Phonology 16 (1999) 29–54. Printed in the United Kingdom

1166

*A gesture-based account of intrusive consonants in English**

Bryan Gick
Yale University and Haskins Laboratories

1 Introduction

A number of recent papers have demonstrated the advantages of using a phonological model incorporating the timing and magnitude of articulatory gestures to account for alternations involving segments such as the English nasals, liquids and glides (e.g. Krakow 1989, Browman & Goldstein 1992, 1995, Sproat & Fujimura 1993, Gick, in press). Some of these works (McMahon *et al.* 1994, McMahon & Foulkes 1995) have made specific reference to the well-known phenomenon of English INTRUSIVE *r*, shown in (1).

(1) *The intrusive* r *of British RP and eastern Massachusetts*

the idea	[aidi(j)ə]	→ the idea is	[aidi(j)əɹɪz]
I draw	[drɔ:]	→ I'm drawing	[drɔ:ɹɪŋ]

However, previous analyses have not linked the intrusive *r* explicitly to other similar processes, nor viewed all of these processes as the natural results of more general principles of phonological organisation. Thus, the intrusive *r* has remained, in the eyes of most linguists, an isolated quirk of English history, or, as one phonologist (McCarthy 1993: 191) has called it, 'the phonologically unnatural phenomenon of *r*-epenthesis'.

The present paper introduces into the discussion of intrusive *r* a recently documented related phenomenon known as INTRUSIVE *l* (Gick 1991, 1997, in preparation, Miller 1993). It is argued that these new facts, in conjunction with current advances in the understanding of articulatory factors in syllable structure, support a view in which the intrusive *r* and *l* are synchronically underlyingly present. Their patterns of appearance and disappearance are argued to be typical of the patterns of timing, augmentation and reduction seen in all segments comprising multiple

* This work was supported by NIH grants DC-02717 and HD-01994 to Haskins Laboratories and by an NSF Graduate Fellowship to the author. Many thanks to Louis Goldstein and Doug Whalen for continued helpful discussion and support, to the editors and three anonymous reviewers for their useful comments, and to Heinz Giegerich for fostering my interest in the very earliest stages of this project.

supralaryngeal gestures (henceforth COMPOSITE SEGMENTS) in similar environments. It will be shown that this approach provides a single account for intrusive *r* and *l* as well as a number of other phenomena from a variety of English dialects, using only principles of gestural organisation needed elsewhere in the phonology.

§2 of this paper covers the descriptive facts of intrusive *r*, and reviews previous analyses. In §3, the details of intrusive *l* are briefly introduced and the relevance of these facts to the interpretation of the intrusive *r* discussed. §4 relates glide formation to the *r* and *l* phenomena, and finally, §5 presents a unified analysis of intrusive consonants and related phenomena based on general properties of gestures in English phonology.

2 Intrusive *r* and other *r* phenomena

INTRUSION typically refers to the presence of a non-historical consonant between two heterosyllabic vowels. Whether this should be considered a case of synchronic insertion or historical reanalysis is a matter to be addressed in the course of this paper. The term intrusive has also occasionally been applied to other types of epenthesis, including a non-historical consonant in final positions (e.g. Bristol *l*; Wells 1982: 344) and a set of excrescent stops, as the [t] in *prin*[t]*ce* (e.g. Clements 1987). The Bristol *l* and similar cases involving final *r* will be referred to as EPITHETIC *l* or *r* in this paper. Other usages of the term to refer to non-sandhi phenomena (such as that described by Clements) will be avoided.

All dialects having intrusive *r* also seem to require two subordinate processes: *r*-vocalisation (the reduction or apparent complete loss of some or all coda /r/'s – also known as *r*-loss, *r*-deletion, *r*-dropping or simply *r*-lessness or non-rhoticity) and linking *r* (the non-deletion or reinsertion of these historical /r/'s when followed by a vowel-initial morpheme). Both of these processes will be discussed below.

2.1 *r*-vocalisation, linking and intrusion

Intrusive *r*, a phenomenon common to many dialects of English, has remained an unresolved topic in phonology. The lack of agreement regarding the intrusive *r* is due in part to the ambiguity inherent in the available data, which has centred almost exclusively on only two dialects: Southern British Received Pronunciation (or RP; e.g. Vogel 1986: §5, Harris 1994: ch. 5, McMahon *et al.* 1994, McMahon & Foulkes 1995) and the dialect of eastern Massachusetts (e.g. Whorf 1943, Kahn 1976, McCarthy 1991, 1993, Halle & Idsardi 1997). No other *r*-intruding dialects have been described which provide evidence bearing conclusively on the controversy.

A full description of all the known facts surrounding intrusive *r* is not necessary here, especially given the wealth of detailed descriptions already available. Jones (1989: §5.3.2–5.4), for example, presents the historical

development of *r*-loss, linking and intrusion in British dialects; McMahon *et al.* (1994) list the facts of the phenomenon as it exists in RP, with mention of a few other dialects, as well as a nearly exhaustive bibliography of recent works of British authorship on the topic (1994: 305); McCarthy (1991, 1993) gives a good account of the *r*-linking and *r*-intrusion phenomena in the dialect(s) of eastern Massachusetts.

2.1.1 r-*vocalisation*. Historically, the first process that led to the present state of the intrusive *r* was the reduction or loss of coda /r/'s before tautosyllabic consonants. Jones (1989: §5.3.3) gives orthoepic evidence of the dropping or reduction of /r/ in this environment from as early as the 15th century, citing such alternate spellings as: ⟨bersel/bessel⟩, ⟨hors/hos⟩, ⟨harsk/haske⟩, ⟨morther/mother⟩ (for 'murder', not 'mother'), ⟨quart/quat⟩, etc. By the 18th century in London the general reduction of coda /r/'s was a thoroughly well-entrenched, if not well-accepted, practice.

In present-day RP, where a vocalised /r/ follows the vowels [a ɔ ə], it has apparently merged into the preceding vowel. Following all other vowels (the high vowels and diphthongal offglides), a historical final /r/ appears as a final schwa. In fact, some researchers (e.g. Giegerich 1997) have argued that the former set of vowels are all phonetically schwa-final, with /a/ and /ɔ/ surfacing as something like [aə] and [ɔə] in intruding dialects.

2.1.2 *Linking* r. Linking *r* is a process common to many *r*-vocalising dialects, whereby a historically attested final *r* is not vocalised (or is vocalised and then reintroduced, depending on the account) when followed by a vowel-initial morpheme. This *r* is generally presumed to be retained or inserted either to serve as a 'hiatus-breaking' element, or to provide a sufficient onset or coda to the following or preceding syllable, respectively. Given the tendency of final glide–*r* combinations to surface phonetically as glide–schwa (as in (2d)), it can be said of most dialects that linking *r* occurs only following the vowels [a ɔ ə], thus:

(2)	*vocalised* r		→ *linking* r	
a. / a __	'mar'	[maː]	→ 'mar is'	[maːɪɪz]
b. / ɔ __	'lore'	[lɔː]	→ 'lore is'	[lɔːɪɪz]
c. / ə __	'coder'	[kodə]	→ 'coder is'	[kodəɪɪz]
d. / ə __	'deer'	[di(j)ə]	→ 'deer is'	[di(j)əɪɪz]

Not all dialects having vocalised coda *r*'s also show linking, as evidenced by examples such as *Mister Adams* [mɪstə(ʔ)ædəmz] in Southern U.S. (Kurath 1964) and African American vernacular dialects (Fasold 1981). In such cases, as there is no surface evidence in any environment for their underlying presence, these final historical *r*'s must be viewed as having been dropped altogether.

2.1.3 *Intrusive* r. Intrusive *r* is typically described as being identical in its surface manifestation to linking *r*, the only difference being that in the

case of intrusion, no final *r* was historically present (e.g. *idea* > *idea*[r] *is*). *r*-intrusion only occurs in dialects having both *r*-vocalisation and *r*-linking. The relationship between vocalised, linking and intrusive *r* has traditionally been attributed to analogy. Wells (1982: 223), for example, states: 'intrusive /r/ arises essentially from the natural tendency to give identical treatment to words with identical endings'. Thus, leaving aside for the moment the important, and often overlooked, step of establishing some motivation for the vocalisation and linking of *r* in the first place, the intrusive *r* appears to follow by analogy on grounds of the phonetic similarity or identity of the forms in isolation (see (3)). In (3), the first row of each pair shows vocalisation and linking; the second row intrusion:

(3) *The linking/intrusive* r *analogical paradigm*

a. / a __	'mar'	[ma:]	→	'mar is'	[ma:ɹɪz]
	'ma'	[ma:]	→	'ma is'	[ma:ɹɪz]
b. / ɔ __	'lore'	[lɔ:]	→	'lore is'	[lɔ:ɹɪz]
	'law'	[lɔ:]	→	'law is'	[lɔ:ɹɪz]
c. / ə __	'coder'	[kodə]	→	'coder is'	[kodəɹɪz]
	'coda'	[kodə]	→	'coda is'	[kodəɹɪz]
d. / ə __	'deer'	[di(j)ə]	→	'deer is'	[di(j)əɹɪz]
	'idea'	[-di(j)ə]	→	'idea is'	[-di(j)əɹɪz]

Again, as with linking, not all dialects that could intrude *r* do so. Some regions that are reported to have vocalisation and linking of *r* without intrusion include areas of New England and the southeastern United States, as well as South Africa (see e.g. Mohanan 1985: 147–148, McMahon *et al.* 1994: 313, note 9, and references therein). Unlike the intrusive *r* dialects, where there is no surface difference between historically *r*-final and non-*r*-final forms, these dialects show clear surface evidence of an underlying distinction.

2.1.4 *Other* r*-insertions.* In addition to the case of intrusion described above, there are also a number of dialects that show a non-historical *r* in other environments. Wells refers to these dialects as HYPER-RHOTIC, citing examples from southwestern England (1982: 343), eastern New England (1982: 522) and parts of the U.S. South (1982: 542). These cases involve the appearance not only of final historical *r*, but also of a historically unattested *r* at the end of certain pre-consonantal or utterance-final words, for which John F. Kennedy gave us numerous examples during the missile crisis in *Cubar*. Unlike intrusion, however, this epithetic *r* is not a sandhi phenomenon. That is, it does not require a following vowel, and is therefore in fundamental conflict with any account claiming that *r* is necessarily inserted to fill hiatus, or to provide an onset for a following onsetless syllable. Harris (1994: 296, note 40) takes the implications of epithetic *r* a step further, stating: 'in such systems, the appearance of *r* in, say, ⟨comma[r]⟩ is not restricted to sandhi contexts, which indicates that it has a lexical source (as opposed to being inserted)'. This possibility will be more thoroughly explored in following sections.

An additional phenomenon commonly thought to be connected to the other types of *r*-insertion discussed above is INTERNAL *r*-EPENTHESIS. Wells (1982: 522) points out sporadically appearing examples from eastern New England, such as *cough* [kɒrf] and *cloth* [klɒrθ], attributing these to essentially the same process of historical vocalisation, merger and reanalysis that leads to intrusive *r*. However, apparently similar cases of internal *r*-epenthesis may be found in dialects that have no history of *r*-loss. Examples include the *wa*[r]*sh*, *Wa*[r]*shington*, *squa*[r]*sh* and *go*[r]*sh* pronunciations known to be a feature of the Midland dialect region of the U.S. (Kurath 1949) and extending at least to St. Louis, Missouri (Murray 1986), but purported to have a much broader distribution. A search of the publicly available archives of the American Dialect Society email list (http://www.americandialect.org/) for the string ⟨warsh⟩ resulted in 34 records since 1992, giving further personal and anecdotal citations of these forms in the U.S. Midwest and inland South (Indiana, Illinois, Minnesota, Michigan, Kansas, Oklahoma, Arkansas) as well as the Northwest (Oregon, Idaho, Northern California). In two cases, these pronunciations even appear in *l*-intruding dialects. Washingtonville, Pennsylvania, a town in an intrusive *l* region, is called *Wa*[r]*shingtonville* by the locals. Similarly, much of the state of Maryland near Washington, DC is basically *l*-intruding, but *Wa*[r]*shington* is still the standard pronunciation. As widespread as this practice is, however, no account has been proposed for this type of insertion.

2.2 Previous accounts of intrusive *r*

According to Wells (see above), the descriptive paradigm in (3) is all that can be said about the intrusive *r*. Because of the apparent mergers of surface forms in the observed dialects, it is impossible to distinguish between the various theories that differ as to the underlying forms, the motivating factors, the syllable function of the intruded element, and so on. McMahon *et al.* (1994: 305) generalise the various approaches to the analysis of intrusive *r* to three types: deletion accounts (where the *r* is underlying following all non-glide-final vowels, and is deleted in some environments), insertion accounts (where *r* is not underlying in any final position, and is inserted under certain circumstances) and combined deletion-insertion accounts (where some surface *r*'s are underlying and some are not). Although there has been much debate on this matter, the majority of recent approaches involve, at some point, the synchronic insertion of some element, usually a complete, phonetic [r], as shown descriptively in (4).

(4) *Descriptive intrusion rule*

$$\emptyset \rightarrow r \;/\; V_1^{\sigma_1} \{\#,+\} \_\_ V_2^{\sigma_2}$$

In (4), V_1 consists of the strictly limited set of vowels /ə a ɔ/ for most intrusive *r* dialects, while V_2 can be any vowel (# and + indicate word and morpheme boundaries, respectively). It may be possible to reduce further the V_1 set to [ə] only, as suggested by Giegerich (1997) (§2.1.1 above), among others. Harris's (1994: 254) FLOATING *r* phonological account also identifies schwa as a crucial factor in understanding intrusive *r*, though not synchronically but historically: 'the historical evidence suggests that intrusive *r* has been around for a long time and that its emergence was originally motivated by a disfavouring of final schwa'. He states on the same page that 'later mergers with vowels from other sources (such as PAW = PORE) would have resulted in the extension of floating *r* to etymologically *r*-less forms'.

Further support for the view that schwa is a key element in the analysis of intrusive *r* can be found in a contemporary description of early 20th-century RP, which seems to have had quite a different synchronic pattern of intrusion. In the 'Explanations' chapter of his *English Pronouncing Dictionary*, Daniel Jones (1928: xvii) observes that a 'considerable proportion' of RP speakers insert an *r* not only in linking environments, but 'also at the end of every other word terminating with «ə». These pronounce **the idea of it** «ðiai'diərəvit».' That is, a 'considerable' number of speakers of 1917 London English exhibited the intrusive *r*, but only following schwa. He goes on to point out that 'a few [RP speakers] extend the practice to all words ending in «ɑ:» and «ɔ:»'. Thus, /ə (a ɔ)/ will be considered to be the set of pre-intrusion vowels.

Harris (1994: 246–247) argues against the synchronic insertion approach of the type illustrated in (4), citing as one problem that such accounts provide no grounds for the selection of *r* as the inserted element. Halle & Idsardi (1997: §2.1) identify a similar problem in McCarthy's (1993) constraint-based analysis. McCarthy (1993: 190), unable to answer the question of, as he succinctly puts it, 'Why r?', resorts to the addition of 'a phonological rule of r insertion ... a phonologically arbitrary stipulation'. However, in view of the variety of dialects in which *r*-intrusion occurs, and the apparent interconnectedness between intrusion and other processes such as vocalisation and linking, it seems that the selection of *r*, although not well understood, is unlikely to be arbitrary.

Perhaps a more basic problem with insertion is the lack of surface evidence for positing an underlying synchronic distinction between historic and non-historic *r*'s.[1] My own observations indicate that naive *r*-

[1] I know of only one claim in the literature that there exists possible surface evidence indicative of an underlying distinction. McCarthy (1991), following an observation previously made by Whorf (1943), identifies a single environment in his eastern Massachusetts dialect where a distinction if pronounced between members of near-minimal pairs spelled with final ⟨r⟩ (e.g. *rear*) and without final ⟨r⟩ (e.g. *Maria*). As a lone counterexample, this example has numerous problems. First, there is every reason to believe that *rear* is monosyllabic, while the corresponding portion of *Maria* is disyllabic; this most simply accounts for the surface distinction between the two. Second, syllabification problems aside, the very limited scope of this surface distinction (it apparently does not apply following any other vowels) is much

intruding speakers, especially children, invariably contend that final vocalised /r/'s have undergone complete phonetic merger with the final vowels [ə a ɔ], as described above in §2.1.1. I know of no perceptual experiments either affirming or contradicting their claim. One case affirming the presence of an underlying /r/ in these forms occurred when I, a native speaker of a non-*r*-vocalising dialect, was living in Britain for the first time. Asking an RP speaker why borrowings such as *taco* and *pasta* were pronounced with an [æ] rather than an [a] as in American English elicited the incredulous response, 'What, you mean with an *r*?'. This is consistent with the general orthographic usage of ⟨r⟩ to designate vowel lengthening in RP English (*har har, blar blar, er*, etc.; cf. U.S. *ha ha, blah blah, uh*). Further along this line, one reviewer of this paper cites experience with non-rhotic subjects who produce forms such as *Lisa*[r] in isolation in careful speech. In any case, if a phonetic merger has in fact taken place, then it can only be on the basis of spelling that some linguists are reluctant to consider that they have phonologically merged as well.

An account of intrusive *r* in which the *r* is lexically, or underlyingly, present would of course avert all of these problems with synchronic insertion, but such a step cannot be taken lightly. A number of researchers have argued that the intrusive *r* is in fact partly or fully present in lexical representations (Mohanan 1985, Harris 1994; also see Giegerich 1997: note 19 and McMahon *et al.* 1994: 305 for others). It seems inescapable, however, that to avoid stipulating the insertion of an *r* in certain environments, one must stipulate its deletion in others. Only if all of the patterns associated with intrusion were shown to be typical of underlying consonants in the same environments could the view be supported that no special allowance is needed to deal with intrusion. After assimilating some additional data in the following sections, we shall return to this possibility in §5.

3 Intrusive *l* and other *l* phenomena

3.1 *l*-vocalisation, linking and intrusion

The intrusive *l* (see (5)) is a surprisingly widespread phenomenon, showing patterns similar to those seen in the intrusive *r*. Its use has been primarily identified with working class and rural dialects in Pennsylvania, Delaware and other areas of the northeastern United States, though it has

---

more consistent with what might be expected of a spelling pronunciation, similar to the dialect-wide Philadelphia resurrection of the /l/ following /a/, in such words as *talk* and *palm*, but not in other environments (e.g. *salmon, half, yolk*, etc.). Third, and most importantly, it is very tenuous to suggest that such a minimal distinction could drive acquisition of the across-the-board underlying distinction necessary to support the synchronic insertion/deletion paradigm that has been proposed for eastern Massachusetts (and, despite the extensive literature on the subject, not even one such example has been cited for any other *r*-intruding dialect).

been reported in all other regions of the country except the Northwest. The few references to the intrusive *l* to date have all been made in unpublished works (Gick 1991, 1997, in preparation, Miller 1993).

(5) *The intrusive* l *of south-central Pennsylvania*

I draw	[dɹɔːʷ][2]	→ I'm drawing	[dɹɔːlɪŋ]
the bra	[bɹaː]	→ the bra is	[bɹaːlɪz]

The intrusive *l*, like the intrusive *r*, can be described as the presence of a non-historical consonant between two vowels, the first of which belongs to a limited set always including /ɔ/, and very rarely /a/ and /ə/ as well. Also reminiscent of the intrusive *r* is the hierarchy of related processes associated with *l*-intrusion, which is fixed in the familiar order of vocalisation, linking and intrusion.

As with *r*, the historical vocalisation of some pre-consonantal coda /l/'s is evident from orthographic artefacts in such words as *half*, *salve*, *salmon*, *talk*, *calm*, *folk*, etc. In some dialects, however, this process has extended to all coda /l/'s (e.g. *drawl* [dɹɔːʷ]). Ash (1982a) says of certain Philadelphia-area dialects: 'the vocalisation of /l/ results from the loss of contact between the tongue and the palate', a description of *l*-vocalisation that is well documented for many English dialects (Giles & Moll 1975, Kahn 1976: 58, 104–105, Ash 1982a, b, Hardcastle & Barry 1989, Browman & Goldstein 1995: 26–27, Narayanan *et al.* 1997: 1070).

Again parallelling the patterns of *r*, in most /l/-vocalising dialects, a brighter allophone of /l/ (one involving a greater tongue tip constriction, among other things) may be heard in intervocalic positions (e.g. *drawling* [dɹɔːlɪŋ]). That is, most dialects with vocalisation of /l/ also show linking (see Ash 1982b for an exceptional dialect).

Finally, in some dialects having both vocalisation and linking, as with intrusive *r*, an [l] can appear intervocalically when a vowel-final word or stem is followed by another vowel. Also similar to the *r* case, as stated above, the intrusive *l* may appear following only a very limited set of vowels.

3.1.1 *Epithetic* l. As with [r], [l] can also undergo epithesis in some *l*-vocalising dialects. The typical case of *l*-epithesis is in the Bristol dialect (Weissmann 1970, Wells 1982), where an [l] is inserted following final schwa, as in *idea*[l] and *area*[l]. A similar phenomenon has been observed in southern Ohio (Lutz 1984, Beverly Flanigan, personal communication) following [ɔ], resulting in utterance-final or preconsonantal pronunciations such as *saw*[l], and misspellings such as ⟨falcet⟩ *faucet*, ⟨papal⟩ *papaw* and ⟨seasowl⟩ *seesaw* (Lutz 1984: 52). Both of these dialect areas have a history of *l*-vocalisation, and both present the same problem for previous syllabically motivated accounts of intrusion: that epithesis occurs regardless of the presence of a hiatus.

[2] A uvularised back rounded vowel is used here for lack of a better device, though upper pharyngeal might be more appropriate for some speakers (Narayanan *et al.* 1997: 1072, 1076).

3.2 Implications of the intrusive *l*

Having set forth the facts of these two similar but independently arising sets of complex patterns, one involving *r* and the other *l*, certain observations can now be made regarding the properties of consonantal intrusion in general. Given the complexity of these processes and their parallel historical developments, it is extremely unlikely that they arose out of separate phonological and phonetic factors. Thus, these factors may no longer be considered properties of *r* – rather, as they apply to multiple segments, they must be considered properties of English.[3] This will sharply narrow the field of possible answers to why and how intrusion occurs. A related issue that may now be more readily addressed is the question raised in §2.2 of how to find a phonological model that can predict *r* as the intruded element. When only *r* was involved, it was possible to argue for *r* simply being the default consonant of English, at least in these dialects. This possibility is ultimately rejected by McCarthy (1993: 190), and is now further weakened by the fact that *l* can also undergo intrusion. Rather, what we must now try to understand are the relationships that apply in both the *r* and *l* cases, such as the connection between vocalisation and linking, and subsequently, intrusion. We must therefore focus on why these processes all occur together, and ideally, how they might be viewed as parts of a single process.

Further important questions regarding intrusive *r* and *l* fall out of a comparison of their surrounding environments. While *r* in most *r*-intruding dialects surfaces following /ə (a ɔ)/, *l* usually appears following /ɔ/ (though it can also occasionally follow /a/ or /ə/ in some dialects). This prompts two questions. First, how are the vowels /ə ɔ a/ different from other final vowels such that they should prompt intrusion? Second, the existence of dialects where *r*- or *l*-intrusion occurs only following /ə/ or /ɔ/ precludes the common interpretation of intrusion as a blind insertion to break up all heterosyllabic VV sequences. Thus, what is this connection between [ə] and [r] and between [ɔ] and [l], and how could it have independently given rise to these two almost identical phenomena? The first of these questions will be addressed in §4, and the second in §5.

4 Glide insertion

According to McCarthy (1993: 176), the set of pre-intrusion vowels /ə ɔ a/ contains 'the only true vowels occurring in word-final position in English'. The term 'true vowels' here presumably refers to the fact that these are all and only the vowels that may appear in English without being

[3] While the factors discussed in the present paper are indeed properties of English, I do not wish to imply that they are exclusively so. As pointed out by one of the reviewers of this paper, similar processes to consonantal intrusion have been observed in, among others, French, Uyghur and possibly Korean, Cantonese and others. While discussion of these goes beyond the scope of the present work, it certainly provides a promising next step for this research programme.

followed by either a high offglide [i̯] or [u̯] or some other consonant in the same syllable (similar observations have been made by Harris 1994: 247 and others). Numerous researchers have identified the high vowels and diphthongs as eliciting an inserted or augmented homorganic glide in prevocalic (sandhi) positions. Broadbent (1991), Harris (1994: 247) and others have explicitly likened this process to intrusive *r*, while still others have referred to these glides directly as 'linking' (Whorf 1943) or 'intrusive' (Bailey 1985: 164). If these final offglides can thus be viewed as adequate postvocalic hiatus-fillers, it may be informative to more carefully investigate them in the same intervocalic positions where intrusion arises.

It is also known that final offglides are more clearly audible (or less vocalised) in some dialects than in others. In particular, such retention of final glides is often cited as a feature of General American English, and is indeed true of many American dialects. Similarly, RP, southern U.S. and other *r*-vocalising dialects are well known for their reduction of final offglides. This pattern of co-occurrence between glide vocalisation and final *r*- and *l*-vocalisation should not be overlooked. A general account giving a uniform treatment to all of these phenomena, and thus accounting for such correspondences, will be outlined below.

5 A gesture-based account

5.1 Properties and types of gestures

McMahon *et al.* (1994: 303) suggest a partial answer to the second question raised in §3.2 (that of the connection between /r/ and schwa), proposing an analysis within ARTICULATORY PHONOLOGY (AP), a phonological model developed by Browman & Goldstein (1986, 1989, 1992). In this model, the fundamental unit is not the segment or the feature, but the articulatory GESTURE. Gestures in this framework consist of the constriction of certain defined regions of the vocal tract by the coordinated effort of some number of independent articulators (say, the upper lip, lower lip and jaw), and are bound to one another only by lexicalised PHASING, or timing, relations. Groups of lexically phased gestures are referred to as CONSTELLATIONS. According to the principles of AP, all gestures present in the phonetic output of a word are specified in its lexical representation. Thus, while the magnitude or timing of a gesture may vary, gestures may not be inserted. The account proposed in this section focuses on two recently identified general properties of gestures, one of which dictates the magnitude or size of different constrictions, and the other of which governs their timing with respect to one another. These have been referred to as FINAL REDUCTION and POSITION-SPECIFIC GESTURAL TIMING (Browman & Goldstein 1992, 1995).

5.1.1 *Final reduction.* Final reduction and position-specific gestural timing are general properties tied to syllable affiliation, and have been observed directly and indirectly in English and a number of other

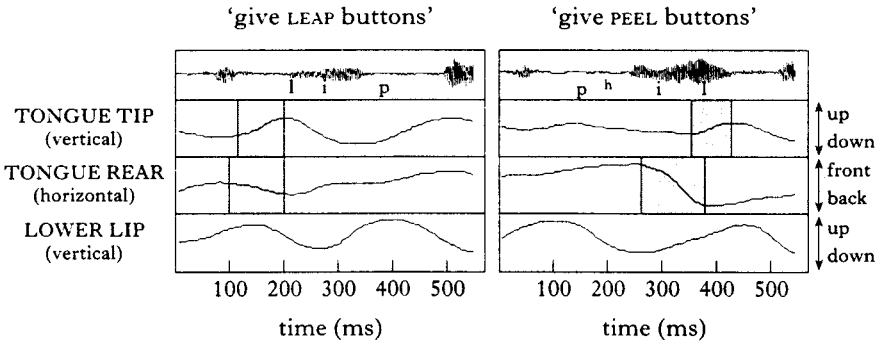

Figure 1
Relative timing of /l/ gestures
(after Browman & Goldstein 1995)

languages (see relevant cross-linguistic data in Anderson 1995, Wang 1995, Walsh Dickey 1997: §§2, 3.3). Final reduction refers to the observation that consonants tend to be produced with less constriction in syllable-final positions than in other positions. This phenomenon has been observed experimentally in stop consonants (Macchi 1988, Turk 1994, Browman & Goldstein 1995), nasals (Browman & Goldstein 1995), /l/ (Giles & Moll 1975, Ash 1982, Hardcastle & Barry 1989: 15, Sproat & Fujimura 1993) and glides (Gick, in press).

An alternative view of this process is INITIAL STRENGTHENING (e.g. Keating *et al.*, in press), whereby domain-initial elements are increased in magnitude, rather than final ones being decreased. The fact that higher-level domains result in greater gestural magnitudes seems to speak in support of this latter view. However, as the outcome of this controversy does not bear substantially on the present discussion, the former term will be applied during the remainder of this paper.

5.1.2 *Gestural timing.* Position-specific gestural timing (see Browman & Goldstein 1995 for a thorough discussion) refers to a property of composite segments, by which the relative timing between component gestures varies predictably depending on syllable position. A preponderance of evidence suggests that both /l/ and /r/, among others, qualify as composite under this definition, with /l/ comprising both tongue tip raising and tongue dorsum backing gestures, and /r/ comprising tongue blade raising, pharynx constriction (tongue root retraction) and possibly lip constriction gestures. This multi-gestural view of the English liquids has been argued on both phonological grounds (see Walsh Dickey 1997 and references therein), and instrumental grounds (Delattre & Freeman 1968, for /r/ and /l/; Narayanan *et al.* 1997, for /r/; Sproat & Fujimura 1993, Browman & Goldstein 1995 and Alwan *et al.* 1997, for /l/).

Figure 1 shows an example of the relative timing of the two gestures of /l/ (tongue tip closure and tongue dorsum backing). Note that in final position (*peel*, image on right) the more front, or typically consonant-like,

tongue tip gesture occurs much later than the back, or more vowel-like, tongue dorsum retraction gesture.

5.1.3 *Two types of gestures: C-gestures and V-gestures.* The two timing patterns shown in Fig. 1 led Sproat & Fujimura (1993) to propose that composite segments are composed of two distinct types of gestures. They refer to these as CONSONANTAL and VOCALIC, respectively, based on the degrees of constriction of the two component gestures of /l/ (a full closure at the tongue tip and an approximant constriction at the tongue dorsum). They further observe that each type of gesture shows a characteristic timing pattern relative to the syllable nucleus, with consonantal gestures tending to be further from the syllable nucleus, while vocalic gestures are closer to it. Gick (in press), in a study of the corresponding gestures in glides, finds that the same patterns hold for /w/. However, as both of the component gestures of /w/ – lip and tongue dorsum – are approximants, Gick suggests that Sproat & Fujimura's constriction degree-based terminology be replaced, recommending the terms C-GESTURE and V-GESTURE, respectively. The latter terms will be used for the remainder of this paper. Thus, based on a comparison of timing patterns with /l/ and nasals, the lip constriction of /w/ is considered a C-gesture, while the raising of the tongue dorsum is a V-gesture.

5.2 Vocalisation

5.2.1 *Vocalisation of* /l/ *and* /w/. §§5.1.1 and 5.1.2 outlined two effects – one of magnitude and the other of timing – that account for differences observed between allophones of liquids and glides in English. Applying these effects to the present discussion, it should now be straightforward to construct an account in which *r*- and *l*-vocalisation follow directly from the processes of final reduction and position-specific gestural timing. Thus, when an /l/ is in syllable-final position, the C-gesture (tongue tip fronting and raising) is both reduced in magnitude (final reduction) and temporally delayed, resulting in what has been traditionally called vocalisation. In this way, for example, the model accurately predicts the process of *l*-vocalisation in Philadelphia and other dialects, which has been described (see §3.1 above) as the reduction of only the tongue tip constriction. This account of the vocalisation of /l/ is basically the one suggested by Browman & Goldstein (1992: 165–167, 1995), Sproat & Fujimura (1993), McMahon *et al.* (1994) and Tollfree (1996: §5.6.iii).

In this analysis, unlike in most theories of allophonic variation, these effects apply not just to the single segment /l/ or its component gestures, but equally across all C-gestures and V-gestures in the system. This, then, accounts for the cross-dialectal co-occurrence of various vocalisation phenomena discussed in §4. Thus, assuming the final offglides to be variants of initial glides /j/ and /w/, exactly the same account can be given for the vocalisation of final glides. Gick (in press) shows evidence of both final reduction and typical syllable position timing effects in /w/ for speakers of three different dialects of American English (as /j/ was found

to have only one gesture, it will not be included in the remainder of this discussion).

5.2.2 *Vocalisation of* /r/ *: experiment 1.* Essentially this same account has been proposed for vocalisation of /r/ (McMahon *et al.* 1994). However, unlike the /l/, glides and nasals, no experimental evidence exists in support of this claim. Because of instrumental limitations in measuring the pharyngeal component of English /r/ during speech production, knowledge of the articulatory workings of /r/ has remained insufficient to provide more than a circumstantial basis for such an analysis. The phonological similarities shared by /r/, /l/ and the glides cited elsewhere in this paper strongly indicate that the same analysis will be supported for /r/ as for /l/ and /w/.

It is true that the difficulty of measuring the pharynx makes it impossible for us to determine the internal timing characteristics of /r/. However, it is still possible to measure variations in magnitude in those gestures of /r/ that are accessible to our techniques. The phonological evidence presented by McMahon *et al.* (1994), and in previous sections of the present paper, support the hypothesis that it is the tongue blade raising gesture of the /r/ that we should expect to categorise as a C-gesture. If this is true, then we must conclude that it is this gesture that is reduced in final positions resulting, in the most extreme cases, in complete *r*-vocalisation.

To test this, tokens of /r/ were collected from two speakers of American English (Subject 1: male, mid-20s, linguistically trained; Subject 2: female, mid-20s; both subjects are phonetically untrained and were unaware of the nature of the study at the time of data collection), to compare tongue tip height in initial allophones with that in final allophones. Stimuli consisted of sentences of the form 'I say _x_ again', where x is one of four minimally distinct nonsensical two-word combinations containing /r/ in initial position (*pa rotter* or *ha rotter*) or final position (*par hotter* or *har hotter*), with stress on the first word of each pair. /h/ was used to prevent resyllabification of final /r/ without interfering with oral articulation. The speaker read sentences aloud from written lists in his normal speech and at a comfortable pace. 10 tokens were collected for each position.

Data was collected using EMMA (electromagnetic midsagittal articulometer – see Perkell *et al.* 1992), a three-coil transmitter system at Haskins Labs. Small receivers were attached to speakers' tongue (tip, blade, body and dorsum), lips, mandible, maxilla and nose bridge. The latter two were used for correction of head movement in the midsagittal plane. Voltages induced in the receivers by three fixed electromagnets situated around the subject's head were used to determine location of the receivers in the midsagittal plane. Movement data was sampled at 500 Hz.

For both subjects, measurements were taken of the position of the tongue tip receiver at its maximum height and frontness during each production of /r/. Tongue tip positions were measured rather than tongue blade for different reasons for each speaker. It is well known that several

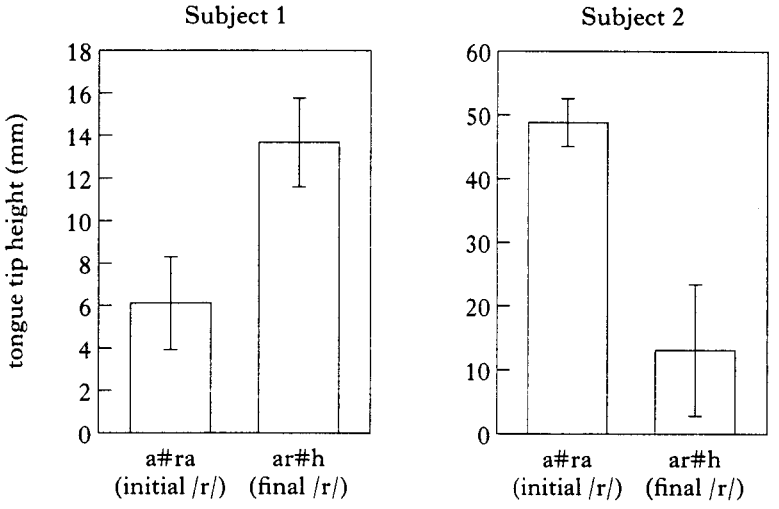

Figure 2
Mean maximum height of tongue tip during productions of /r/ in initial and final position

distinct techniques may be employed in the production of English /r/ (Delattre & Freeman 1968: 41, Hagiwara 1995, Alwan *et al.* 1997: 1079). Based on two-dimensional simulation of receiver movement using Haskins' MAVIS (Multiple Articulator Visualization) software, each speaker's technique became quite clear. Subject 1 appears to use what Hagiwara (1995: §§6.3–6.4) refers to as either the TIP DOWN or BLADE UP configuration, while for Subject 2, the tongue blade is preceded by the tip, resulting in Hagiwara's TIP UP pattern. Thus, for Subject 1, the most effective measure location for the tongue gesture is the tongue blade, while for Subject 2, the tongue tip provides the better metric. Unfortunately, the receiver used for the tongue blade for Subject 1 was corrupt, resulting in out-of-bounds measurements, and had to be discarded. For these reasons, the tongue tip measurements were used for both speakers, though for Subject 1 we expect to see lowering and fronting of this receiver as the blade increases in height, while for Subject 2 we expect raising.

Standard one-way analysis of variance (ANOVA) was used to test for significant variance between initial and final allophones. One initial token was excluded for Subject 1 because of ambiguous location of movement extremum.

Results of this experiment showed evidence for the final reduction effect for both subjects. As predicted, tongue tip positions in initial allophones of /r/ for Subject 1 were lower ($p < 0{\cdot}0001$) and more fronted ($p = 0{\cdot}0452$) than those in final allophones (frontness and height were negatively correlated; $r^2 = 0{\cdot}807$). For Subject 2, vertical tongue tip position was greater in initial allophones than in final allophones ($p < 0{\cdot}0001$). Mean

vertical magnitudes are shown for both subjects in Fig. 2 (error bars show 95% confidence interval). These results support the final reduction analysis of r vocalisation proposed by McMahon *et al.*

5.3 Merger

Since the present account considers intrusive consonants to be lexically present, they are undifferentiated from linking consonants. While this eliminates a great deal of the complication of most analyses, the resulting simpler model must still be able to account adequately for all of the facts which previous accounts have; and as was shown above for vocalisation, where possible, the phenomena should be motivated by processes already observed or needed elsewhere in the phonology.

McMahon *et al.* extend their hypothesis for /r/-vocalisation to include a subsequent merger with schwa (1994: 303–304). Drawing on the comparatively limited research on the articulation of /r/ (mainly Delattre & Freeman 1968), McMahon *et al.* analyse the /r/ as 'a constellation of two gestures... one in the region of the hard palate, and the other in the mid-pharynx' (1994: 303). They propose that 'the pharyngeal constriction component of the /r/ is articulatorily rather similar to the constriction for schwa',[4] concluding (1994: 304): 'if the palatal gesture was reducing in magnitude, then one would expect the pharyngeal gesture for the /r/ to be heard as a schwa'. This applies both to the underlying schwa (e.g. the final sound in *drama*) and to the schwa offglides in their account following /a/ and /ɔ/. Now, if indeed these claims regarding the articulatory similarity between /r/ and schwa prove to be true, they will fit with the prediction made previously in this paper that there must historically have been such a connection. Desirable as this may be, though, no direct physical evidence exists in support of this claim. Without the use of X-ray, or until the temporal resolution of MRI or other imaging methods is substantially improved, speech activity in the pharyngeal region will not be directly measurable (see e.g. Westbury *et al.* 1995: 56, Alwan *et al.* 1997). However, based on the similar vocalisation processes described in the previous section, the connection hypothesised by McMahon *et al.* between /r/ and schwa can now be looked for in /l/.

The suggestion made by McMahon *et al.*, restated, is that if the tongue raising gesture (C-gesture) of the /r/ were reduced sufficiently, the remaining pharyngeal constriction (presumably a V-gesture) would be perceptually sufficiently similar to schwa to precipitate a merger in some dialects. Recall that the same phonological connection was predicted between /l/ and /ɔ/ on the basis of the distributional evidence discussed

[4] In an early X-ray pilot film of American English vowels produced by researchers at Haskins Laboratories, a substantial upper pharyngeal retraction is apparent in both initial and final schwas for all four of the subjects filmed. This retraction is particularly clear in schwas following high vowels with advanced tongue root (as in the sequence [ítə, útə]). MRI data more specifically tailored to the testing of McMahon *et al.*'s proposal has been collected and is currently being processed by the author and collaborators.

in §3.2. If the tongue tip (C-gesture) of the /l/ were reduced or eliminated, the remaining vowel gesture would be a low back and, depending on the dialect, rounded vowel. Of all the vowels in the American English inventory, this result most resembles /ɔ/ (see note 4). Thus, assuming the articulatory structure of /r/ to be parallel to that of /l/, the proposed historical merger of /r/ with schwa follows in the same way. The fact that these correspondences fall out from processes already needed for other aspects of the phonology speaks strongly in favour of this type of analysis.

5.4 Intrusion: ambisyllabic consonants in gestural theory

At this point, the account of McMahon *et al.* diverges from the present one. Theirs, as most other analyses, posits *a priori* that the intrusive *r* is not lexically present, and they must therefore find a way to supply the missing gesture in those environments where it appears. Adopting a hybrid derivational approach to manipulate the units of AP, McMahon *et al.* conclude that the only way to get the apparent insertion of the missing consonantal gesture of /r/ is to allow for the wholesale insertion of gestures, if under certain restricted circumstances. These circumstances would be limited by licensing gestural insertion rules only at one level within a Lexical Phonology framework (see Zsiga 1993, McMahon & Foulkes 1995 and Tollfree 1996 for similar approaches). While this account may be superior to previous theories in accounting in the way described above for *r*-vocalisation and the subsequent historical merger with schwa, it ultimately falls prey to the same arbitrariness problem of McCarthy's analysis – i.e. the arbitrary insertion of a gesture requires as much extra theoretical apparatus as the insertion of any other phonological unit (McMahon *et al.* 1994: 300).

In defence of the present theory, what must now be demonstrated is that uncontroversially underlying consonants independently show the same patterns of reduction and augmentation observed in intrusive consonants, If so, then no further apparatus needs to be invoked to account for these patterns. In the past, such a hypothesis would have been untestable, both because of the limitations of speech science technology, and because *r* was the only intrusive consonant known in the literature. However, as it has been shown above that other segments – *l*, and probably the glides as well – follow the same patterns as intrusive *r*, and since the articulation of these segments has recently begun to be better understood, it is now possible to propose tests for the applicability of the proposed model to consonantal intrusion in general. In the previous two sections, it was demonstrated that known principles of articulatory organisation account for both the vocalisation of liquids and glides and their subsequent merger with final vowels, without the need for any new theoretical apparatus. Extending this account to intrusion, if all of the patterns associated with intrusive consonants are predictable from the model, and can be shown to be consistent with those of underlying elements, this will speak strongly in favour of the view of intrusive consonants as lexically present.

Before we can look for behaviour like that of intrusive consonants, however, we must first determine specifically where to look and what to look for. First, where does intrusion arise? If an intrusive consonant is underlyingly present, it must be a final element, the presence of which is obscured in certain environments due to vocalisation. Its presence becomes apparent only when it is followed by a vowel-initial word or morpheme. Thus, if we are to compare it to other underlying consonants which are not fully obscured by the process of vocalisation, these other consonants must be word-final, they must be composite segments, and they must be followed by vowel-initial words. Such segments have often been referred to as AMBISYLLABIC in the phonological literature (e.g. Kahn 1976), referring to the theoretical notion that they may be affiliated with both the preceding and following syllables.

Second, how does sensitivity to a following vowel affect the actual production of a final composite segment? An acoustic analysis of intrusive *r* (McCarthy 1993: 179) found the intrusive *r* in *saw*[r] *eels* to be 'considerably more vocalic, with more energy at all frequencies' than the initial allophone in *saw reels*. Acoustic studies of the ambisyllabic allophones of /l/ (Lehiste 1962, Bladon & Al-Bamerni 1976) report similar results, indicating that these allophones fall between initial and final allophones in terms of their acoustic darkness. The proposed model should represent the effect of a following vowel on a final allophone in a way consistent with these findings. We know from the discussion of final reduction in §5.1.1 that the C-gestures of composite segments in final position are reduced in magnitude compared with initial allophones. It has also been proposed that these variations in magnitude are determined by phasing relations with a preceding or following vowel. If indeed the segments in question are ambisyllabic (phased simultaneously to both preceding and following vowels) – a supposition supported by the acoustic studies cited above – we should expect to find that the magnitude of the C-gestures in this position will be intermediate between those of initial and final allophones.

5.4.1 *Evidence from* /w/ *and* /l/. In a recent EMMA experiment, Gick (in press) looked at both /w/ and /l/ (the two composite glide and liquid segments of English for which all gestures have been measured) in non-intruding dialects of American English. Initial allophones in nonsense word pairs, such as the /l/ in *pa lotter* or the /w/ in *pa wadder*, were compared with final allophones, as in *pall hotter* and *pow hotter*. Examples of final allophones were also measured without the intervening *h*, as in *pall otter* and *pow otter* (in the dialects of the speakers, all of these sets are minimally distinct). /h/ was used to fill hiatus without interfering with oral articulations. In the absence of this hiatus-filler, the final /l/ or /w/ was presumed to be ambisyllabic. /w/ is the ideal object for study because of the unique characteristic that both of its component gestures are approximants, allowing gradient effects of magnitude to be captured that might be obscured in a complete closure (a problem with measuring the

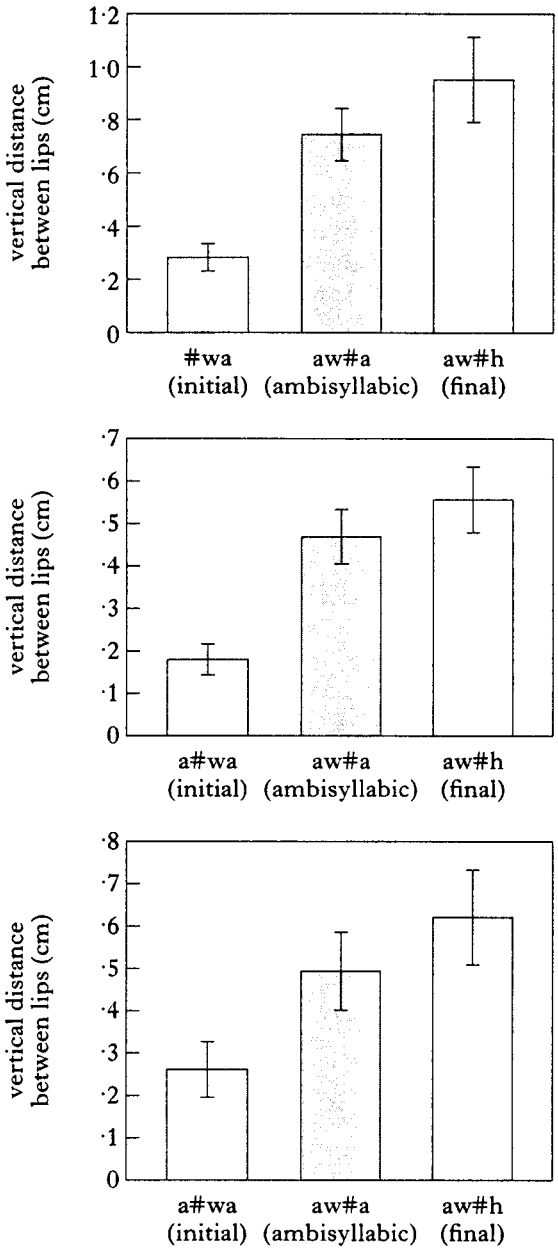

Figure 3
Mean magnitude of lip gesture of /w/ in initial, ambisyllabic and final allophones for three speakers (from Gick, in press)

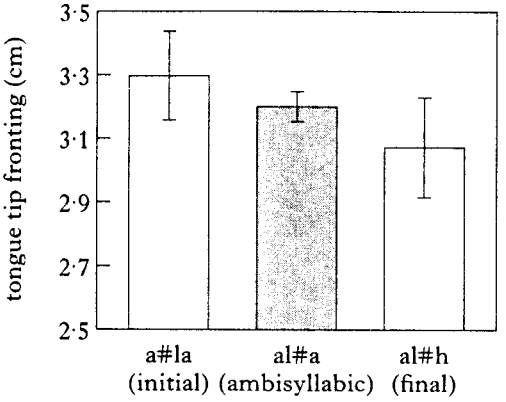

Figure 4
Mean magnitude of tongue tip gesture for /l/ in initial, ambisyllabic and final allophones (from Gick, in press)

tongue tip closure of /l/). Results for /w/ for three speakers are shown in Fig. 3.

In the same study, measurements were also made of the tongue tip (C-gesture) for one speaker's production of /l/, shown in Fig. 4. The effect is less prominent for /l/ than for /w/, possibly because the C-gesture of /l/ is a complete closure, and variations in constriction must therefore be very small. Gick also reports measurements of V-gestures (tongue dorsum) for /w/ and /l/, finding, as predicted, that they are not sensitive to following context.

5.4.2 *Evidence from /r/: experiment 2.* In §5.2.2, initial and final allophones of /r/ were compared for two subjects. Using the same methods described there, tongue tip height during /r/ was measured in ambisyllabic position, allowing comparison of all three allophones (*pa rotter*, *par hotter* and *par otter*). Results in Figs 5 and 6 show that the intermediate effect seen in the C-gestures of ambisyllabic allophones of /w/ and /l/ is also present in /r/. Fig. 5a shows the effect in tongue tip height for Subject 1 (post hoc tests indicate $p < 0{\cdot}02$ for all contrasts), while Fig. 5b graphs both mean height and fronting maxima for the same subject, illustrating that the relationship between lowering and fronting is maintained (the difference in frontness is significant for initial *vs.* final allophones, but not between these and the ambisyllabic allophone). Figure 6 shows the tongue tip height effect for Subject 2 (post hocs show $p < 0{\cdot}02$ for all contrasts).

All of these results support the hypothesis that the C-gestures, and only the C-gestures, of composite segments of English have a greater magnitude (are more tightly constricted) in initial allophones, a significantly reduced magnitude in final allophones and an intermediate magnitude in inter-vocalic allophones. This accounts for the strong initial allophones,

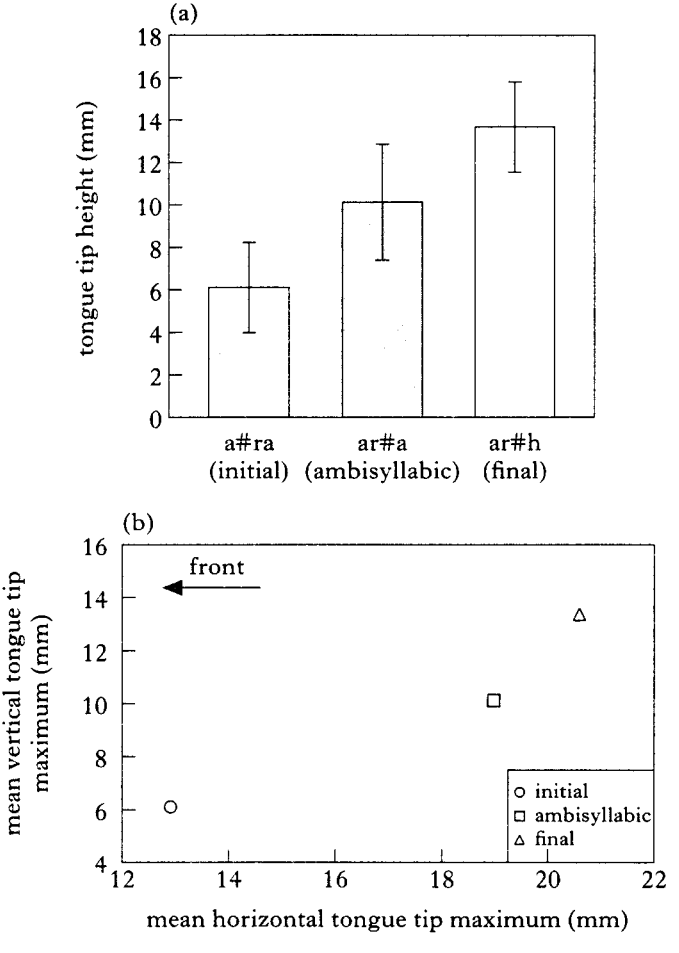

Figure 5
(a) Mean tongue tip height for Subject 1 for /r/ in initial, ambisyllabic and final allophones; (b) Mean tongue tip height and frontness for /r/ in initial, ambisyllabic and final allophones

reduced (sometimes to perceptual absence) final allophones and intermediate ambisyllabic (linking or intrusive) allophones observed to varying degrees in lexically present composite segments in all dialects of English. Thus, we conclude that linking and intrusion are simply extreme versions of the same pattern seen in non-intruding dialects. It is only in those dialects where final /r/ never surfaces (such as the southern U.S. and African American vernacular dialects described in §2.1.2) that the historical /r/ must be said to be lexically absent. In this way, both the historical development and the synchronic behaviour associated with

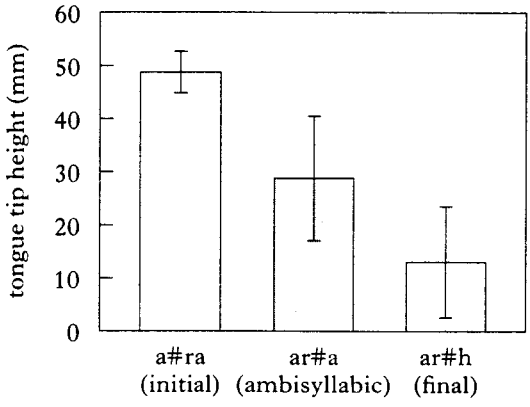

Figure 6
Mean magnitude of tongue tip gesture for Subject 2 for /r/ in initial, ambisyllabic and final allophones

intrusion fall out of a simple model based on previously attested properties of gestures.

5.5 Summary of the gesture-based account of intrusion

To summarise the account of intrusion described in this section, vocalisation (final reduction plus temporal lag) applies to C-gestures in final allophones in all dialects of English thus observed, in some cases to the point where the gestures are perceptually obscured. Even in cases where the segment is perceptible in final positions (as with General American /r/, /l/ or /w/), final allophones followed by a vowel (ambisyllabic allophones) will show less reduction.

Historically, in cases of complete perceptual vocalisation, if the remaining V-gesture sounded sufficiently similar to another vowel in the system, a merger (neutralisation) took place in many dialects between forms containing final vocalised /r/ or /l/ and those containing final schwa or /ɔ/, respectively. In some dialects, however (such as South African English; see §2.1.3), despite substantial vocalisation, no merger took place. In a dialect having the merger, however, a child learning the language could interpret the resulting final vowel in one of two ways: either as an underlying /r/ or /l/ in all cases, resulting in both linking and intrusion (as in RP and eastern Massachusetts), or as an underlying vowel, resulting in neither linking nor intrusion (as in southern American and Black American Vernacular dialects; see §2.1.2). The fact that all three of these types of dialects exist, but other logical possibilities do not (e.g. intrusion in non-vocalising dialects, intrusion in non-linking dialects), provides an additional typological argument in favour of this analysis.

Aside from this merger and the resulting cross-dialectal typology, the so-called intrusive consonants are uninteresting, being only a lexical /r/ or

/l/ following the gestural patterns of any composite segment of English. The glides and nasals have not been claimed to have undergone similar mergers, but otherwise follow the same patterns of magnitude and timing in initial, final and ambisyllabic positions. This account is consistent with Harris's (1994: ch. 5) 'floating *r*' account, in which he states that historically schwa-final words are 'in fact *r*-final, even though the consonant is lexically floating and thus not always phonetically interpreted'. The important difference between Harris's account and the present one is that the experimental data presented above and elsewhere shows that no special phonological status is needed for *r*, *l* or glides in order to get the behaviour of intrusive *r*, intrusive *l* or glide insertion.

5.6 Other insertions

Now that we have a clear story for intrusion, how does the present model deal with the other types of apparent *r*- and *l*-epenthesis discussed in §§2.1.4 and 3.1.1?

5.6.1 *Epithesis.* Unlike most analyses of intrusive *r*, the present account does not require that the non-historical *r* or *l* necessarily arises in intervocalic positions, as is required by analyses where the intruded element is argued to arise out of a need for an onset to the following syllable, or as a hiatus filler. For, in the present account, there are no limitations on what the timing and magnitude properties of a particular dialect must be, simply that they must apply consistently across all C- and V-gestures within that dialect. Thus, cases of epithesis such as the Bristol *l* (where *tango* and *tangle* both surface with final [l]) and the numerous dialects containing examples of final epithetic *r* (see §2.1.4) fall out as naturally from the present account as do the intervocalic intrusive cases. Furthermore, this model predicts that the different practices of timing or magnitude in epithetic dialects should be reflected elsewhere in the phonology. Thus, for example, such a dialect should not fully vocalise postvocalic glides. Interestingly, Bristol not only has *l*-intrusion, but is also one of the southern English dialects where final *r* is pronounced.

5.6.2 *Internal epenthesis: gestural overlap.* Internal *r*-epenthesis of the American *wa*[r]*sh* variety, on the other hand, seems to belong to a different process altogether. It does not apply consistently following certain vowels, nor does it apply at any kind of a syllable, morpheme or word boundary. Rather, it seems to appear only in the very specific environment /a, (ɔ) __ ʃ/. This distribution suggests a different analysis, for which a gestural representation is also particularly well suited. Browman & Goldstein (1990) discuss examples such as English *tense*, which often arises with an epenthetic, or excrescent, [t], creating mergers in many dialects of such pairs as *tense* and *tents*, *presence* and *presents*, and so on. They account for these as simple cases of gestural overlap, whereby the vocal fold spreading for the voiceless /s/ begins before release of the tongue tip closure for the /n/, resulting in the percept of a /t/ (a voiceless

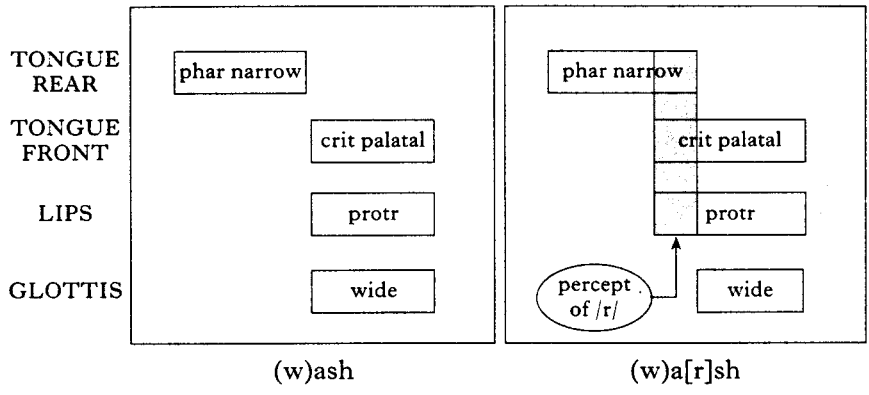

Figure 7
Gestural score of *wash* and *wa*[r]*sh*

tongue tip closure and release). The *wa*[r]*sh* cases can be given a similar account: since [r] is composed of pharyngeal, tongue blade and probably lip gestures, any time these three gestures happen simultaneously, an [r] is predicted to result. We already know from our previous discussion that /a/ and /ɔ/ have a pharyngeal component. Now, what of the gestures of /ʃ/? This consonant consists of both tongue blade raising and lip rounding. Thus, in any dialect where there is an audible overlap between the vowel (or its schwa offglide) and the oral gestures of the /ʃ/, we should expect the percept of an /r/, and this is precisely what we do get. Fig. 7 shows a GESTURAL SCORE illustrating the timing involved.

This same gestural overlap analysis can be given for other examples of non-historical *r*'s. Louis Goldstein (personal communication) suggests a similar analysis for pronunciations such as *oyster* [ɔ˞stɔ˞] and *boil* [bɔ˞l]. Thus, if the temporal coordination of the component gestures of the diphthong [ɔj] is slightly altered in a particular dialect so that all three gestures overlap (the upper pharyngeal and lip gestures of [ɔ] and the tongue blade raising of [j]), then the same perceptual [r] as in *wa*[r]*sh* will result. Again, as with epithesis, the presence of this kind of overlap in a dialect predicts that there will be other evidence of the same timing tendencies elsewhere in the same dialect, allowing for further testing of this hypothesis.

6 Conclusion

This paper has attempted to demonstrate how previously observed principles of gestural organisation – final reduction, position-specific timing and overlap – can provide a simple and unified account for a variety of phenomena involving the liquids and glides of English. The various historical and synchronic facts and recent experimental findings reported

on in this and other papers combine to support the hypothesis that these phenomena, including intrusion, glide formation and other cases of epenthesis, are simply the audible results of patterned variations in timing and magnitude affecting all composite segments of English.

REFERENCES

Alwan, Abeer A., Shrikanth S. Narayanan & Katherine Haker (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II: the rhotics. *JASA* **101**. 1078–1089.

Anderson, S. R. (1995). Intra-syllabic gestural relations. Paper presented at Haskins Laboratories.

Ash, Sharon (1982a). *The vocalization of /l/ in Philadelphia*. PhD dissertation, University of Pennsylvania.

Ash, Sharon (1982b). The vocalization of intervocalic /l/ in Philadelphia. *The SECOL Review* **6**. 162–175.

Bailey, Charles-James N. (1985). *English phonetic transcription*. Dallas: Summer Institute of Linguistics & University of Texas at Arlington.

Bladon, R. A. W. & A. Al-Bamerni (1976). Coarticulation resistance in English /l/. *JPh* **4**. 137–150.

Broadbent, J. (1991). Linking and intrusive r in English. *UCL Working Papers in Linguistics* **3**. 281–302.

Browman, C. P. & L. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* **3**. 219–252.

Browman, C. P. & L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* **6**. 201–251.

Browman, C. P. & L. Goldstein (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 341–376.

Browman, C. P. & L. Goldstein (1992). Articulatory phonology: an overview. *Phonetica* **49**. 155–180.

Browman, C. P. & L. Goldstein (1995). Gestural syllable position effects in American English. In F. Bell-Berti & L. J. Raphael (eds.) *Producing speech: contemporary issues. For Katherine Safford Harris*. Woodbury, NY: American Institute of Physics. 19–34.

Clements, G. N. (1987). Phonological feature representation and the description of intrusive stops. *CLS* **23:2**. 29–50.

Delattre, P. & D. C. Freeman (1968). A dialect study of American *r*'s by x-ray motion picture. *Linguistics* **44**. 29–68.

Fasold, R. W. (1981). The relation between black and white speech in the South. *American Speech* **56**. 163–189.

Gick, B. (1991). A phonologically motivated theory of consonantal intrusion and related phenomena in English. Ms, University of Edinburgh.

Gick, B. (1997). The intrusive l. Paper presented at the Annual Meeting of the American Dialect Society, Chicago.

Gick, B. (in press). Articulatory correlates of ambisyllabicity in English glides and liquids. In Local *et al.* (in press).

Gick, B. (in preparation). The American intrusive l. Ms, Yale University.

Giegerich, Heinz (1997). The phonology of '/ɔ:/' and '/ɑ:/' in RP English: Henry Sweet and after. *English Language and Linguistics* **1**. 25–47.

Giles, Stephen B. & Kenneth L. Moll (1975). Cinefluorographic study of selected allophones of English /l/. *Phonetica* **31**. 206–227.