near-obsession*). Open form compounds (e.g., *Far East*) or closed form compounds (e.g., *firearm*) were identified visually while reading/listening to the texts. Moreover, in the case of orthographic <r> or <re> followed by *he*, *him*, *his*, *her*, *himself*, *herself*, *have*, *has* or *had*, the corresponding texts were listened to in order to decide whether the /h/ had been

dropped. If so, the expression was considered as a potential context of /r/-liaison. This inspection revealed, as Bauer has suggested (1984: 77), that it is probably the exception rather than the rule (at least in broadcast speech) to delete /h/ in these cases.

Next, texts were analysed auditorily for the occurrence/non-occurrence of intrusive /r/ in the potential contexts previously identified. Decisions were generally quickly reached as to the appearance or not of an instance of /r/-liaison. In most cases, the phonetic identity of the /r/ is a post-alveolar approximant [ɹ] (as was the case of Bauer's 1984 study), with a few instances of slightly retracted place of articulation, making it similar to a retroflex approximant [ɻ] (but not quite) and a few instances of voiced alveolar taps [ɾ] in very conservative RP speakers. Realizations of /r/ as a labio-dental approximant [ʋ] or as a truly retroflex variant [ɻ] were generally considered as a feature of a non-RP accent and speakers who produced them were not analysed.

Although decisions were generally quickly made regarding the presence/absence of an instance of /r/-liaison, auditory analysis of the data was occasionally considered insufficient. In these cases, spectrographic analysis of the relevant sound files was carried out using the *Speech Filing System* (SFS), a free program for speech research developed at UCL (see URL 4). Two types of cases needed particular attention. The first was uncertainty about the presence of /r/. In this respect, it has long been noted that /r/ is characterized by a low F3 which, although it is not true for all rhotics in the languages of the world (see e.g., Ladefoged and Maddieson 1996: 244; Lindau 1985), is a well-justified specification at least for American English retroflex and British English post-alveolar approximants (*ibid.*). Thus visual inspection of the corresponding spectrogram(s) and presence/absence of a prominent low F3 was considered the criterion for determining the presence/absence of an r-sound. The second case involved analysis of sequences of /r/+ə/+potential /r/-liaison, for which some authors (e.g., Brown 1988: 145; Lewis 1977: 30–31) claim that speakers often omit the schwa and prolong the /r/, making it syllabic. In these cases, presence of a long [ɹ] was considered as an instance of /r/-liaison, previous to the elision of schwa.

Online Appendix A ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) exemplifies the way the texts were dealt with for the analysis. This appendix contains tables that include titles of the newscasts next to a text reference number arbitrarily assigned to each text. Also, the gender of the speaker is specified as well as the relevance of the text for the study (i.e., whether it was analysed or not). Also, Online Appendix B ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) shows, for female and male speakers separately, some of the potential contexts of linking /r/ with a

specification of whether an r-sound was used or not in each case. Online Appendix C ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) does the same for intrusive /r/ as well as indicating the identity of the vowel at the end of the linking syllable. Limitations of space prevent the full set of data to be published in the paper version of this study. Thus, all appendices ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) are published online and can also be obtained from the authors of the paper on request.

## 2.2. *Results and discussion*

As a preliminary clarification regarding the analysis of the data used for this study, it should be mentioned that the speakers studied did not have a chance to produce the same number or linking and/or intrusive /r/s. In fact, each of the 129 speakers could produce any number of linking and intrusive /r/s so the units of analysis should be speakers ( $N = 129$ ) and not individual instances of /r/ production. Given this, to investigate the research questions formulated in this study, data from all 129 speakers regarding the relevant research questions are presented although it was with the data computed separately for each speaker that statistical tests were carried out. The computation of data separately for individual speakers was carried out by adding, for each individual speaker, the number of potential cases of the different conditions investigated (linking /r/ in general, linking /r/ in the expressions *for example* and *the/a number of*, intrusive /r/, linking syllables with or without /r/ in the onset for intrusive /r/ and linking /r/, linking syllables ending in back or central vowels for intrusive /r/, linking /r/ in compounds and non-compounds, etc.).

With regard to the first research question in this study (i.e., whether there are differences in the use of linking /r/ and intrusive /r/), the rates of use of these two phenomena were calculated for the speakers as a single group. The results obtained show that the absence of linking /r/ is very frequent given that the percentage of linking /r/ in the corpus (58%) is less than two thirds of all potential cases (570 out of 984 potential cases—see table 3 below). This relatively low percentage of linking /r/ could be due to the careful, speech-conscious style of the newscasts. In fact, it has been claimed that “use of linking/intrusive /r/ is a feature of fluent colloquial style, and is not so common in careful declarative style” (Brown, 1988: 145),<sup>3</sup> which may be due to the fact that the latter

---

3. That linking /r/ is very common in fluent colloquial style seems to be the case, at least, in some southern British English accents. In a study of English analysing informal conversation of 10 adolescents of British White ethnicity from the Fens and 22 from the London Borough of Tower Hamlets, for instance, linking /r/ was found to occur at a rate of 96% and 95% respectively.

may be delivered at a lower rate of speech and “slower speech seems to produce fewer [r]s” (McMahon 2000: 249–250; see also Heselwood 2006: 92–93). In this respect, studies with scripted speech read by newscasters have shown similar tendencies in other languages, with speech being closer to canonical forms than casual speech (see e.g., Torstensson 2004).<sup>4</sup>

The low percentage of linking /r/ found in this study could also be partly due to the tendency among some speakers to eliminate intrusive /r/ from their speech. Having been told it is slovenly or vulgar many speakers might eliminate intrusive /r/ only “at the expense of eliminating linking /r/’s too” (Wells and Colson 1971: 95) and this tendency might perhaps be more marked in a speech-conscious, careful declarative style typical of scripted news-reading.

The results obtained also show that intrusive /r/ is not a very frequent phenomenon in broadcast RP, and is perhaps (like linking /r/) less frequent than in colloquial, unscripted speech (Brown 1988: 145). In a corpus of over 50,000 words and out of 165 potential cases, there were only 52 actual instances of intrusive /r/ (32%), less than a third of all potential cases (see also Table 3). This finding is similar to percentages of intrusive /r/ occurrence found in other non-rhotic accents such as Tyneside, with a rate of around 20% (e.g., Watt and Milroy 1999) or New Zealand English, with a rate of around 30% (Hay and Warren 2002).

In order to find out whether there are statistically significant differences between rates of occurrence of linking /r/ and intrusive /r/, the relative rate of occurrence of linking /r/ and intrusive /r/ was computed separately for each speaker. For this analysis, only speakers who had at least one potential case of intrusive /r/ were considered and of those, only the speakers who had three or more potential cases of linking /r/, which ruled out 78 speakers. Thus, the performance of only 61 speakers was analysed (see also Table 3). This performance amounted to 678 potential and 397 actual cases (59%) in the case of linking /r/, the mean being  $0.589 \pm 0.245$  (*SD*). In the case of intrusive /r/, the 61 speakers produced 148 potential and 49 actual instances (33%), with a mean of  $0.314 \pm 0.419$  (*SD*). A Wilcoxon signed ranks test was applied, which showed that the differences in percentage between actual linking /r/ and intrusive /r/ in the corpus are not random ( $Z = -3.827$ ,  $p = 0.001$ ) for  $N = 61$  subjects.

---

4. It should also be pointed out that the absence of linking /r/ (or intrusive /r/) from some of the potential contexts does not necessarily mean that no hiatus-breaker is employed. In fact, it has been pointed out that when linking (but also intrusive) /r/ are not used, sometimes speakers use a glottal stop (e.g., Allerton 2000; Foulkes 1998) and such is the case in the present study.

Table 3. Number of potential /r/-liaison contexts (linking /r/ and intrusive /r/), actual instances and rates of use for 129 and 61 speakers

| No. of speakers | <i>Linking /r/</i> |              |                      | <i>Intrusive /r/</i> |              |                      |
|-----------------|--------------------|--------------|----------------------|----------------------|--------------|----------------------|
|                 | Potential cases    | Actual cases | Rate of actual cases | Potential cases      | Actual cases | Rate of actual cases |
| <i>N</i> = 129  | 984                | 570          | 58%                  | 165                  | 52           | 32%                  |
| <i>N</i> = 61   | 678                | 397          | 59%                  | 148                  | 49           | 33%                  |

The hypothesis entertained for the first research question investigated in this study (i.e., that there would be significant differences between the rates of use of linking /r/ and intrusive /r/) is therefore confirmed when the rates of usage by speakers as a group are analysed.

An alternative approach to the one presented above regarding the issue of whether there are significant differences between the rates of usage of linking /r/ and intrusive /r/ is to look at differences between speakers or in the same speaker across different texts produced by the individual. With the data available, though, this approach has some problems and cannot be satisfactorily explored. To start with, there are very few speakers (21%) who produce more than one or two texts so the data produced by a given newsreader should be analysed jointly, independent of the number of texts produced by that newsreader. This makes even more sense given that no text was produced by two or more speakers so individual performances cannot be compared in this respect. A further complication is that the number of potential cases of intrusive /r/ for individual speakers is very low. Table 4 shows the different numbers of potential contexts of intrusive /r/ in the corpus studied and the number of speakers who had a given number of potential contexts in their data. The table also shows the number of speakers who never used an intrusive /r/ and the number of speakers who always used it in their potential cases (table 5 shows the same information in relation to linking /r/). As the data shown in the tables reveal, only three out of the 129 speakers studied had more than four potential cases of intrusive /r/. Four cases is already a very small number to draw any reliable conclusions regarding cross-speaker differences so patterns of cross-speaker variation for intrusive /r/ cannot be studied satisfactorily. More substantial evidence can be obtained, however, looking at patterns of cross-speaker variation for linking /r/. In this respect, two aspects are of particular interest: whether variability is categorical (i.e., there is always variability in speakers) and whether there are significant differences across speakers in their production of linking /r/.

Table 4. *No. of potential cases of intrusive /r/ in the corpus studied, no. of speakers who had a given number of potential contexts in their data and number of speakers who never used intrusive /r/ or who always used it in their potential cases*

| No. of potential contexts | <i>Females</i> |                               |                                | <i>Males</i> |                               |                                |
|---------------------------|----------------|-------------------------------|--------------------------------|--------------|-------------------------------|--------------------------------|
|                           | Speakers       | Never an intrusive /r/-linker | Always an intrusive /r/-linker | Speakers     | Never an intrusive /r/-linker | Always an intrusive /r/-linker |
| 0                         | 21             | —                             | —                              | 37           | —                             | —                              |
| 1                         | 11             | 9                             | 2                              | 23           | 17                            | 6                              |
| 2                         | 2              | 1                             | —                              | 12           | 6                             | 4                              |
| 3                         | 2              | 1                             | —                              | 10           | 6                             | 2                              |
| 4                         | 2              | 1                             | —                              | 6            | 2                             | 1                              |
| 8                         | —              | —                             | —                              | 1            | —                             | —                              |
| 13                        | —              | —                             | —                              | 1            | —                             | —                              |
| 15                        | —              | —                             | —                              | 1            | —                             | —                              |

Table 5. *No. of potential cases of linking /r/ in the corpus studied, no. of speakers who had a given number of potential contexts in their data and number of speakers who never used linking /r/ or who always used it in their potential cases*

| No. of potential contexts | <i>Females</i> |                     |                      | <i>Males</i> |                     |                      |
|---------------------------|----------------|---------------------|----------------------|--------------|---------------------|----------------------|
|                           | Speakers       | Never an /r/-linker | Always an /r/-linker | Speakers     | Never an /r/-linker | Always an /r/-linker |
| 0                         | 1              | —                   | —                    | 1            | —                   | —                    |
| 1                         | 3              | 1                   | 2                    | 9            | 3                   | 6                    |
| 2                         | 3              | 2                   | —                    | 7            | 1                   | 3                    |
| 3                         | 3              | 1                   | 1                    | 14           | 1                   | 5                    |
| 4                         | 5              | 1                   | 1                    | 7            | 1                   | 1                    |
| 5                         | 8              | —                   | —                    | 13           | 1                   | 3                    |
| 6                         | 2              | —                   | —                    | 8            | —                   | 2                    |
| 7                         | 4              | —                   | 1                    | 1            | —                   | —                    |
| 8                         | 2              | —                   | —                    | 5            | —                   | —                    |
| 9                         | 1              | —                   | —                    | 3            | —                   | 1                    |
| 10–20                     | 6              | —                   | —                    | 16           | —                   | —                    |
| 21–30                     | —              | —                   | —                    | 4            | —                   | —                    |
| +30                       | —              | —                   | —                    | 3            | —                   | —                    |

Regarding whether variability is categorical, the data shown in table 4 lead us to think that, although linking /r/ may be a notable field of cross-speaker variation (Lewis 1975: 39), its variation is not as great as to include speakers who never use it at all or who always use it. On the one hand, there appear to be no speakers who never use linking /r/. As table

4 shows, there are no speakers with six or more potential contexts that never use linking /r/. With fewer contexts, the presence of no /r/-linkers might be due to the few potential cases available per speaker. However, an interesting fact is that, out of the 21 speakers (males and females combined) who had five potential cases for linking /r/, there was only one newsreader, Paul Keller, who did not produce any linking /r/ at all.<sup>5</sup> This speaker does not contribute any more texts in the BBC WN in other years so he cannot be studied any further. With four or fewer contexts, reliable conclusions cannot be drawn since the contexts are too few. However, out of 53 speakers, only 11 (21%) do not produce any linking /r/ in their data, although it is relatively easy, using search facilities within the BBC WN site or on the Internet generally, to find further newscasts/interviews by most of those 11 speakers with contexts where they do produce some linking /r/s.<sup>6</sup> In sum, the data obtained seem to support the claim that it may be rare to find BBC newsreaders who never produce any linking /r/ in their speech. This in turn seems to put into question Jones's observation in the middle of the twentieth century that "a great many Southern people" did not use linking /r/ "at all" (Jones 1956: 113) or that there appeared then to be "an *increasing tendency*, especially among younger people, not to use linking /r/ at all"—emphasis added—(Jones 1960: 197). These supposed tendencies, if they exist at all presently, do not seem to apply to the set of BBC newsreaders investigated in this study.

On the other hand, as table 4 also shows, it seems that there are no speakers who always produce linking /r/. There is one speaker in the corpus, Andrew North, with nine potential and nine actual linking /r/s. However, this speaker sometimes fails to produce linking /r/ in texts from the BBC NW in years other than 2004 or 2005,<sup>7</sup> and so is the case with one out of seven speakers, Naomi Grimley, with seven potential cases and seven linking /r/s, who fails to make some links in other recordings available on the Internet.<sup>8</sup> Two out of ten speakers with six

---

5. "A test to be British" (31 October 2005).

6. E.g., Frances Harrison: "Refugee crisis in Sri Lanka" (20 June 2001) four out of six contexts: "about a quarter are living ...", "then there are at least half a million ...", "... but it's clear a huge proportion", and "waiting for an end to ...". Mark Dummett: "Bush visits Hyderabad" (3 March 2006), one out of two potential contexts: "the home to an *ever-expanding* middle class ..."

7. E.g., In "More severe petrol shortages in Iraq" (11 June 2007).

8. E.g., "Tony Blair: Past, Present and Future" at [http://www.kcrw.com/news/programs/tp/tp070510tony\\_blair\\_past\\_pres](http://www.kcrw.com/news/programs/tp/tp070510tony_blair_past_pres). Naomi Grimley misses three out of seven potential cases of linking /r/: "a mixture of humility and apology ...", "reforms that Tony Blair instituted ...", and "... without having a proper elector ...".

potential contexts, James Helm and Richard Black, also fail to make links in other recordings of theirs,<sup>9</sup> and so is the case of the three speakers, Tom Gibb, Rodney Smith and Andrew Marr, with five potential cases.<sup>10</sup> With four or fewer potential cases, reliable conclusions cannot again be drawn but out of 53 speakers, only 19 (36%) always produced a linking /r/ in their contexts but this may be due, as discussed above, to the few potential cases available. In fact, using again search facilities within the BBC WN site or on the Internet generally, many of these speakers can be found to miss some of their potential contexts.<sup>11</sup> In sum, these data again seem to support the claim that it may be rare, if not impossible, to find BBC newsreaders who always produce linking /r/ in all potential contexts.

Regarding the issue of to what extent linking /r/ varies across speakers, only those speakers with at least six potential cases in the data contributed to the corpus were analysed. This restriction was established since we believe that to better test differences across subjects in percentage of /r/-liaison usage, the more potential cases that each speaker has in his/her speech, the more representative the rate of actual usage will be for that speaker. It was also chosen because with speakers with five or fewer potential cases discarded, there was still a sizeable number of speakers (60, 47% of all the population studied) that could be analysed. With this criterion in mind, a boxplot was generated, which can be seen in Figure 1 below. The boxplot depicts graphically a five-number summary comprising the smallest observation, lower quartile, median, upper quartile and largest observation. Therefore the boxplot shows the whole range of values for a given group of speakers as well as the interquartile range (i.e., the difference between the first and the third quartile, which includes about 50% of the data). The boxplot also identifies outliers, i.e., single observations at an appreciable distance from most others.

- 
9. E.g., James Helm: “Atlantic rowers” (9 January 2006) failure to produce linking /r/ in three out of five potential contexts: “was floating beside their upturned boat...”, “. . . at the end of November and were heading . . .”, and “. . . lost the use of the rudder on their small rowing . . .”. Richard Black: “IRA: the mechanism of decommissioning” (24 October 2001), failure to produce one out of two potential contexts: “. . . Cambodia, Mali, El Salvador and many other nations . . .”.
  10. E.g., Rodney Smith: a) “Gorbachev enters software piracy argument” (7 February 2007) in which he fails to produce linking /r/ in the two potential cases. Tom Gibb: “IMF blocks loan to Argentina” (7 December 2001) in which the speaker failed to produce one of the three potential contexts: “. . . currency or adopt the dollar as the official currency . . .”.
  11. E.g., Russel Padmore at [http://www.bbc.co.uk/worldservice/learningenglish/radio/specials/1549\\_weekender\\_extra/page40.shtml](http://www.bbc.co.uk/worldservice/learningenglish/radio/specials/1549_weekender_extra/page40.shtml)

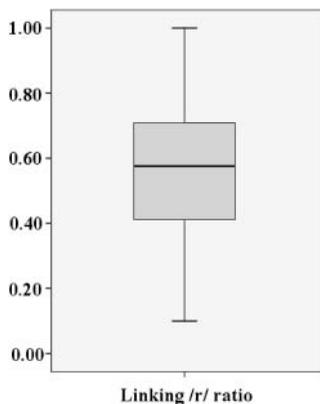


Figure 1. Box plot for Linking /r/ ratio

As Figure 1 shows, the frequency of linking /r/ usage is displayed on the  $y$ -axis and the distribution for the linking /r/ ratio on the  $x$ -axis. Also, the data obtained from the statistical analysis show that the highest ratio is 1 and the lowest is 0.1. The median, represented as a black horizontal line inside the vertical rectangle on the linking /r/ ratio distribution is 0.58 and the mean 0.57, a figure which coincides with the percentage of linking /r/ usage in the analysis of all the data of speakers as a group (see above). The range is 0.90, the variance 0.49 and standard deviation 0.221. As Figure 1 also shows, there are no outliers or observations numerically distant from the rest of the data. Moreover, a histogram showing the number of speakers who produced a given rate of linking /r/ against the linking /r/ ratio is shown in Figure 2. This figure shows a normal distribution. However, to test the normality of the distribution a Shapiro–Wilk  $W$ -test, which tests the null hypothesis that a sample  $x_1 \dots x_n$  derives from a normally distributed population, was carried out. The Shapiro–Wilk  $W$ -test showed that the data do not differ significantly from a normal distribution ( $W = 0.978$ ,  $p = 0.351$ ). The conclusion that can be drawn from the test performed is that, from the data that are available, there is no evidence that there may be two different groups or populations in the rate of linking /r/ usage.

The second research question investigated in this study refers to whether avoidance of /r/-liaison is more common in female than in male speakers. Considering the performance of the 129 speakers of the unabridged corpus, out of 216 potential cases of linking /r/ in the females' data, 120 linking /r/'s were produced (56%). In the males' data, out of 768 potential cases, 450 linking /r/'s were used (58%). The analysis of

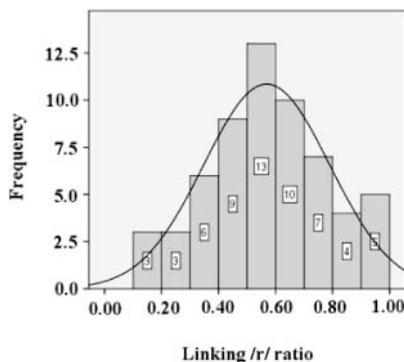


Figure 2. Histogram showing the number of speakers who produced a given rate of linking /r/ against the linking /r/ ratio

the data also reveals that out of 29 potential cases of intrusive /r/, females produced only seven (24%) whereas men produced 45 out of 136 potential cases (33%). These results are summarized in Table 6 below.

With the aim of finding out whether there were statistically significant differences between male and female speakers in /r/-liaison usage, the relative rates of occurrence of linking /r/ and intrusive /r/ were computed separately for each speaker. For this analysis again, only speakers who had at least one potential case of intrusive /r/ were considered and of those, only the speakers who had three or more potential cases of linking /r/, which ruled out 78 speakers so the performance of only 61 speakers was analysed, with 46 males and 15 females in total (see also Table 6).

Table 6. Number of potential /r/-liaison contexts, actual instances and percentages of use: females, males, and both groups combined for 129 subjects and for 61 subjects

| No. of speakers | Gender group          | Linking /r/     |              |                      | Intrusive /r/   |              |                      |
|-----------------|-----------------------|-----------------|--------------|----------------------|-----------------|--------------|----------------------|
|                 |                       | Potential cases | Actual cases | Rate of actual cases | Potential cases | Actual cases | Rate of actual cases |
| <i>N</i> = 129  | <i>N</i> = 38 Females | 216             | 120          | 56%                  | 29              | 7            | 24%                  |
|                 | <i>N</i> = 91 Males   | 768             | 450          | 58%                  | 136             | 45           | 33%                  |
| <i>N</i> = 61   | <i>N</i> = 15 Females | 125             | 70           | 56%                  | 27              | 7            | 26%                  |
|                 | <i>N</i> = 46 Males   | 553             | 327          | 59%                  | 121             | 42           | 35%                  |

In the case of 15 females, there were 125 potential and 70 actual instances (56%) of linking /r/ and 27 potential and seven actual instances (26%) of intrusive /r/ (see also Online Appendix D <http://dx.doi.org/>

10.1515/COGL.2009.031\_suppl-1), with a mean of  $0.579 \pm 0.205$  (*SD*) for linking /r/ compared to a mean of  $0.244 \pm 0.382$  (*SD*) for intrusive /r/. For males, there were 553 potential and 327 actual instances (59%) of linking /r/, with a mean of  $0.592 \pm 0.259$  (*SD*) compared to 121 potential and 42 actual instances (35%) of intrusive /r/, with a mean of  $0.337 \pm 0.431$  (*SD*). Two Wilcoxon signed ranks tests were applied which showed that, with two gender groups of speakers, the differences between linking /r/ and intrusive /r/ are also statistically significant for females ( $Z = -2.274$ ,  $p = 0.023$ ), for  $N = 15$  speakers, and males ( $Z = -3.094$ ,  $p = 0.002$ ) for  $N = 46$  speakers. Moreover, two Mann-Whitney tests were performed to find out whether there were significant differences between males and females in the production of linking /r/ and intrusive /r/. The results of these tests show that there are no differences between males and females in the production of linking /r/ ( $U = 332.5$ ,  $p = 0.833$ ) and in the production of intrusive /r/ ( $U = 301.0$ ,  $p = 0.410$ ). Thus, our initial hypothesis that female speakers would produce fewer intrusive /r/s than male speakers is not confirmed.

As far as intrusive /r/ is concerned, two explanations seem to be readily at hand for the observed patterns. On the one hand, taking for granted that the phenomenon is equally stigmatized for both groups, intrusive /r/ might tend to be less frequent in the female population but the male newreaders analysed might be, as a group, more conscious of the stigmatized nature of the phenomenon, conditioned by the formal context of news-reading. If this were so, males might make more efforts to avoid intrusive /r/ with the result that the differences between the rates of occurrence of /r/ between both gender groups are not statistically significant. On the other hand, and also taking for granted that intrusive /r/ is equally stigmatized for both groups, it might be claimed that once females have access to a labour market traditionally monopolized by men, the former tend to imitate men in different ways, including males' speech habits (Coates 1993: 10). If this applies to /r/-liaison, it would mean that women tend to use intrusive /r/ as often as men in imitation of the latter.

The third research question investigated in this study was whether intrusive /r/ is more common after central or after back vowels. In the unabridged corpus analysed, the phonemic contexts after which an intrusive /r/ could have been inserted after a central vowel are: a) /ə/ preceded by a consonant (C+/ə/); b) stressed diphthongal /ɪə/ and /eə/; c) unstressed, disyllabic /i.ə/; d) disyllabic /i:.ə/ (stressed on /i:/); and e) /ɜ:/ . As Table 7 below shows, most potential cases involve C+/ə/ (101), and disyllabic /i.ə/ (39). The other phonemic contexts are exemplified by few potential cases. This is within what can be expected a priori, since intrusive /r/ has often been claimed to be extremely rare after (final) central

Table 7. *Potential cases of intrusive /r/, actual occurrences and rate of actual cases by previous vowel context depending on the preceding vowel context (central/back vowel)*

| Preceding vowel context |          | Intrusive /r/ after central/back vowels |              |                      |
|-------------------------|----------|---|--------------|----------------------|
|                         |          | Potential cases                         | Actual cases | Rate of actual cases |
| Central vowels          | C+/ə/    | 101                                     | 26           | 26%                  |
|                         | /ɪə/     | 5                                       | 2            | 40%                  |
|                         | /i.ə/    | 39                                      | 11           | 28%                  |
|                         | /i:ə/    | 2                                       | 0            | 0%                   |
|                         | /eə/     | 1                                       | 0            | 0%                   |
|                         | Combined | 147                                     | 39           | 27%                  |
| Back vowels             | /ɔ:/     | 18                                      | 13           | 72%                  |

vowels other than schwa. This is particularly true of /eə/, /ʊə/, or /ɜ:/, simply because native words containing those vowels at morpheme boundaries and not followed by historical /r/ are almost non-existent (Brown 1988: 150; Collins and Mees 2003: 105; Wells and Colson 1971: 95).<sup>12</sup>

The potential back-vowel contexts after which an intrusive /r/ could have been inserted were 18 (far fewer than those involving central vowels) and they all involve the vowel /ɔ:/.<sup>13</sup> It is interesting to note that no potential contexts for intrusive /r/ after /ɑ:/ were found in the corpus despite the fact that the literature typically discusses /ɔ:/ and /ɑ:/ together, offering as many examples of the latter as of the former and suggesting perhaps that intrusive /r/ is as common after /ɑ:/ as after /ɔ:/.

Table 7 shows that the rate of intrusive /r/ after central vowels, combining the individual results of the different final phonemic contexts identified in the corpus is 27%, with little difference between female and male newsreaders (23% and 28% respectively). This contrasts with the high percentage of intrusive /r/ after the back vowel /ɔ:/ (i.e., 72%). However, in order to find out whether there are significant differences between the rate of occurrence of intrusive /r/ after back vowels and central vowels, the occurrence of intrusive /r/ in these cases was computed separately for each speaker. For this analysis, only speakers who had at least one potential case of intrusive /r/ after central and back vowels were considered.

12. In “Arafat aides leaves on Paris trip” (8<sup>th</sup> Nov 2004) by James Reynolds. For Andrew Marr see “Yoko Ono on the Andrew Marr Show” on Youtube.com (added 2 October 2007) at [http://www.youtube.com/watch?v=ZMwGy429L\\_U](http://www.youtube.com/watch?v=ZMwGy429L_U).

13. Real instances of intrusive /r/ in these 18 potential cases were two examples of *saw[r]* *a*, four instances of *withdraw[r]al*, and one instance of the expressions *draw[r] up*, *law[r] in*, *claw[r] its way*, *draw[r] in*, *Pattern Law[r] Office*, *law[r] enforcement*, and a *flaw[r] in*.

For the nine speakers who complied with the criterion, there were 27 potential and 11 (41%) actual instances of intrusive /r/ after central vowels, with a mean of  $0.553 \pm 0.443$  (*SD*) while there were 12 potential and nine (75%) actual cases of intrusive /r/ after central vowels, with a mean of  $0.667 \pm 0.500$  (*SD*). Table 8 below summarizes these data and Online Appendix E ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) also shows the data per speaker. A Wilcoxon signed ranks test was applied, which shows that the differences in percentage between back vowels before intrusive /r/ and central vowels also before intrusive /r/ is not statistically significant ( $Z = -0.542$ ,  $p = 0.688$ ) for  $N = 9$  subjects.

Table 8. *Potential cases of intrusive /r/, actual occurrences and rate of actual cases by previous vowel context depending on the preceding vowel context (central/back vowel) for 9 speakers (and also for 129 speakers)*

| No. of speakers | Preceding vowel context | Intrusive /r/ after central/back vowels |              |                      |
|-----------------|-------------------------|---|--------------|----------------------|
|                 |                         | Potential cases                         | Actual cases | Rate of actual cases |
| $N = 9$         | Central vowels          | 27                                      | 11           | 41%                  |
|                 | Back vowels             | 12                                      | 9            | 75%                  |
| $N = 129$       | Central vowels          | 147                                     | 39           | 27%                  |
|                 | Back vowels             | 18                                      | 13           | 72%                  |

The findings obtained do not confirm the hypothesis entertained in this study, i.e., that intrusive /r/ would be significantly more common after back vowels than after central vowels since back vowels and post-alveolar approximants are phonetically similar in that they share a low F3. One reason for the unexpected result could be the relatively few number of speakers (nine) and data investigated (39 potential instances). Thus this issue needs exploring further with a greater number of potential cases, which can be obtained by enlarging our corpus. Another explanation of a non-phonetic nature for the results obtained could be the alleged greater stigmatization of intrusive /r/ after back vowels. In this respect, the specialized literature (e.g., Fox 1978: 74; García-Lecumberri and Maidment 2000: 34; Lewis 1977: 30) has often claimed that stigmatization is greater after back vowels because the lexical items in which the latter are found as potential contexts for intrusive /r/ are less frequent than those items in which a potential context for intrusive /r/ is found after a central vowel. This claim is based on the assumption that speakers' attention would be drawn to items with a back vowel because of their rare character, and speakers would try to avoid intrusive /r/ as much as they do with

intrusive /r/ after central vowels (e.g., Brown 1988: 150; Collins and Mees 2003: 105; Crystal 1984: 42; Gimson 1980: 39, 302–303; Lewis 1975: 40–41).

The fourth research question investigated in this paper addressed the issue of whether /r/-liaison is generally avoided or not when the syllable that would make the link begins with /r/.

In the unabridged corpus, there are 27 potential cases of intrusive /r/ in which the syllable likely to make the link begins with /r/ and nine instances in which the link is actually made (33%) while in the corpus of potential intrusive /r/ excluding the cases with /r/ in the onset the potential cases were 138 and the actual cases 43 (31%). In addition, there are 15 potential cases of linking /r/ in which the last syllable making the potential link begins with /r/, with five (33%) actual instances. The rate of linking /r/ excluding potential cases of linking /r/ preceded by /r/ is 58% (969 potential and 565 actual cases). Table 9 below shows, for 129 speakers, the number and rate of use of intrusive /r/ and linking /r/ in syllables that do and do not begin with /r/. These data seem to refute the claim that the link is “generally made” when the syllable begins with /r/ (Lewis 1975: 38) and provides more support to the less strong claim that linking /r/ is “not as a rule inserted” in this context (Jones 1956: 112).

Table 9. *Number of potential intrusive /r/ and linking /r/ cases with /r/ in the onset and /r/-less onset, actual instances and rates of use for 129 speakers*

| N° of speakers | Phenomenon    | Subcorpus        | Cases and rate  |                  |                       |
|----------------|---------------|------------------|-----------------|------------------|-----------------------|
|                |               |                  | Potential cases | Actual instances | % of actual instances |
| N = 129        | Intrusive /r/ | /r/ in the onset | 27              | 9                | 33%                   |
|                |               | /r/-less onset   | 138             | 43               | 31%                   |
|                | Linking /r/   | /r/ in the onset | 15              | 5                | 33%                   |
|                |               | /r/-less onset   | 969             | 565              | 58%                   |

As far as the hypothesis considered in this study is concerned (i.e., that /r/-liaison will be generally avoided when the syllable that would make the link begins with /r/), the results obtained seem to provide some support to the hypothesis, particularly in the case of linking /r/ since the percentage of use in the corpus (58%) is higher than in the subcorpus of 15 potential cases when the syllable that would make the link begins with /r/ (33%). However, to find out whether the differences in percentages are statistically significant, the relative rate of occurrence of linking /r/ and intrusive /r/ when the syllable with the potential link began with

Table 10. Number of potential intrusive /r/ cases with /r/ in the onset and /r/-less onset, actual instances and rates of use for 129 and 18 speakers

| N° of speakers | Phenomenon    | Subcorpus        | Cases and rate  |                  |                       |
|----------------|---------------|------------------|-----------------|------------------|-----------------------|
|                |               |                  | Potential cases | Actual instances | % of actual instances |
| N = 18         | Intrusive /r/ | /r/ in the onset | 25              | 9                | 36%                   |
|                |               | /r/-less onset   | 49              | 21               | 43%                   |
| N = 13         | Linking /r/   | /r/ in the onset | 15              | 5                | 33%                   |
|                |               | /r/-less onset   | 287             | 172              | 60%                   |

/r/ or did not was computed separately for each speaker. Table 10 summarizes these data, Online Appendix F ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) shows the potential and actual cases for both intrusive /r/ and linking /r/ separately and Online Appendix G ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) presents the data per speaker for linking /r/ and intrusive /r/ also separately. With this restriction, only 13 speakers had one or more potential cases of linking /r/ with /r/ in the onset of the linking syllable and one or more potential cases of linking /r/ when the onset of the linking syllable did not contain /r/. These also had 15 potential and five (33%) actual instances of linking /r/ in items with /r/ in the onset, with a mean of  $0.345 \pm 0.455$  (*SD*) but 287 potential and 172 (60%) actual instances in /r/-less onsets, with a mean of  $0.559 \pm 0.149$  (*SD*).

In the case of intrusive /r/, only 18 speakers had one or more potential cases of intrusive /r/ preceded by /r/ in the onset and one or more potential cases of intrusive /r/ without /r/ in the onset. These had 25 potential and nine (36%) actual instances when there is an /r/ in the onset, with a mean of  $0.361 \pm 0.465$  (*SD*) and 49 potential and 21 actual instances (43%) in /r/-less onsets, with a mean of  $0.481 \pm 0.412$  (*SD*). A Wilcoxon signed ranks test was performed. This text shows that the difference between the rates of intrusive /r/ when the linking syllable begins with /r/ and when it does not is not statistically significant either ( $Z = 0.178$ ,  $p = 0.906$ ). A second Wilcoxon signed ranks tests was performed, revealing too that the differences between the rates of linking /r/ when the linking syllable begins with /r/ and when it does not is not statistically significant ( $Z = -0.1758$ ,  $p = 0.800$ ) for  $N = 14$  speakers.

These results seem to suggest that, if there is any tendency to avoid r-sounds in the nearby contexts, the effect is not appreciated in the corpus analysed for intrusive /r/. This may be due to the relatively few potential intrusive /r/ cases when the syllable contains /r/ in the onset in the

unabridged corpus (i.e., 27) and also in the performance of the 18 speakers analysed (i.e., 25). If there is any tendency to avoid r-links when an r-sound is at the beginning of the syllable that would make the link, this tendency is again not clearly revealed for linking /r/ in the data analysed. This might again be due to the few potential linking /r/ cases in syllables that begin with /r/ in the unabridged corpus and in the 13 speakers investigated (i.e., 15). Ideally, the archive should then be expanded to include more potential cases of /r/-liaison with /r/ in the onset to obtain more conclusive evidence on the effect of the presence of /r/ at the beginning of the syllable that would make the link.

The fifth research question looked at whether /r/-liaison is more frequent in words with bound morphemes, compounds and collocations than in expressions with morphologically unrelated morpheme boundaries and no evident high frequency of occurrence.

With regards to words with bound morphemes, the assumption is that since these morphemes cannot stand freely (only exceptionally in speech), symbolic units that contain them are learned as single units since there is no pause in between the morphemes and they have a single rhythmic structure as well as a coherent communicative intention in a given communicative context (Tomasello 2003: 63). Thus, symbolic units are normally experienced as single units, even if they can be decomposed into their constituent elements; thus, /r/ has a higher likelihood of becoming entrenched if the /r/ was not there historically (or remaining entrenched if it originally was).

In our corpus, there were 150 cases of potential linking /r/ across two morpheme boundaries (when one of them is a bound morpheme), 11 in prefixes or 139 suffixes. The prefixes involved were: *over-*, *under-*, *super-*, *hyper-* and *inter-*; the suffixes were: *-able*, *-ably*, *-al(ly)*, *-ent*, *-er*, *-est*, *-ing*, *-ity*, *-ic*, *-ish*, *-ist*, *-ism*, and *-ial*. In all cases, a linking /r/ was produced. This confirms the hypothesis that there is no variability in prefixed or affixed words where /r/ has historical antecedence (e.g., Heselwood 2006: 81; Lecumberri and Maidment 2000: 34; Wells 1982: 224; Wells and Colson 1971: 94). Historically, this can be taken as a resistance of /r/ in those positions in the overall tendency to drop /r/ in current non-rhotic English. Synchronically, it may be claimed that analogical processes are operating to maintain /r/ as a categorical hiatus-breaking strategy word-medially in words with bound morphemes that did not exist when /r/ started to be dropped in current non-rhotic English accents.

As far as word-internal intrusive /r/ is concerned, there are only five potential cases in the corpus analysed in words with bound morphemes (see Online Appendix H [http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)). All the cases involve the suffixes *-al* (four instances of the word

*withdrawal*) and *-ing* (one instance of the word *drawing*). Intrusive /r/ was found in all the instances of *withdrawal* but it did not occur in the word *drawing*. Thus, despite the few cases involved, intrusive /r/ reaches 80% of all actual usage word-internally.

The data obtained (summarized in table 11 below) show then that use of linking /r/ is categorical across morpheme boundaries in words with bound morphemes (100% of all cases) and apparently very frequent in the case of intrusive /r/ (80%). The rates obtained are much higher than the rate of linking /r/ in the corpus not including the bound morphemes items in the linking /r/ data (58%) or in the intrusive /r/ data (30%).

Table 11. *Potential cases of linking and intrusive /r/, actual instances and rate of actual instances for items with bound morphemes and for the corpus without bound morphemes*

| Subcorpus                      | Linking /r/     |              |                      | Intrusive /r/   |              |                      |
|--------------------------------|-----------------|--------------|----------------------|-----------------|--------------|----------------------|
|                                | Potential cases | Actual cases | Rate of actual cases | Potential cases | Actual cases | Rate of actual cases |
| Items with bound morphemes     | 150             | 150          | 100%                 | 5               | 4            | 80%                  |
| Corpus without bound morphemes | 984             | 570          | 58%                  | 160             | 48           | 30%                  |

As the data obtained reveal, intrusive /r/ does not seem to be a particularly rare or infrequent phenomenon word-internally position in lexical items with bound morphemes. However, the few items of potential intrusive /r/ in word-internal position do not permit any strong claims to be made about this issue. The tentative conclusion that can be drawn, though, is that intrusive /r/ may be gaining ground at word-internal morpheme boundaries. This would be in line with recent comments in the literature that “many speakers of present-day RP pronounce /r/ in this sort of word” (García-Lecumberri and Maidment 2000: 34). Moreover, this is not necessarily in contrast with opinions such as that there may be “more sentiment against intrusive /r/ word-internally than across word boundaries” (Wells 1982: 225) even if it is “sometimes frowned upon when it occurs within words” (García-Lecumberri and Maidment 2000: 34). As is well known, speakers’ attitudes to language do not necessarily match usage and the former are often biased by their sense of linguistic norm.

Regarding compounds, it could be expected that /r/-liaison might be more frequent in these kind of lexical items than across word boundaries

with no morphological relationship. The assumption is that, although the constituents of a compound can stand freely, each compound has its own rhythmic structure and its meaning is typically non-compositional; thus a compound is an independent symbolic unit. As in the case of words with bound morphemes, the morphemes of a compound are more strongly ‘glued’ together than any simple sequence of two words with a potential linking /r/ at their boundary. Not surprisingly, Jones (1960: 196) claims that /r/ is “... *generally* inserted in compound words” (emphasis added). This comment suggests that linking /r/ in compounds may not be a categorical phenomenon, but, rather that it is very frequent.

Our unabridged corpus for 129 speakers contained 26 compounds with potential linking /r/ (see also Online Appendix I [http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) out of which 20 actual instances of linking /r/ were found (77%); see Table 12 below. Since the rate of occurrence of linking /r/ compounds is higher than the rate of linking /r/ in the corpus excluding compounds (i.e., 57%), this seems to confirm the hypothesis that linking /r/ is frequent in compounds and Jones’s (1960) claim. However, in order to find out whether there were statistically significant differences in the rate of occurrence of linking /r/ in compounds as compared with the corpus without compounds, the relative rate of occurrence of linking /r/ in compounds and in non-compounds was computed separately for each speaker. For this analysis, only speakers who had at least one potential case of linking /r/ in a compound as well as in non-compounds were considered. As table 12 also shows, 20 speakers met this criterion, producing 26 compounds that were potential instances of linking /r/ and 20 actual cases (77%), with a mean of  $0.775 \pm 0.413$  (*SD*). These 20 speakers also produced 202 potential and 106 (52%) actual instances of linking /r/ in non-compounds, with a mean of

Table 12. *Number of potential linking /r/ cases and actual instances as well as rate of actual instances in compounds and in the corpus without compounds for 129 and 20 speakers*

| N° of speakers | Subcorpus                | Linking /r/            |                         |                          |
|----------------|--------------------------|------------------------|-------------------------|--------------------------|
|                |                          | No. of potential cases | No. of actual instances | Rate of actual instances |
| N = 129        | Compounds                | 26                     | 20                      | 77%                      |
|                | Corpus without compounds | 958                    | 550                     | 57%                      |
| N = 20         | Compounds                | 26                     | 20                      | 77%                      |
|                | Corpus without compounds | 202                    | 106                     | 52%                      |

0.494 ± 0.296 (*SD*); see also Online Appendix J ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) for details. A Wilcoxon signed ranks test was performed, showing that the difference in percentage between actual linking /r/ in compounds and in non-compounds is statistically significant ( $Z = -2.250$ ,  $p = 0.023$ ) for  $N = 20$  subjects.

The results obtained then may indicate that the morphemes of a compound are more strongly ‘glued’ together than any simple sequence of two words with a potential linking /r/ at their boundary so linking /r/ may be somehow entrenched. However the non-categorical nature of linking /r/ in compounds (unlike in polymorphemic words with bound morphemes) suggests that certain word boundary phenomena operate not only at word boundaries but also at word-internal morpheme boundaries in compounds.

A study of intrusive /r/ could not be carried out with compounds since only one case was found in the corpus in which no intrusive /r/ was found.<sup>14</sup> A relatively common compound containing a potential intrusive /r/ mentioned in the literature is *law and order* but no instance of it was found in the corpus analysed. An inspection of the BBC WN corpus since 1999 until December 2007 yielded only two instances, in both of which an intrusive /r/ was pronounced;<sup>15</sup> however the data are too limited to draw any conclusions.

Finally, the frequency of /r/-liaison in collocations or sequences of words which co-occur more often than would be expected by chance could also be expected to be higher than in the rest of the corpus. The assumption is that it is precisely the frequency of such co-occurrence that may lead to entrenchment of the /r/. In this respect, occasional comments in the relevant literature seem to suggest that speakers use /r/-liaison more frequently in common expressions (Jones 1956: 113).

As far as linking /r/ is concerned, Jones (1956: 113) mentions *after all* and *for example* as common expressions where linking /r/ is regularly found. In the corpus analysed, an expressions such as *after all* could not be satisfactorily studied since it appears only on one occasion. The same happens with another expression such as ‘a matter of’, found only twice. After inspecting the corpus, only four expressions seemed to be relatively frequent to merit analysis: *for example/instance* and *the/a number of* (see Online Appendix K [http://dx.doi.org/10.1515/COGL.2009.031\\_](http://dx.doi.org/10.1515/COGL.2009.031_)

---

14. *euthanasia-in-cambodia-dot-com*, in “Webmaster faces legal action” (11 November 2005) by Guy de Launey

15. In “Bangladeshi Budget” (13 June 2003) and “Darfur Crisis Getting Worse” (25 June 2007).

supp-1).<sup>16</sup> Regarding the expressions *for example/instance* (analysed together), 13 instances were found in the corpus and in all of them a linking /r/ was found (100%). In the expressions *the/a number of* (also analysed together), found on 24 occasions, linking /r/ was used in 23 cases (92%). These percentages (see tables 13 and 14 below) suggest that linking /r/ may be highly frequent in collocations (categorical or nearly categorical in some cases) and more frequent than the average percentage of linking /r/ in the corpus without the collocations studied (56%).

In order to find out whether there were statistically significant differences between linking /r/ in *for example/instance* and in the corpus without these expressions, the relative rates of occurrences for these two conditions were computed separately for each speaker and the same was done for the collocations *the/a number of*. For these analyses, only speakers who had at least one potential case of linking /r/ in *for example/instance* or *the/a number of* as well as other potential cases of linking /r/ were considered. 12 speakers met this criterion for the expressions ‘for example/instance’ and 19 for the expressions *the/a number of*. Regarding *for example/instance* (see also Table 13), the 12 speakers produced the same number of potential and actual cases (i.e., 13) as in the unabridged

Table 13. *Number of potential cases and actual cases of linking /r/ as well as rate of actual cases in the expressions for example/instance and in the corpus without these expressions for 129 and 12 speakers*

| No. of speakers | Subcorpus                               | Linking /r/     |              |                      |
|-----------------|---|-----------------|--------------|----------------------|
|                 |   | Potential cases | Actual cases | Rate of actual cases |
| N = 129         | <i>for example/instance</i>             | 13              | 13           | 100%                 |
|                 | Corpus without <i>for ex. /instance</i> | 971             | 557          | 57%                  |
| N = 12          | <i>for example/instance</i>             | 13              | 13           | 100%                 |
|                 | Corpus without <i>for ex. /instance</i> | 248             | 153          | 62%                  |

16. Out of all the occurrences of the lexeme ‘number’ (suffixed occurrences included), i.e., 36 instances, the collocation “number/s of” occurs on 25 occasions (“number of” 24, “numbers of” 1), i.e., 69%. In the BNC, out of 61,012 instances of the lexeme ‘number’ (also suffixed occurrences included), the collocation “number/s of” occurs 38,294 times (“number of” 34734, “numbers of”, 3560), i.e., 63%. Also, in our corpus, out of all the occurrences of the lexeme “example” (suffixed occurrences included), i.e., 11, the collocation “for example” occurs 9 times, i.e., 82%. In the BNC, out of 43,086 occurrences of the lexeme ‘example’, the common expression ‘for example’ occurs 23,755 times, i.e., 55%.

corpus, with a mean of  $1.00 \pm 0.00$  (*SD*), but 248 potential and 153 (62%) actual instances of the subcorpus without these expressions, with a mean of  $0.615 \pm 0.185$  (*SD*); see also Online Appendix L ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)). A Wilcoxon signed ranks test shows that the difference in rate between actual linking /r/ in *for example/instance* and the subcorpus excluding these expressions is statistically significant ( $Z = -2.936$ ,  $p = 0.001$ ) for  $N = 12$  subjects.

In the case of *the/a number of* (see also Table 14), the 19 speakers produced 24 potential and 21 (88%) actual instances, with a mean of  $0.894 \pm 0.267$  (*SD*), compared to the 298 potential and 182 (61%) actual instances of linking /r/ excluding *the/a number of*, with a mean of  $0.611 \pm 0.259$  (*SD*); see also Online Appendix L ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)). A Wilcoxon test shows that the difference is also statistically significant between actual linking /r/ in *the/a number of* and the subcorpus without the former ( $Z = -2.767$ ,  $p = 0.03$ ) for  $N = 19$  speakers.

Table 14. *Number of potential cases and actual cases of linking /r/ as well as rate of actual cases in the expressions the/a number of and in the corpus without these expressions for 129 and 19 speakers*

| No. of speakers | Subcorpus                             | Linking /r/     |              |                      |
|-----------------|---------------------------------------|-----------------|--------------|----------------------|
|                 |                                       | Potential cases | Actual cases | Rate of actual cases |
| $N = 129$       | <i>the/a number of</i>                | 24              | 21           | 88%                  |
|                 | Corpus without <i>the/a number of</i> | 960             | 549          | 57%                  |
| $N = 19$        | <i>the/a number of</i>                | 24              | 21           | 88%                  |
|                 | Corpus without <i>the/a number of</i> | 298             | 182          | 61%                  |

Combining the results of the four expressions, 26 speakers met the criterion of producing at least one of the collocations studied as well as potential case of linking /r/ that was not a collocation. As table 15 shows (and Online Appendix L ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) for details), these speakers produced 37 collocations that were potential instances of linking /r/ and 34 actual cases (92%), with a mean of  $0.989 \pm 0.207$  (*SD*) as well as 366 potential and 215 (59%) actual instances of linking /r/ excluding the collocations, with a mean of  $0.599 \pm 0.245$  (*SD*). Again, a Wilcoxon test shows that the difference is statistically significant between actual linking /r/ in the four collocations combined and the subcorpus excluding those collocations ( $Z = -3.784$ ,  $p = 0.001$ ) for  $N = 26$  subjects.

Table 15. Number of potential cases and actual cases of linking /r/ as well as rate of actual cases in the expressions for example/instance and the/a number of (combined) and in the corpus without these expressions for 129 and 26 speakers

| No. of speakers | Subcorpus                       | Linking /r/     |              |                      |
|-----------------|---------------------------------|-----------------|--------------|----------------------|
|                 |                                 | Potential cases | Actual cases | Rate of actual cases |
| N = 129         | Collocations combined           | 37              | 34           | 92%                  |
|                 | Corpus without the collocations | 947             | 536          | 57%                  |
| N = 26          | Collocations combined           | 37              | 34           | 92%                  |
|                 | Corpus without the collocations | 366             | 215          | 59%                  |

Unfortunately, no data can be offered regarding collocations with potential intrusive /r/. One likely candidate often discussed in the relevant literature is *the idea of* but the latter could not be studied satisfactorily since only one instance of it was found in the corpus. However, another inspection of the BBC WN corpus from the years 1999 to 2007 shows that the expression occurs 11 times and that intrusive /r/ is found on seven occasions (64%).<sup>17</sup> This figure is lower than that of expressions such as *for example/instance* or *the/a number of*, but double the percentage of intrusive /r/ in the corpus (32%).

The results obtained can be interpreted, as explained in section II, as the categorical or nearly categorical lexicalization or entrenchment of /r/-liaison in words with bound morphemes in the case of intrusive /r/ and in the case of linking /r/ in collocations and compounds. Related to this, it could be claimed that the differences initially observed in the rates of usage of linking /r/ and intrusive /r/ in the corpus (58% and 32% respectively) could be due to the high entrenchment of a small number of items with a high lexical frequency (Ewa Dąbrowska, p.c.). Thus, in order to find out whether there are also significant differences between linking /r/ and intrusive /r/ excluding the influence of compounds and collocations, the relative rate of occurrence of the relative rate of occurrence of linking /r/ and intrusive /r/ was computed separately for each

17. Intrusive /r/ found in “Iran Petrol Rationing” (29 June 2007), “Arab Firm Delays Us Ports Deal” (24 February 2006), “Royal Mail Prize For Coming To Work” (2 May 2005), “Conference On Iraq’s Future” (29 April 2003), “The Future of Cyprus”—twice examples—(10 March 2003) and “Controversial Russian Stamps” (1 May 2002). No intrusive /r/ found in “United States ‘Career Brides’” (11 July 2003), “Anti-Gravity Device Could Change Air” (29 July 2002), “Background To The Russian Elections” (23 March 2000) and “Rush For Hong Kong Rail Shares” (28 September 2000).

speaker who had potential cases of both linking /r/ and intrusive /r/ excluding compounds and collocations (combined). Thus, the performance of 60 speakers was analysed (only one speaker, James Westhead, was excluded from a similar comparison made for research question one above). These produced 632 potential and 358 (57%) actual instances of linking /r/, with a mean of  $0.575 \pm 0.254$  (*SD*) and 147 potential and 49 (33%) actual instances of intrusive /r/, with a mean of  $0.320 \pm 0.420$  (*SD*); see table 16 and Online Appendix M ([http://dx.doi.org/10.1515/COGL.2009.031\\_supp-1](http://dx.doi.org/10.1515/COGL.2009.031_supp-1)) for details. A Wilcoxon signed ranks test was applied, which showed that the differences in percentage between actual linking /r/ and intrusive /r/ in the corpus are not random ( $Z = -3.432$ ,  $p = 0.001$ ) for  $N = 60$  subjects. Thus there seem to be significant differences between intrusive /r/ and linking /r/ independently of the role of entrenchment in a small number of high lexical frequency items.

Table 16. *Number of potential /r/-liaison contexts (linking /r/ and intrusive /r/), actual instances and rates of use for 129, 61 speakers (those with at least three potential cases of linking /r/) and 60 speakers (those with at least three potential cases of linking /r/ excluding collocations and compounds)*

| No. of speakers | Linking /r/     |              |                      | Intrusive /r/   |              |                      |
|-----------------|-----------------|--------------|----------------------|-----------------|--------------|----------------------|
|                 | Potential cases | Actual cases | Rate of actual cases | Potential cases | Actual cases | Rate of actual cases |
| $N = 60$        | 632             | 358          | 57%                  | 147             | 49           | 33%                  |

### 3. Conclusion

The phenomenon of /r/-liaison has long been the focus of academic interest, although detailed empirical studies of its use in non-rhotic English have so far been infrequent. Given this scarcity of empirical evidence, the present paper has tried to look into the usage patterns of /r/-liaison in RP through the analysis of a corpus of news archives from the BBC WN site.

The study was conducted from the theoretical standpoint of usage-based Cognitive Linguistics (e.g., Geeraerts in preparation; Langacker 1999; Tummers et al. 2005) with the methodological implication that it is necessary to study actual language use or usage events (the actual instantiations of the language system) with appropriate methodological and analytic tools. Thus, at a methodological level, the study reported in this paper can be considered as a contribution to usage-based Cognitive Linguistics in that it looks at real language data that instantiate a given

linguistic phenomenon as the inevitable source of knowledge about the language system.

The data-driven approach to /r/-liaison presented in this study has looked at different aspects that determine variability in this phenomenon. We believe that a proper understanding of the phenomenon of /r/-liaison and its variability requires the investigation of different factors since language itself emerges from the interaction of varied inherent and experiential factors of different sorts, i.e., biological, behavioural, psychological, social, cultural and communicative (Langacker 1991: 1).

With regard to the sociolinguistic component of /r/-liaison, the present study has produced three main findings. The first is that the difference between the rate of occurrence of both linking /r/ and intrusive /r/ is statistically significant, which reveals that, although the two processes appear to be phonetically identical, other factors may be causing the difference. In this respect, the differences found seem to indicate that intrusive /r/ remains a somehow stigmatized phenomenon and that people's ideas about the correctness of intrusive /r/ are affecting the linguistic usage patterns in their speech. In fact, Wells (1994b) suggests that "... in spite of its prevalence in RP (and other non-rhotic accents), intrusive /r/ does remain to some extent the object of overt stigmatization" (p. 201). This interpretation could be supported by the fact that linking /r/ occurs at a rate (58%) possibly lower than would have been expected in colloquial speech. This lower incidence of linking /r/ could be explained as a result of the tendency to avoid intrusive /r/ from their speech at the expense of eliminating linking /r/s too (Wells and Colson 1971: 95), a possible feature of scripted broadcast speech.

The fact that intrusive /r/ may be stigmatized can be related, as stated in the introduction, to the sociolinguistic variable 'level of instruction', particularly in relation to literacy levels. In this respect, speakers who have a higher level of instruction tend to use more prestigious forms and adjust more to linguistic norms than those with a lower level of instruction. Intrusive /r/ could be a non-prestigious phenomenon because, to the educated, literate and somehow spelling-conscious, it may clearly appear to be a vulgarism (see e.g., Crystal 1984: 36; Jones 1956: 114; Knowles 1987: 134; Wells 1982: 224), or as something to be avoided. Since RP is defined sociolinguistically as a social accent and is often associated with education, the set of speakers studied, professionals recording scripted news for a prestigious public broadcasting company may well be considered educated and perhaps spelling-conscious.

It should be pointed out that, if lack of prestige is a characteristic of intrusive /r/ for the RP speakers studied, this view ultimately derives from speakers' knowledge and/or beliefs about the relationship between

spelling and pronunciation and how spelling should capture such associations. Related to this, speakers' conceptions about language usage (e.g., what they think is correct/incorrect, what they know about spelling, etc.) can be thought of as 'folk theories' around which cognitive psychology has long claimed and shown that concepts are organized (see e.g., Lin and Murphy 1997; Murphy 1993; Rips 1989, 1995) and which other cognitive scientists have discussed under various names for different types of knowledge structures.<sup>18</sup> These folk theories are sets of beliefs which, according to Rips (1995), may be sketchy, naïve, stereotyped, or incorrect, and as a result are "a host of mental explanations rather than a complete, organized, scientific account" (Murphy and Medin 1985: 312). The existence of folk theories in phonology have already been discussed in relation to the assignment of allophones to phoneme categories (Mompeán 2004). If folk phonological theories also exist for /r/-liaison and these include ideas about how correct or incorrect intrusive /r/ is, these theories might explain the fact that, despite being essentially the same phenomenon phonetically (as explained above), there are significant differences in the rate of use of linking /r/ and intrusive /r/.

The second main sociolinguistic finding of this study is that there do not seem to be any speakers who always use linking /r/ or who never use it. Therefore variability in /r/ liaison usage seems to be inherently categorical within speakers. The data also do not provide any evidence that there may be two or more different groups or populations in the rate of linking /r/ usage. Unfortunately, the few potential cases of intrusive /r/ per speaker do not allow us to arrive at any reliable conclusions about the variability of its use within speakers. In any case, the results obtained with regard to cross-speaker variation contribute to our understanding of language-internal variation regarding /r/-liaison. This kind of variation was typically ignored in Saussurean and Chomskyan linguistics, with their dichotomies of *langue* and *parole* or *competence* and *performance* and their focus on analysis carried out at the level of 'a language' providing a picture of a supposedly homogeneous and idealized speech community. However, there is growing interest in Cognitive Linguistics in the rich and complex patterns of intralingual variation (see e.g., Geeraerts in preparation; Kristiansen and Dirven, 2006) and, as long as Cognitive Linguistics takes the claim that it is a usage-based approach to language, this study exemplifies the need to take into account the rich and complex patterns of intralingual variation.

---

18. These include 'frames' (e.g., Fillmore 1985), 'scripts' (e.g., Schank and Abelson 1977), 'mental models' (e.g., Barsalou 1992) or 'idealized cognitive models' (Lakoff 1987).

Finally, the third finding of a sociolinguistic nature is that there are no statistically significant differences between males and females in /r/-liaison usage. This finding was obtained after looking at speakers as two gender groups but also considering individual differences within each group. Taking for granted that the phenomenon is equally stigmatized for both groups, the finding obtained might be due to the fact that intrusive /r/ might tend to be less frequent in the female population but the male newsreaders analysed might be, as a group, more conscious of the stigmatized nature of the phenomenon, conditioned by the formal context of news-reading. If this were so, they might make more efforts to avoid it with the result that the differences between the rates of occurrence of /r/ are not statistically significant. An alternative explanation for the results obtained could be that once females have access to the labour market, traditionally monopolized by men, the former tend to imitate men in different ways, including males' speech habits. In any case, the finding regarding gender differences in /r/-liaison variability also represents an instance of the interest in intralingual variation that is growing in Cognitive Linguistics.

Regarding the phonetic side of /r/-liaison variability, the present study has yielded two main findings. The first finding is that intrusive /r/ does not seem to be more common after back vowels than after central vowels, as could be expected from the acoustic similarity between back vowels and post-alveolar approximants. One reason for the failure to notice the effect of the similarity mentioned above could be, as explained above, that intrusive /r/ may be more stigmatized after back vowels since there are relatively few potential post-back-vowel cases of intrusive /r/ in the language, drawing thus speakers' attention because of their rare nature so speakers would try to avoid intrusive /r/ as much as they do with intrusive /r/ after central vowels. The second finding is that /r/-liaison is often avoided when the syllable that would make the link begins with /r/, which can be explained as a tendency to avoid similar sounds in adjacent positions. However, the results obtained do not reveal a statistically significant difference between both conditions but the few data analysed suggest that this question should be further explored in the future.

The main finding of this study regarding the usage-based aspect of /r/-liaison is that the latter seems to be more frequent in words where the potential context of /r/-liaison has a high degree of entrenchment or repeated use. This seems to be the case of word-internal intrusive /r/, compounds and collocations. However, the difference between rates of occurrence of linking /r/ and intrusive /r/ cannot be attributed exclusively to the high entrenchment of a small number of items with

a high lexical frequency since, excluding compounds and collocations, there are still significant differences between linking /r/ and intrusive /r/.

Despite the evidence found and the answers given to the research questions in this study, the latter has certain limitations that should be acknowledged. One clear limitation is that some of the phenomena investigated (e.g., /r/-liaison in word-internal position, syllables making the potential link beginning with /r/, potential cases of post-back-vowel intrusive /r/, etc.) were exemplified by very few items; as a result the conclusions drawn from the evidence obtained are limited or inconclusive. This suggests that the corpus should be enlarged to obtain more evidence regarding the issues mentioned above. This is feasible since the BBC WN website contains newscasts from 1999 till now.

A virtue but also a limitation of the study is that it only looked at one accentual variety. This study controlled for the variable of the speaker's accent but, as a feature of most non-rhotic English and not exclusively of RP, /r/-liaison could also be studied across accentual varieties. In this respect, future studies could be carried out to compare variability in /r/-liaison usage in different non-rhotic accents.

Another limitation of the present study is that diachronic change could not be investigated since the year of birth of the speakers is not available. In this respect, speakers' age has been claimed to affect variability in /r/-liaison usage. In the sixties, Jones (1960: 197) claimed that there appeared to be then "... an increasing tendency, especially among *younger* people, not to use linking /r/ at all" (emphasis added). However, a few years later, Bauer (1984) found no evidence of a correlation between the use of linking /r/ and the year of birth of the speaker, the year in which the recordings he analysed were made or the speaker's age at the time of recording. However, Bauer found some (limited) evidence with respect to intrusive /r/ usage suggesting that it was more common among speakers born after 1940. An analysis and comparison of RP speech from current and previous decades, should then be conducted to investigate the influence of the age factor in /r/-liaison usage. This analysis may provide data regarding changes in the use of /r/-liaison by different generations of broadcasters.

A further limitation of this study is that it did not look at potential differences in the occurrence of /r/-liaison due to stylistic variation. As mentioned above, it has already been suggested that use of linking/intrusive /r/ is a feature of fluent colloquial style, and is not so common in careful declarative style (e.g., Brown 1988: 145). This is contrary to Ramsaran's (1983) claim that the use of linking /r/ does not vary with formality in RP. However, since the corpus analysed in this study

contains a formal, careful, declarative style, where speakers tend to be speech-conscious, it cannot offer a picture of /r/-liaison usage in informal, colloquial, spontaneous speech. Further research with non-scripted speech, for instance, is needed.

Some further thematic questions that have been left unexplored regarding the phonetic aspect of /r/-liaison are whether rhythmic factors (e.g., the stressed/unstressed nature of the syllable following the syllable that makes the /r/-link) or intonational factors (e.g., presence/absence of an intonation boundary between the syllable making the link and the following vowel) condition /r/-liaison usage in some way.

Apart from the thematic limitations of the study mentioned above, the present study has some methodological and/or analytic limitations. One such limitation is that the statistical analyses carried out represent bivariate analyses (or the simultaneous analysis of two variables), used to uncover whether one variable (e.g., speaker sex) is related to another variable (e.g., rate of linking /r/s). However, this analytical approach may be modest for a complex phenomenon such as /r/-liaison where many different variables may obviously interact. Thus, a multivariate analysis (i.e., the observation and analysis of three or more variables at a time) would be a better analytical approach. However, given that some of the factors investigated are exemplified by very few potential contexts and/or instances in the corpus, we believe that to better test the possible influence of a third variable on a given original bivariate relationship or to test the joint effects of two or more variables upon a dependent variable a set of data elicited from informants in experimental settings should be used. This kind of data may allow us to control the number of potential cases of some of the phenomena under investigation. Thus two directions for future research are to obtain data under experimental conditions and to carry out multivariate analyses thereof. Thus, corpus-based and experimental studies may complement each other and provide a better understanding of variability in /r/-liaison. Moreover, other data-collection procedures could be used to investigate related issues such as the degree of prestige/stigmatization speakers attach to phenomena such as intrusive /r/. As a case in point, questionnaires could be used to survey speakers' conceptions and perceptions about /r/-liaison usage to uncover, for instance, what speakers' attitudes are towards intrusive /r/.

All the limitations mentioned above suggest directions for future research. However, despite its limitations, this study has provided an empirical perspective on a phenomenon so far hardly ever approached from data-driven research. Thus the study as a whole demonstrates how productive corpus-based methods can be for theoretical issues and it

represents a contribution, as claimed above, to usage-based Cognitive Linguistics.

Received 17 January 2007

University of Murcia

Revision received 15 May 2009

## References

- Allerton, David J. 2000. Articulatory inertia vs. 'Systemzwang': Changes in liaison phenomena in recent British English. *English Studies* 81(6). 574–581.
- Ashby, Michael & John Maidment. 2005. *Introducing phonetic science*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 1984. Linking /r/ in RP: Some facts. *Journal of the International Phonetic Association* 14. 74–79.
- Barsalou, Lawrence W. 1992. *Cognitive psychology. An overview for cognitive scientists*. Hillsdale, NJ: Lawrence Erlbaum.
- Blumstein, Sheila E. 1991. The relation between phonetics and phonology. *Phonetica* 48. 108–119.
- 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. 144–151.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Coates, Jennifer. 1993. *Women, men and language*. London: Longman.
- Collins, Beverley & Inger M. Mees. 2003. *Practical phonetics and phonology. A resource book for students*. London: Routledge.
- Croft, William. 2005. Logical and typological arguments for Radical Construction Grammar. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions*, 273–314. Amsterdam & Philadelphia: John Benjamins.
- Crystal, David. 1984. Should intruders keep out? In David Crystal (ed.), *Who cares about English usage?*, 36–44. London: Penguin.
- Donegan, Patricia. 1993. On the phonetic basis of phonological change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 98–130. London: Longman.
- Dubois, Sylvie & Barbara Horvath. 1999. When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change* 11. 287–314.
- Fillmore, Charles D. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6. 222–255.
- Foulkes, Paul. 1998. English [r]-sandhi: A sociolinguistic perspective. *Leeds Working Papers in Linguistics and Phonetics* 11. 78–95.
- Foulkes, Paul & Gerard J. Docherty. 2000. Another chapter in the story of /r/: 'Labiodental' variants in British English. *Journal of Sociolinguistics* 4(1). 30–59.
- Fox, Anthony. 1978. To 'r' is human? Intrusive remarks on a recent controversy? *Journal of the International Phonetic Association* 8(1–2). 72–74.
- García-Lecumberri, María L. & John Maidment. 2000. *English transcription course*. London: Arnold.
- Geeraerts, Dirk. 2003. Cultural models of linguistic standardization. In René Dirven, Roslyn Frank & Martin Pütz (eds.), *Cognitive models in language and thought. Ideology, metaphors and meanings*, 25–68. Berlin & New York: Mouton de Gruyter.

- Geeraerts, Dirk. 2005. Lectal variation and empirical data in Cognitive Linguistics. In Francisco Ruiz de Mendoza Ibáñez & M. Sandra Peña Cervel (eds.), *Cognitive Linguistics. Internal dynamics and interdisciplinary interaction*, 163–189. Berlin & New York: Mouton de Gruyter.
- Geeraerts, Dirk. in preparation. Recontextualizing grammar: Underlying trends in thirty years of Cognitive Linguistics. In Elzbieta Tabakowska, Lukasz Wiraszka & Michal Choinski (eds.), *Cognitive Linguistics in action: From theory to application and back* (Applications of Cognitive Linguistics).
- Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16. 29–54.
- Gick, Bryan. 2002. The American intrusive *l*. *American Speech* 77. 67–183.
- Giegerich, Heinz J. 1999. *Lexical strata in English: Morphological causes, phonological effects*. Cambridge studies in linguistics. Cambridge: Cambridge University Press.
- Gimson, Alfred C. 1980 [1962]. *An introduction to the pronunciation of English*, 3rd edn. London: Edward Arnold.
- Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
- Hay, Jennifer & Paul Warren. 2002. Experiments on /r/-intrusion. *Wellington Working Papers in Linguistics* 14. 47–58.
- Hay, Jennifer & Andrea Sudbury. 2005. How rhoticity became /r/-sandhi. *Language*. 81(4). 799–823.
- Heselwood, Barry. 2006. Final schwa and r-sandhi in RP English. *Leeds Working Papers in Linguistics and Phonetics* 11. 78–95.
- Hougaard, Anders. 2005. Conceptual disintegration and blending in interactional sequences: A discussion of new phenomena, processes vs. products, and methodology. *Journal of Pragmatics* 37(10). 1653–1685.
- Johansson, Stig. 1973. Linking and intrusive /r/ in English: A case for a more concrete phonology. *Studia Linguistica* 27. 53–68.
- Jones, Daniel. 1956 [1909]. *The pronunciation of English*, 4th edn. London: Cambridge University Press.
- Jones, Daniel. 1960 [1918]. *An outline of English phonetics*, 9th edn. Cambridge: Heffer.
- Keating, Pat A. 1991. On phonetics/phonology Interaction. *Phonetica* 48. 221–222.
- Knowles, Gerald. 1987. *Patterns of spoken English: An introduction to English phonetics*. London & New York: Longman.
- Kohler, Klaus J. 1995. Phonetics: A language science in its own right. In Elenius Dins Kjell & Peter Branderus (eds.), *Proceedings of the XIII International Congress of Phonetic Sciences*, vol. 1, 35–39. Stockholm.
- Kristiansen, Gitte. 2003. How to do things with allophones. Linguistic stereotypes as cognitive reference points in social cognition. In René Dirven, Roslyn Frank & Martin Pütz (eds.), *Cognitive models in language and thought. Ideology, metaphors and meanings*, 69–120. Berlin & New York: Mouton de Gruyter.
- Kristiansen, Gitte & René Dirven (eds.). 2006. *Cognitive sociolinguistics. Language variation, cultural models, social systems*. Berlin & New York: Mouton de Gruyter.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language, Variation and Change* 2. 205–254.
- Ladefoged, Peter. 2001 [1975]. *A course in phonetics*, 4th edn. Fort Worth, TX & Orland, FL: Hartcourt Brace Jovanovich College Publishers.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, vol. 1: *Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar*, vol 2: *Descriptive application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1999. A dynamic usage-based model. In Ronald W. Langacker, *Grammar and conceptualization*, 91–145. Berlin & New York: Mouton de Gruyter.
- Lewis, J. Windsor. 1975. Linking /r/ in the General British Pronunciation of English. *Journal of the International Phonetic Association* 5(1). 37–42.
- Lewis, J. Windsor. 1977. The r-link business—A reply. *Journal of the International Phonetic Association* 7(1). 28–31.
- Lin, Emilie L. & Gregory L. Murphy. 1997. Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance* 23. 1153–1169.
- Lindau, Mona. 1985. The story of /r/. In *Phonetic linguistics: Essays in honour of Peter Ladefoged*, Victoria A. Fromkin (ed.), 157–168. Orlando, FL: Academic Press.
- McMahon, April. 2000. *Lexical phonology and the History of English*. Cambridge: Cambridge University Press.
- Mohanan, Karuvannur P. 1986. *The theory of Lexical phonology*. Dordrecht: Reidel Publishing Co.
- Mompeán, Jose. A. 2004. Category overlap and neutralization: The importance of speakers' classifications in phonology. *Cognitive Linguistics* 15(4). 429–469.
- Moreno-Fernández, Francisco. 1998. *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel Lingüística.
- Mugglestone, Lynda. 2003. 'Talking proper': *The rise of accent as a social symbol*, 2nd edn. Oxford: Oxford University Press.
- Murphy, Gregory L. 1993. Theories and concept formation. In Ivan Van Mechelen, James Hampton, Ryszard S. Michalski & Peter Theuns (eds.), *Categories and concepts: Theoretical views and inductive data analysis*, 173–200. New York: Academic Press.
- Murphy, Gregory L. & Douglas L. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review* 92. 289–316.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Ohala, John J. 1987. Explanation in phonology: Opinions and examples. In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds.), *Phonologica 1984*, 215–225. London: Cambridge University Press.
- Ohala, John J. 1990. There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* 18. 153–171.
- Peterson, Gordon E. & Barney, Harold L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 22. 369–381.
- Pierrehumbert, Janet. 2000. The Phonetic grounding of phonology. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5. 7–23.
- Ramsaran, Susan. 1983. Stylistic variation in RP. In Matthew P. R. Van Den Broeke & Ariel Cohen (eds.), *Proceedings of the X International Congress of Phonetic Sciences*, 624–627. Dordrecht: Foris.
- Rips, Lance J. 1989. Similarity, typicality, and categorization. In Stella Vosniadou & Andrew Ortony (eds.), *Similarity and analogical Reasoning*, 21–59. New York: Cambridge University Press.
- Rips, Lance J. 1995. The current state of research on concept combination. *Mind and Language* 10. 72–104.
- Roach, Peter, James Hartman & Jane Setter. 2006. *English Pronouncing Dictionary with CD-ROM*, 17th edn. Cambridge: Cambridge University Press.

- Rohrer, Tim C. 2005. Embodiment and experientialism. In Dirk Geeraerts & Herbert Cuyckens (eds.), *The Oxford handbook of Cognitive Linguistics*, 48–81. Oxford: Oxford University Press.
- Schank, Roger C. & Robert P. Abelson. 1977. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Torstensson, Niklas. 2004. Newstalk: Reductions in careful, scripted speech. In *Proceedings of FONETIK 2004*, 16–19. Department of Linguistics, Stockholm University, [http://www.ling.su.se/fon/fonetik\\_2004/torstensson\\_fonetik2004.pdf](http://www.ling.su.se/fon/fonetik_2004/torstensson_fonetik2004.pdf)
- Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Tummers, José, Kris Heylen & Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1. 225–261.
- Uffmann, Christian. 2007. Intrusive /r/ and optimal epenthetic consonants. *Language Sciences* 29. 451–476.
- URL 1 BBC World Service website. *Learning English* section (last accessed 20 October 2008). [www.bbc.co.uk/worldservice/learningenglish/newsenglish/](http://www.bbc.co.uk/worldservice/learningenglish/newsenglish/)
- URL 2 BBC World Service website. *Learning English* section. Archive from 2004 (last accessed 20 October 2008). [www.bbc.co.uk/worldservice/learningenglish/newsenglish/witn/archive\\_2004.shtml](http://www.bbc.co.uk/worldservice/learningenglish/newsenglish/witn/archive_2004.shtml)
- URL 3 BBC World Service website. *Learning English* section. Archive from 2005 (last accessed 20 October 2008). [www.bbc.co.uk/worldservice/learningenglish/newsenglish/witn/archive\\_2005.shtml](http://www.bbc.co.uk/worldservice/learningenglish/newsenglish/witn/archive_2005.shtml)
- URL 4 Free software Speech Filing System (SFS) UCL website (last accessed 20 October 2008). <http://www.phon.ucl.ac.uk/resource/sfs/>
- Vennemann, Theo. 1972. Rule inversion. *Lingua* 29. 209–242.
- Watt, Dominic J. L. & Ann L. Milroy. 1999. Patterns of variation and change in three Tyneside vowels: Is this dialect levelling? In Peter Foulkes & Gerard J. Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 25–46. London: Arnold.
- Wells, John C. 1982. *Accents of English*, 3 vols. Cambridge: Cambridge University Press.
- Wells, J. C. 1994a. “Transcribing Estuary English—A discussion document”. *Speech, Hearing and Language: UCL Work in Progress* 8. 259–267.
- Wells, J. C. 1994b. “The cockneyfication of RP?” In Gunnel Melchers & Nils-Lennart Johannessen (eds.), *Nonstandard varieties of language: Papers from the Stockholm symposium*, 198–205. Stockholm: Almqvist and Wiksell International.
- Wells, John C. & Greta Colson. 1971. *Practical phonetics*. London: Pitman Publishing.