Perception of intrusive /r/ in English by native, cross-language and cross-dialect listeners

Annelie Tuinman and RSNHolger Mitterer and RSNAnne Cutler and RSN

Citation: The Journal of the Acoustical Society of America **130**, 1643 (2011); doi: 10.1121/1.3619793
View online: http://dx.doi.org/10.1121/1.3619793
View Table of Contents: http://asa.scitation.org/toc/jas/130/3
Published by the Acoustical Society of America

Perception of intrusive /r/ in English by native, cross-language and cross-dialect listeners

Annelie Tuinman[a)]
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

Holger Mitterer
Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

Anne Cutler[a),b)]
MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith South, NSW 2751, Australia

(Received 18 October 2009; revised 5 May 2011; accepted 3 July 2011)

In sequences such as *law and order*, speakers of British English often insert /r/ between *law* and *and*. Acoustic analyses revealed such "intrusive" /r/ to be significantly shorter than canonical /r/. In a 2AFC experiment, native listeners heard British English sentences in which /r/ duration was manipulated across a word boundary [e.g., *saw (r)ice*], and orthographic and semantic factors were varied. These listeners responded categorically on the basis of acoustic evidence for /r/ alone, reporting *ice* after short /r/s, *rice* after long /r/s; orthographic and semantic factors had no effect. Dutch listeners proficient in English who heard the same materials relied less on durational cues than the native listeners, and were affected by both orthography and semantic bias. American English listeners produced intermediate responses to the same materials, being sensitive to duration (less so than native, more so than Dutch listeners), and to orthography (less so than the Dutch), but insensitive to the semantic manipulation. Listeners from language communities without common use of intrusive /r/ may thus interpret intrusive /r/ as canonical /r/, with a language difference increasing this propensity more than a dialect difference. Native listeners, however, efficiently distinguish intrusive from canonical /r/ by exploiting the relevant acoustic variation. *© 2011 Acoustical Society of America.* [DOI: 10.1121/1.3619793]

PACS number(s): 43.71.Hw, 43.71.Es [RSN] Pages: 1643–1652

I. INTRODUCTION

Margaret Thatcher's legendary nickname *Laura Norder* not only reflected her political preferences, but also the fact that, like most of her compatriots, she pronounces *law and order* with an intrusive /r/ between *law* and *and*. Most British English dialects are nonrhotic, i.e., they have the phonotactic constraint whereby /r/ can occur in word onsets but not at the end of words; thus the /r/ in *real* /ril/ is pronounced but the /r/ in a citation-form utterance of *hear* /hɪə/ is not, even though the spelling of *hear* ends with the letter *r*.

A pronounced word-final /r/ does appear in British English, for instance when a word such as *hear* precedes another word beginning with a vowel (e.g., *hear it* is spoken as /hɪərɪt/). This is known as "linking /r/." A pronounced /r/ also appears in the similar effect known as "intrusive /r/," whereby /r/ may be inserted after a nonhigh vowel (e.g., ə, a, ɔ, or diphthongs ending in [ə]) and before a vowel-initial word (Giegerich, 1992; Cruttenden and Gimson, 1994). Both of these expressions of /r/ result from so-called connected speech processes. As the *law and order* case illustrates, an intrusive /r/ can be reinterpreted as a syllable onset (the default realization of /r/ in British English). This is more likely to happen with intrusive than with linking /r/, given that intrusive /r/ has no reflection in spelling, whereas the spelling offers a source for a spoken linking /r/.

Connected speech processes feature in all languages, as far as is known, and their effects have not been ignored by speech perception research. It is clear from many recent studies that native listeners are able to derive the correct interpretations of utterances in which such processes have in some way altered the canonical pronunciation that a word sequence might otherwise receive (e.g., Spinelli *et al.*, 2003; Mitterer *et al.*, 2006a; Mitterer and Ernestus, 2006; Connine *et al.*, 2008, for liaison in French, assimilation in Hungarian, /t/ reduction in Dutch, and schwa deletion in American English, respectively). In some cases it has been established which acoustic cues are exploited to this end (e.g., liaison consonants are systematically shorter than canonically realized consonants, and listeners base interpretations of otherwise ambiguous sequences such as *trop artisan* versus *trop partisan* on this durational difference; Spinelli *et al.*, 2003).

No such evidence is as yet available for the case of intrusive /r/, however, and the first aim of the current study is to redress this omission. A prior report suggests that intrusive and canonically pronounced /r/ differ acoustically (Cruttenden and Gimson, 1994); we test for the presence of such acoustic

a)Also at: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.
b)Author to whom correspondence should be addressed. Electronic mail: Anne.Cutler@mpi.nl

differences in production, and examine in our first perceptual experiment whether native listeners make use of these acoustic parameters in interpreting a realized /r/ as intrusive (*saw ice*) versus canonical (*saw rice).*

In a second perceptual test we present the same acoustically manipulated materials to proficient users of English as a second language (L2), whose exposure to English is predominantly to the British variety, but in whose native language, Dutch, /r/ is pronounced word-finally and no connected speech process resembles intrusive /r/. Finally, our third perception test examines the responses, again to the same materials, of speakers of American English, i.e., native speakers of the target language English, who will have been exposed to varying varieties of this language, including some nonrhotic varieties with intrusive /r/, but who are not themselves native users of the British variety.

The processing of speech input by L2 listeners has been widely researched, with particular attention to the consequences of inventory differences between languages (e.g., Polka, 1995; Guion *et al.,* 2000; Best *et al.,* 2001); without doubt the most well-known and well-studied such inventory mismatch is the English /r/-/l/ contrast for Japanese listeners, (e.g., Underbakke *et al.,* 1988; Bradlow *et al.,* 1997; Ingram and Park, 1998, Cutler *et al.,* 2006). Speech processing by listeners with a different dialect of the same language has received somewhat less research attention, but recent studies have documented challenges presented by this less radical mismatch also. Sumner and Samuel (2009), for instance, found that speakers of a General American dialect can have problems in recognizing words produced with a New York accent, Floccia *et al.* (2006) likewise found an initial processing cost for a different dialect in word recognition tasks, and Otake and Cutler (1999) found that cross-dialect perception of Japanese words exhibited lower sensitivity to information in the signal (d') and a higher degree of bias (β) toward lexical knowledge. Cross-dialect difficulties in phoneme perception, however, seem to be of lesser magnitude than cross-language difficulties. Cutler *et al.* (2005) asked Australian, American, and Dutch listeners to identify American English vowels in meaningless CV and VC syllables, and found that overall, the Australian and American listeners performed equally well, although the Australians were systematically affected by the tendency to greater vowel tenseness in their native dialect (reporting, for example, /ɔ/ as /ɑ/ more than vice versa). The Dutch listeners' overall performance, however, was significantly worse, suggesting that language differences have greater consequences than dialect differences for the perception of phonemes.

In neither the cross-language nor the cross-dialect case has there been substantial attention to the processing of connected speech phenomena. We found one cross-dialect study of cues to syntactic structure, namely, palatalization and intervocalic flapping in American English (Scott and Cutler, 1984), and one cross-language study of a word-level process in French (Darcy *et al.,* 2007). In the former study, British English listeners failed to use cues used by American English listeners, while in the latter study, compensation for consonant-to-consonant voicing assimilation—which occurs in French, but not in English—was observed in native French listeners and also, to a lesser degree, in English learners of French (note, however, that general auditory processes contribute to this compensation; Mitterer *et al.,* 2006a,b).

The present study, then, addresses the /r/-insertion found in nonrhotic British English and how it is perceived, by native listeners and by listeners with another language or another dialect. This insertion process induces acoustic evidence for /r/ that cannot be attributed to an underlying or orthographic representation. On the basis of the above literature summary we can make certain, albeit cautious, predictions of our likely results.

First, studies of native listeners' processing of other connected speech phenomena suggest that our British English listeners should be able to make effective use of whatever acoustic cues there are that distinguish intrusive from canonical /r/. That is, if we succeed in establishing the presence of the acoustic differences predicted by Cruttenden and Gimson (1994), and manipulate these differences along a continuum from one type of /r/ to the other, the native listeners' responses should essentially track our manipulation.

Second, based on L2 listening studies with phonetic segments, we predict that our cross-language listeners will not succeed in matching this native sensitivity. The /r/ insertion process is, as noted, unknown in Dutch. Dutch dialects exhibit conspicuous variation in the way /r/ is produced (van de Velde, 1996; van Bezooijen, 2005), but no dialect is nonrhotic in the way British English is. There is thus no scope in Dutch for either linking /r/ or intrusive /r/. In the three cases *meer* ("more"), *meer tijd* ("more time") and *meer appels* ("more apples") the final /r/ in *meer* is pronounced, and in no variety of Dutch is /r/ ever inserted in contexts such as *na appels* "after apples" (Giegerich, 1992; Gussenhoven and Broeders, 1997; Collins and Mees, 1999). For these listeners we predict that exposure will be, as in the phoneme perception cases, of little relevance. The target English pronunciation taught in Dutch schools and universities is actually British English (see, e.g., Gussenhoven and Broeders, 1997; Collins and Mees, 1999), and British radio and television can be received in all households; note further that foreign-language productions shown on Dutch television channels are subtitled, never dubbed. Despite this opportunity for wide exposure, Dutch listeners consistently fail to distinguish minimal word pairs in English as a result of phoneme inventory mismatches (Weber and Cutler, 2004; Broersma and Cutler, 2011), and we predict that the /r/-insertion process will also mismatch with the L1 expectations to an extent that these listeners' identification responses will differ significantly from those of the native listeners.

Third, on the basis of the existing small sample of interdialectal studies we predict that the cross-dialect listeners should at least outperform the cross-language listeners. As noted above, vowel inventory mismatches across dialects have a lesser effect on vowel identification than mismatches across languages, and the word processing effects reported by Sumner and Samuel (2009), Floccia *et al.* (2006) and Otake and Cutler (1999) were all of lesser magnitude than the comparable cross-language interference effects (e.g., Ingram and Park, 1998; Broersma and Cutler, 2011); further, the cross-dialect disadvantage in the Floccia *et al.* study was in fact only temporary.

Moreover, opportunity for exposure may be more relevant in the case of varieties of the same language than in the case of different languages, since there is evidence for rapid perceptual adaptation across varieties (Evans and Iverson, 2004). Our American listeners were tested in Philadelphia, where a rhotic variety is the norm. Most dialects of American English are rhotic (i.e., the /r/ in *more* is always pronounced in the English translation equivalents of the above Dutch utterances: *more, more time, more apple*s). These varieties are unlikely to exhibit intrusive /r/ even in casual speech, as a complete search of the Buckeye corpus of conversational speech (Pitt *et al.,* 2007) revealed. We isolated from this corpus all 4698 instances in which /r/-insertion would be possible between a word ending in a non-high vowel followed by a vowel-initial word (the licit context for the British English process). In none of these is there /r/-insertion. In three cases, /r/ is transcribed at the word boundary, but this always stems from the words involved rather than from insertion (e.g., "camera either" produced as "camreither").

However, some US dialects are nonrhotic (these are said to show "r-dropping" i.e., are distinguished from the rhotic majority, while British varieties that are rhotic are likewise distinguished from the local nonrhotic majority by being referred to as "r-pronouncing," e.g., Wolfram and Schilling-Estes, 2006). These would show a British-like pattern whereby *more* would have no /r/, *more apples* a (linking) /r/, and *saw apples* possibly an intrusive /r/. Note that even the Boston accent, perhaps the most widely referred-to nonrhotic US variety, is moving toward rhoticity (Wells, 1982; Trudgill, 1986; Irwin and Nagy, 2007). Nevertheless, American listeners are likely to have heard nonrhotic varieties both of American English, and, via the media at least, of British English. This may even suffice to support performance parallel to that of the British native listeners.

At issue in our study is the extent to which listeners report the presence of a phonetic segment /r/ in materials varying in amount of acoustic evidence for intrusive versus canonical /r/ at a word boundary. Given that we predict that at least one listener group could have difficulty exploiting our acoustic manipulation, we incorporate further variation in the materials as a basis for comparison with the acoustic /r/ evidence. Previous reports suggest that non-native listeners may rely to a greater extent than native listeners on both orthographic information (Escudero *et al.,* 2008) and semantic information (Bradlow and Alexander, 2007). Thus we vary orthographic support for a final /r/ (i.e., we contrast the word *more* versus the word *saw* before the boundary) and semantic support for a canonical /r/ interpretation (we compare a context weakly favoring *rice* after the boundary with a context offering no such support). For each listener group we then compare whether their responses are based on acoustic evidence alone, or take account of orthographic and semantic factors as well.

II. ACOUSTIC ANALYSES

A. Materials

We created 27 pairs of English sentences contrasting an onset and an intrusive /r/. An example sentence is *"And then Emma (r) ejected the cassette*." In all pairs, a member of a minimal pair such as *eject/reject* followed a word ending on a low vowel (e.g., *Emma).* Trivially, the r-initial member of the pair in the sentence will trigger pronunciation of an /r/. More importantly, the vowel-initial member of the pair together with the preceding low vowel (in this case, the last vowel of *Emma*) creates a context in which intrusive /r/ can appear. The sentences are listed in the Appendix.

A list was constructed in which all sentences occurred, in random order and, to obscure the purpose of the study, interspersed among 208 filler sentences. The list was recorded by a female native speaker of British English from the London area who normally produces intrusive /r/s in her casual speech. Each crucial sentence with /r/ was recorded at least twice. Further repetition was required for any token produced with hesitation. The complete data set for acoustic analysis contained 127 tokens: 72 sentences with an /r/-initial word, 55 with a vowel-initial word. One /r/-initial word was preceded by an intake of breath and was not measured. Of the 55 tokens with vowel-initial words, three contained an intervocalic glottalization. The remaining 52 contained an intrusive /r/.

B. Measurements and discussion

We measured both the duration of each intrusive or onset /r/, and the decrease in intensity from the vowel preceding /r/ to the lowest point in the /r/. All /r/ tokens were measured by the first author and one token of each sentence was also measured by the second author as a reliability check. The correlation between the two measurements was high (duration: $r = 0.72$; intensity difference: $r = 0.95$). All measurements are listed in the Appendix.

Both predicted differences appeared in this speaker's productions. Overall, onset /r/s were longer [$F(1,124) = 8.74$, $p < 0.01$]; the measured onset /r/s were on average 89 ms, while the intrusive /r/s averaged 69 ms. Onset /r/s also displayed a larger intensity decrement from the preceding vowel to the lowest point [$F(1,124) = 10.79$, $p < 0.01$]; the mean intensity decrement for onset /r/s was 7.9 dB, and for intrusive /r/s 2.2 dB.

To ascertain the potential usefulness of these patterns for listeners, we calculated a power estimate, the Cohen's d difference score (mean difference divided by standard deviation; Cohen, 1992). Values of Cohen's d above 0.8 are held to indicate a large effect size. Cohen's d for the durational difference was 1.6, and for the intensity decrement 1.9. Thus these measured differences between onset and intrusive /r/s have an effect size that is sufficiently large to support potential perceptual use by listeners. Whether listeners—native and non-native—can indeed distinguish between the two types of /r/ on the basis of acoustic evidence was then tested in perception experiments, using a 2AFC task.

III. PERCEPTION EXPERIMENTS

A. Participants

The three perception experiments involved respectively 18 native speakers of British English recruited from the participant pool of the Laboratory of Experimental Psychology of the University of Sussex, 18 native speakers of Dutch from the Max Planck Institute participant pool, and 14 native speakers of American English from the participant pool of the Institute for Research in Cognitive Science of the University of Pennsylvania. None reported any hearing impairment. All were volunteers and were paid a small fee for their participation. The Dutch participants had a high level of proficiency in English as a second language; on average, they had received seven years of English instruction in primary and secondary education.

B. Materials

Four experimental sentence frames (see Table I) were constructed, crossing an orthographic bias with a semantic bias for the perception of /r/. In each sentence token presented in the perception tests, listeners' task was to judge whether they heard *ice* or *rice*.

The orthographic factor contrasted the words *saw* versus *more* preceding the target word *(r)ice*. As the phrase *more ice* includes an /r/ in the spelling, a perceived /r/ in the speech signal can be attributed to *more,* while in the case of *saw ice,* an /r/ sound cannot be ascribed to spelling. An orthographic bias should therefore manifest itself in terms of more reports of *rice* after *saw* than after *more*.

The semantic manipulation contrasted sentence frames with the context *the social worker* and *given to the poor* versus frames with *the little girl* and *given to her brother*. To ascertain the extent of the bias, 15 native speakers of Dutch and 16 native speakers of English (5 Australasian, 3 British, 8 US) rated the acceptability of the words *ice* and *rice* in four (written) sentence contexts, on a seven-point scale. The contexts were constructed by crossing the two frames with the two adjectives *extra* or *more* before the critical word. An analysis of variance (ANOVA) with sentence subject (*girl, social worker*), adjective (*extra, more*) and object (*ice, rice*) as independent variables revealed similar patterns for the two language groups. For neither group was there a significant effect of adjective ($F < 1$ for Dutch speakers, $F\ [1,15] = 4.1$, $p > 0.05$ for English speakers), nor did this factor interact with other factors (Fs $[1,14] < 3$, $pMin > 0.1$; Fs $[1,15] < 4$, $pMin = 0.07$); that is, the ratings were unaffected by whether the sentence referred to *extra* or *more (r)ice*. However, there was for both groups a significant interaction of the subject and object factors ($F\ [1,14] = 74.1$, $p < 0.001$; $F\ [1,15] = 27.3$, $p < 0.001$), with higher ratings for *rice* (6.5 for Dutch speakers, 6.6 for English speakers) than for *ice* (2.9, 3.8) when *social worker* was the subject, but no significant preference for either object (*rice:* 5.4, *ice:* 5.9 for Dutch speakers, *rice:* 6.1, *ice:* 5.7 for English speakers) when *girl* was the subject. The *social worker* sentences were thus assumed to be somewhat biased to *rice* rather than *ice,* while the *little girl* sentences were unbiased. A cross-group analysis revealed the semantic bias effect to be stronger for the Dutch than for the English group ($F\ [1,29] = 6.5$, $p < 0.05$).

Although our acoustic measurements showed that both the duration of /r/ and the intensity drop from vowel to /r/ differed across intrusive versus canonical cases, intensity and duration trade off for perception of events shorter than 200 ms (Coren *et al.*, 1994, pp. 234–235). This means that within this range, altering duration of an intensity drop has the same perceptual effect as altering the extent of the intensity drop. For this reason, we chose to vary our materials along a single parameter, namely, duration.

To implement the acoustic manipulation, we constructed our sentences using MBROLA (Dutoit *et al.*, 1996), a speech synthesizer based on diphone concatenation. The same female native speaker of British English who produced the materials for the acoustic measurements served as a model. This speaker first recorded the four sentence frames in Table I several times. The most fluent recordings without hesitations were chosen as the model sentences. The durations of all segments and the course of the pitch contour of each sentence were measured. These measurements were then used to create synthetic tokens of the same sentences, with MBROLA taking as input the list of phonemes plus prosodic information (phoneme durations, pitch contour) to produce speech samples. Four native British listeners then judged these output tokens and further improved them by adjusting the phoneme durations and pitch contour to render them even more natural. Final versions were then created in which all constant portions were identical in duration across the four sentences. Thus every occurrence of *saw* had exactly the same duration in each sentence, and so did every occurrence of *more*. Similarly, the duration of the phonemes /ai/ and /s/ was constant in all versions of *(r)ice*. The only non-constant parameter was the duration of /r/; this varied from 25 ms to 121 ms in seven quadratic steps (see Fig. 1). This range of durations ran from somewhat longer than the shortest of our speaker's measured tokens of intrusive /r/ at 10 ms, to somewhat shorter than her longest measured canonical /r/ at 128 ms.

C. Design and procedure

There were three independent variables: (1) the duration of /r/ at the critical word boundary; (2) the orthographic

TABLE I. Sentence materials.

	Orthography	
Semantic Context	*Ice* bias	*Rice* bias
Less *rice* bias	The little girl saw more ice/rice was given to her brother	The little girl saw ice/rice was given to her brother
More *rice* bias	The social worker saw more ice/rice was given to the poor	The social worker saw ice/rice was given to the poor

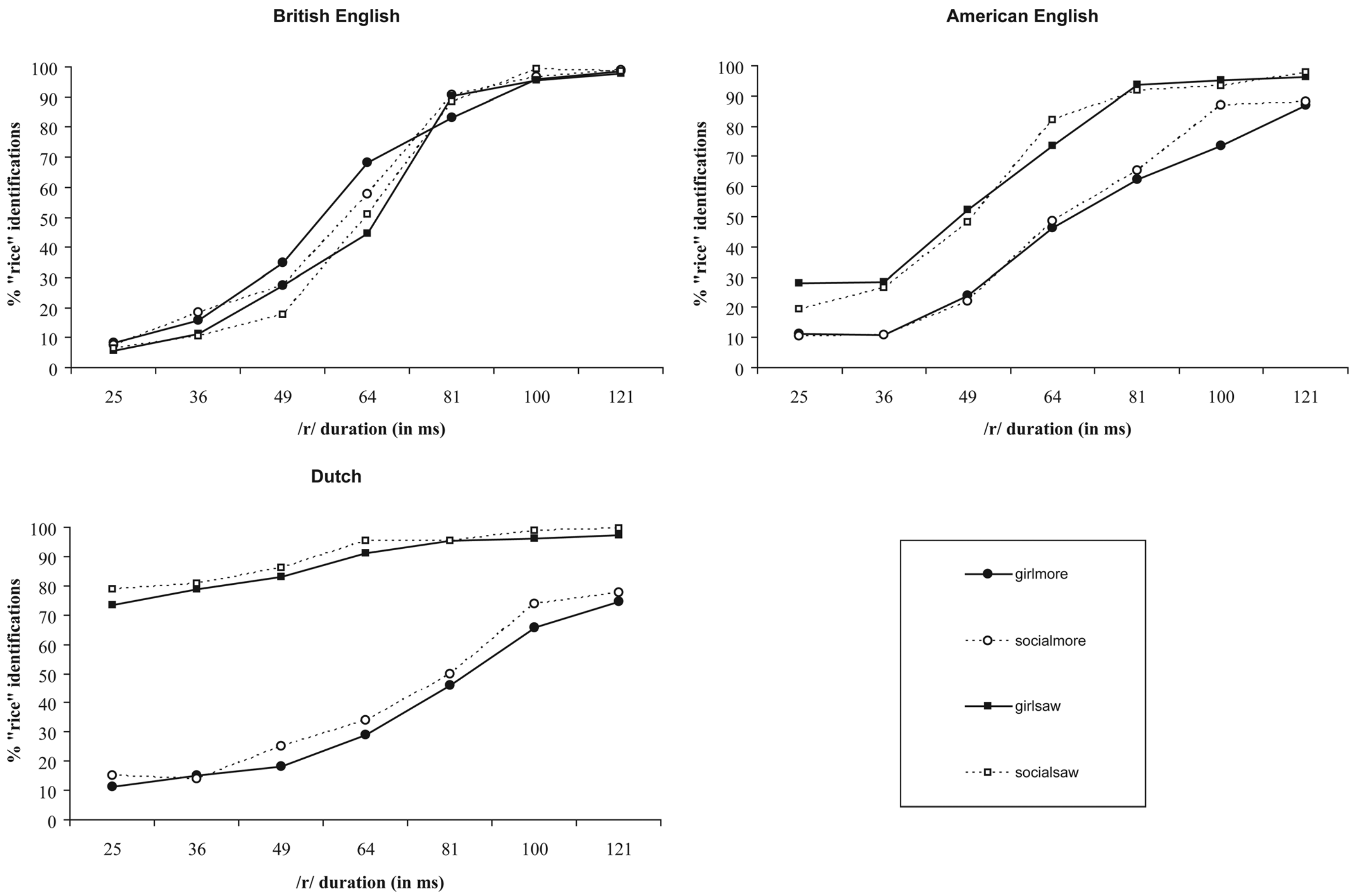

FIG. 1. Percentages of *rice* identifications in perceptual experiments with British English, American English and Dutch listeners, as a function of the duration of /r/ at the word boundary in the input, the orthography of the pre-boundary word (*more* vs *saw),* and the semantic context of the sentence (*the little girl … to her brother* vs *the social worker … to the poor).*

manipulation: sentences with *more (r)ice* versus with *saw (r)ice;* (3) the semantic manipulation: sentences with *the little girl* versus *the social worker* as subject. The dependent variable was the percentage of *rice* responses in each cell of the design.

In all three countries, the experiments were conducted on a standard PC running the NESU software. Participants were tested one at a time in a quiet room. They wore Sennheiser headphones, sat at a comfortable reading distance from the computer screen and had a two-button response box in front of them. Instructions were given in English (also to the non-native participants); these informed participants that on each trial they would hear an English sentence and see the words *ice* and *rice* on the computer screen. They were asked to press the right button labeled *rice* if the sentence they heard contained the word *rice,* and to press the left button labeled *ice* if the sentence contained the word *ice*. The 280 stimulus sentences (seven /r/-durations $\times$ four sentences $\times$ 10 repetitions) were presented in a random order which was different for every participant.

The experiments started with four practice trials. Each trial (practice or experimental) began with 150 ms of blank screen, after which the words *ice* and *rice* appeared in the upper left and upper right corners of the screen. After a further 450 ms the sentence was presented over the headphones. From the onset of /r/ in the sentence, participants had four seconds to press one of the buttons. Their response caused the other word to be removed from the screen, so that it could be seen that the answer had been registered by the computer. If participants did not respond within this time limit, a stopwatch appeared on the screen to remind them to react more rapidly. After each button press, the next trial started after a one-second interval. Participants could take a break after every 50 trials and continued when they were ready. Additionally, reaction times (RTs) were recorded from onset of the target word *(r)ice,* although the percentage of *rice* judgments was the principal dependent variable.

D. Results and discussion

Figure 1 shows the percentage of *rice* judgments as a function of the three independent variables, separately for each experimental population. It is clear at a glance that the three groups' response patterns all differed from each other. British English listeners produced four closely similar functions for the four sentences, with a smooth categorical curve along the durational continuum. For both the other groups there was, however, separation between the response functions for sentences with *saw* (squares) versus *more* (circles), suggesting that each of these groups was affected by the orthographic factor. However, the two latter groups' functions are far from identical; they differ in relative categorical shape, in degree of separation between the *more* and the *saw* sentences, and in whether or not there is a difference between the *little girl* and the *social worker* sentences.

1. British English listeners

The native listeners' responses were analyzed first. Individual means for each combination of /r/-duration, orthography, and semantic context were calculated. These means were then logistically transformed (see, e.g., Dixon, 2008, for the necessity of such transformation) and subjected to an ANOVA with Orthography, Context, and /r/-duration as independent variables. This analysis revealed no significant main effect of Orthography [$F(1, 17) = 4.4$, $p > 0.05$], no main effect of Context ($F < 1$), but a significant effect of /r/-duration [$F(6, 102) = 379.1$, $p < 0.001$], reflecting the increasing proportion of *rice* responses as duration of the /r/ increased.

In addition, there were two significant interactions: a two-way interaction between orthography and /r/-duration and a three-way interaction between orthography, context, and /r/-duration. To examine these interactions, ANOVAs with the independent variables orthography and context were conducted on all levels of the /r/-duration continuum. Effects of orthography and context were only observed at steps 2-4 of the /r/-duration continuum (as is visible in the British English panel in Fig. 1). At step 2, there was an effect of orthography [$F(1, 17) = 8.2$, $p < 0.05$], with more *rice* responses when there was an orthographic bias *against* an onset-/r/ (15% *rice* responses after *more* vs 9% *rice* responses after *saw;* note that these percentages are based on the estimated marginal means transformed back from log odds to a percentage value). At step 3, there were effects of both orthography [$F(1, 17) = 5.9$, $p < 0.05$]—again in the unexpected direction (29% *rice* responses after *more* vs 19% after *saw*)—and context [$F(1, 17) = 12.2$, $p < 0.01$], also in the unexpected direction (*girl:* 35% *rice* response, *social worker:* 27%). At step 4, there was a significant interaction between orthography and context [$F(1, 17) = 11.3$, $p < 0.01$]: the effect of context was in the expected direction for *saw* sentences, but in the opposite direction for *more* sentences (see Fig. 1).

The native listeners' response pattern is thus driven by the acoustic evidence for /r/, with no systematic effect of either of our other two manipulations. These listeners treat longer /r/ as a canonical /r/ in onset position, shorter /r/ as an intrusive /r/.

2. Non-native listeners

The performance of the Dutch listeners and of the American English listeners was then compared to that of the native listener group. Table II shows the interaction terms (with the language factor) arising from these two analyses. It can be seen that the two groups differ in similar direction (though often in differing degree) from the British English listeners.

In both cases the interaction of language and /r/-duration was highly significant, reflecting the lesser dependence of the two non-native groups' responses on the duration manipulation, compared with the native group. In the figure, this can be seen as flattening of the response curves across the duration continuum, in comparison with the categorical function evident in the native responses. In both cases, also, there was an interaction of language with orthography; although there was no main effect of the orthographic factor in the native responses, there was such an effect for each of the non-native groups. This can be seen in the figure in the separation of the response functions for *saw* sentences versus for *more* sentences for each of the non-native groups, but not for the native group. Further, in both cases there is a three-way interaction of language, orthography and /r/-duration; this reflects the fact that, as noted above, there was a reverse effect for a few points on the /r/ continuum for the native listeners, but always an effect in the predicted direction for the other two listener groups. This interaction for the Dutch listeners is stronger and also captures the flattening of the *saw* sentence curves visible in the figure.

Two interactions with language appeared only for the Dutch listeners: they showed a main effect of semantic context (visible in Fig. 1 as a consistent slightly higher proportion of *rice* responses to *social worker* than to *girl* sentences) shown by neither of the other groups, and they alone showed a three-way interaction of language, context and /r/-duration. The latter reflects the fact that the context effect for Dutch listeners was constant over the /r/-duration continuum, which was not the case for British English listeners (again, see above).

Thus neither non-native group produced a response pattern mimicking that of the native group. The Dutch group,

TABLE II. Interaction terms for ANOVAs comparing British English listeners with Dutch and with American English listeners. Asterisks denote significance level of interaction (* =< 0.05; ** =< 0.01; *** =< 0.001).

	Native Language			
	Dutch		American English	
Interaction terms for the comparison with BE listeners	df	F	df	F
Orthography	1,34	140.2***	1,30	8.1**
Semantic Context	1,34	4.7*	1,30	< 1
/r/-Duration	6,204	44***	6,180	5.7***
Orthography * Context	1,30	< 1	1,30	4.1
Orthography * /r/-Duration	6,204	19.7***	6,180	2.3*
Context * /r/-Duration	6,204	2.4*	6,180	1.6
Orthography * Context * /r/-Duration	6, 204	1.4	6, 180	2.1

with a different native language, differed from the native listeners more strongly than did the American group with only a varietal separation. This difference of relative similarity to the native group can be seen in Fig. 1, and is evident in the F-values which differ by more than an order of magnitude across the two analyses in Table II. Finally, note that the mean RTs of the British and American English listeners also did not differ (688 vs 690 ms), while the responses of the Dutch listeners were much slower (mean 952 ms). Subanalyses of the British English data revealed no difference between participants with slow versus fast mean response times; neither of these subsets showed either an effect of context or an effect of orthography in the expected direction.

IV. GENERAL DISCUSSION

The results of our perceptual experiments closely followed the predictions laid out in the Introduction. First, the durational differences between intrusive and canonical /r/, revealed by our acoustic analyses, formed as predicted the basis of the response pattern by native listeners. Second, listeners with another language failed as predicted to match the native sensitivity to the acoustic patterning of this connected speech process. Third, the responses of cross-dialect listeners resembled the native pattern, though again they did not match it.

In confirmation of the claims in phonetic descriptions of intrusive /r/ (Cruttenden and Gimson, 1994), the speaker whose utterances we analyzed made consistent and substantial acoustic distinctions between the tokens of canonical /r/, in word onset position, and the tokens of /r/ she intruded between vowels at a word boundary. We predicted on the basis of the perceptual findings from other connected speech processes that the difference between the two types of /r/ should then be easily accessible for native listeners. In this respect, intrusive /r/ thus patterns similarly to other linking phenomena that have been tested in both speech production and perception, such as liaison in French, whereby segments not pronounced in citation-form utterances of a word will surface in running speech when the word precedes a word-initial vowel. Such liaison segments are also significantly shorter than the same segments' canonical pronunciations, and listeners resolve ambiguity by exploiting the segment duration (Spinelli *et al.,* 2002, 2003). In that there is always an orthographic source for the surfacing segment, the phenomenon of liaison in French actually resembles British English linking /r/ more closely than intrusive /r/; we would predict that British listeners could similarly exploit segment duration to identify a linking /r/. The present results demonstrate that they can certainly distinguish intrusive from canonical /r/ on the basis of acoustic evidence alone.

The responses of the British English native listeners were highly categorical: short /r/s were reported as intrusive, long /r/s as canonical. These listeners made no use of the other factors that we had built into the materials; the semantic context had no impact on their responses at all, and the only statistically noticeable effect of the orthography comparison was in the direction opposite to what an orthographic sensitivity would induce. In fact, the latter result could also have arisen purely from durational processing. In natural speech, a sentence complement that does not begin with a conjunction is often signaled by a prosodic break, in the form of pre-boundary lengthening and initial strengthening of the post-boundary onset phoneme (Cho *et al.,* 2007). After *saw,* but not after *more,* the critical word *(r)ice* in our materials began the complement. If British English listeners interpreted a long /r/ after *saw* as a preboundary-lengthened intrusive /r/ rather than as an (insufficiently lengthened) onset /r/, their tendency to fewer *rice* responses after *saw* just in the ambiguous portion of the continuum would, like every other aspect of their response pattern, attest to their fine-grained sensitivity to the durational structure of their native speech.

Neither the sensitive appreciation of the acoustic realization of /r/ shown by the native British listeners, nor their concentration on the acoustic information to the exclusion of other variation, were replicated in the two non-native listener groups. Here our predictions were again borne out in that while both groups showed similar differences from the British listeners, the cross-language difference was much larger than the cross-dialect difference. First, both non-native groups produced shallower identification functions than the British group over the /r/ duration continuum; that is, although they too attended to the acoustic manipulation, their use of it was less sensitive than that of the native listeners. Second, both American and Dutch listeners were affected by the written form of the words they heard, and gave fewer onset /r/ responses when the /r/ sound could have been attributed to the letter *r* in *more,* which, as we saw, the native listeners did not do at all. In each case the discrepancy between the native and the non-native results was greater for the cross-language than for the cross-dialect comparison. Finally, further differences appeared in the cross-language but not in the cross-dialect comparison: Dutch listeners were the only ones who showed sensitivity to the semantic bias built into our materials, and were the only ones to allow orthography to modulate their use of durational information.

Our prediction that listeners with another language would fail to show nativelike performance with intrusive /r/, despite high proficiency in the target language, was motivated by the evidence from non-native phoneme and word perception. In the particular case of Dutch listeners to English, high listening proficiency and wide exposure to native input does not stop these listeners experiencing substantial confusion between minimal word pairs differing only in a contrast that appears in English but not in Dutch (e.g., the vowels in *cattle* versus *kettle,* or the word-final voicing distinction in *robe* vs *rope;* Broersma and Cutler, 2008, 2011). Our present results suggest that unfamiliar phonological processes can cause similar confusion. Although, as already noted, British English is widely available in the Dutch media and is the target pronunciation taught in Dutch schools and universities, the process of /r/-insertion, like all other casual speech phenomena, is not explicitly taught. Moreover, the process is completely absent from Dutch.

Previous findings also underlay our decision to incorporate in our materials both an orthographic and a semantic manipulation, and as in the prior cases (Escudero *et al.*, 2008; Bradlow and Alexander, 2007), each manipulation had a significant effect on the response patterns of our cross-language listeners. Though offline judgements had suggested that English-speakers were also sensitive to our semantic manipulation (albeit to a lesser extent than Dutch-speakers), in practice neither English-native group based their responses on this factor.

Contrasting with the performance of our Dutch listeners, advanced L2 learners tested by Darcy *et al.* (2007) could to some degree compensate for a type of assimilation unfamiliar from their L1, as noted in the Introduction. However, although both assimilation and /r/-insertion fall into the general class of connected speech phenomena, they are in many ways different. Assimilation alters the nature of the evidence for a segment (e.g., its place of articulation or voicing), but insertion adds evidence that is in principle compatible with an additional segment. Assimilation occurs widely across languages and listeners can compensate for unfamiliar assimilation processes in languages they do not know (e.g., general auditory processes allow Dutch listeners with no knowledge of Hungarian to compensate for the Hungarian assimilation process by which /lr/ sequences become /rr/; Mitterer *et al.,* 2006a,b). The /r/-insertion process, in contrast, is far less common. It is certainly not an automatic process in speech production; the cross-boundary vocalic sequences that trigger it in British English also feature in varieties of English which do not show this insertion, as well as in Dutch and other languages which likewise do not show it. Further, though the process is widespread in British English (Foulkes and Docherty, 1999), some speakers variably suppress it for socially motivated reasons (Broadbent, 1991; Gussenhoven and Broeders, 1997). (Note that the resulting variable occurrence may make it yet more difficult for non-native listeners to interpret an intrusive /r/ correctly when they hear one.) Our results show that the perceptual processing of /r/-insertion patterns quite differently from that of assimilation; non-native listeners do not respond to it as native listeners do.

In the cross-dialect case, however, our study suggests that the listening difficulty resulting from /r/-insertion is not of the same order as the difficulty experienced by the cross-language listeners. In this our results are in line with those of Cutler *et al.* (2005), who found that a dialect mismatch caused less difficulty in phoneme perception than a language mismatch. In our experiment, American listeners did not attain native listener levels in their categorization of intrusive versus canonical /r/, and they were sensitive to the orthographic factor ignored by the native listeners, but the strongest influence on their responses came, as with the native listeners, from the durational evidence for /r/.

From the small literature on cross-dialect speech perception, it appears that different varieties of the same language can encourage attention to different cues to the same contrast (e.g., Miller and Grosjean, 1997; Kirby, 2010); but the flexibility that speakers of the same language show in speech perception can allow rapid adjustment, especially if contextual cues are available. Thus within regional dialects of American English, Clopper and Bradlow (2006) found that perception of predictable sentences in noise was not adversely affected by a speaker/listener mismatch in regional dialect, and Sumner and Samuel (2009) found that General American speakers' difficulties in recognizing words produced with a New York accent could easily be removed by semantic priming. Otake and Cutler (1999) found that suprasegmental structure differences across accent did not adversely affect accuracy of interpretation in minimal-pair decisions, although, as already noted, signal-detection measures showed that listeners with a mismatching dialect were both less sensitive to acoustic cues to the distinction in question, and more susceptible to bias. Floccia *et al.* (2006) observed that adjustment to regional accent occurred during their experiments. Within regional dialects of British English, Evans and Iverson (2004) showed that listeners could alter their goodness rating for vowels according to the dialect (northern versus southern) of the carrier phrase in which the tokens occurred. Interestingly, this pattern of alteration was shown by northern speakers who had been living for years in a southern dialect area, but not by northern speakers who had remained living in the north (despite the fact that the southern variety is dominant in British media so that the northern listeners would have been regularly exposed to it).

This latter result raises the possibility that the value of exposure may principally be felt at the higher exposure ranges. Immersion, in other words, can produce learning though more or less regular encounters with a particular variety cannot. Further evidence for this may be found in the findings of Scott and Cutler (1984), who studied British listeners' perception of connected speech processes in American English. British listeners are (and were at the time of Scott and Cutler's testing) regularly exposed to American speech via television and cinema. Nonetheless, the British-resident listeners did not succeed in matching American listeners' performance at deriving syntactic information from the patterns of cross-word-boundary flapping (e.g., to distinguish *If you want to eat, early lunch will be served* from *If you want to eat early, lunch will be served).* A group of British residents of Chicago, however, did show nativelike ability to do this, and interestingly, their ability to do it was not correlated with length of residence, which varied from one to 30 yr. The perceptual learning which, within the native language, efficiently supports adjustment to new talkers (Norris *et al.*, 2003), can lead to adjustment to a new variety if the majority of talkers one encounters every day use that variety.

The perceptual adjustment to dialectal varieties that these findings indicate does not appear to hold for different late-learned languages, however. Immersion does not remove the perceptual difficulty of phoneme repertoire mismatch for second-language listeners (e.g., Ingram and Park, 1998). The advantage of a largely identical phoneme repertoire, vocabulary and higher level linguistic processes across varieties of one and the same language supports perceptual learning for speech in a way that cannot be equaled across languages. The connected speech processes that affect phoneme realization, we conclude, pattern similarly in this respect to the more well-studied processes of phoneme identification.

In conclusion, our results show that native listeners efficiently exploit the durational cues which are reliable correlates of /r/-insertion. Listeners of another language exploit these cues less effectively, and fall back on semantic context and orthographic information. Listeners with a different dialect are more native-like than listeners with a different language. That is, native listeners keep *Laura Norder* in her place, listeners with another dialect also know where *Laura* should go, and only non-native listeners are not sure what to do with her.

ACKNOWLEDGMENTS

This research was supported by a research stipend from the Max Planck Society and a Fulbright grant, both awarded to the first author. We thank Alan Garnham for making the British English listener population available, Delphine Dahan for making the American English listener population available, Leah Roberts for recording the speech materials, and Laurence Bruggeman for technical assistance. The acoustic measurements of /r/ were reported to the 16th International Conference of Phonetic Sciences, Saarbrucken (Tuinman *et al.*, 2007).

APPENDIX

Sentence	Mean /r/ duration (ms)	Mean intensity decrement (dB)
A few days ago, I saw aces when I looked at my cards.	56.5	1.3
A few days ago, I saw races from the British Superbike Championship.	69.8	11.1
I heard that Canada aided the area of Lesotho in Africa.	63.3	2.0
I heard that Canada raided the area of Lesotho in Africa.	84.9	7.1
I really thought that I saw air burning.	71.8	2.8
I really thought that I saw rare animals.	93.6	6.2
Did you know that the terracotta ear of the statue was broken?	64.7	1.1
Did you know that the terracotta rear of the statue was broken?	80.8	8.7
My brother likes extra ice when he has dinner.	64.7	2.8
My brother likes extra rice when he has dinner.	83.3	8.0
I think that Rebecca owes a lot to London.	64.8	2.8
I think that Rebecca rows a lot to London.	86.8	8.9
In north Malaysia itches are number one on the list of annoyances.	65.8	4.4
In north Malaysia riches are frowned upon.	100.5	7.5
And then Emma ejected the cassette.	60.6	1.3
And then Emma rejected the cassette.	69.3	5.2
I read that people from China etch whenever they feel like doing so.	64.6	1.3
I read that people from China retch whenever they feel like doing so.	88.1	7.0
My youngest sister saw odes to Rome made by many different people.	68.6	0.2
My youngest sister saw roads to Rome made by many different people.	96.5	9.1
I saw on Discovery Channel that people in Panama ate cows' eyes in former times.	91.3	−2.2
I saw on Discovery Channel that people in Panama rate David as the nicest city.	107.7	1.3
My favorite grandma aged because her dog suddenly died.	87.4	2.4
My favorite grandma raged because her dog suddenly died.	112.8	4.2
Do you think it was sepia ash on her body?	72.6	0.3
Do you think it was sepia rash on her body?	88.1	6.1
We had not expected it, but suddenly Natasha emitted a cry of pain.	58.8	2.9
We had not expected it, but suddenly Natasha remitted the money to us.	71.4	3.6
Usually, Sheila elates everybody with her presence.	69.4	7.8
Usually, Sheila relates well to other people.	79.6	12.4
The Broadway musical "Clarissa" evolves into a deeply moving metaphor for the struggles of an entire generation.	66.8	1.1
The Broadway musical "Clarissa" revolves around a young American girl in London.	100.4	8.1
I read that the president of Algeria evokes the idea of using nuclear weapons.	66.2	3.2
I read that the president of Algeria revokes his decision to cut taxes.	87.3	7.8
I heard that the Australia alley is a nice street to live in.	74.5	2.0
I heard that the Australia rally lasts two weeks.	94.1	10.8
For Buddhists in India enunciation is of great importance.	73.5	1.4
For Buddhists in India renunciation is part of their daily lives.	96.7	7.8
We asked the children to draw apt presents for Mother's Day.	66.2	5.2
We asked the children to draw wrapped presents for Mother's Day.	80.4	11.9
Obviously, the bourgeois eye expert was bragging about his salary to his Harvard friends.	88.0	0.0
Obviously, the bourgeois rye expert was bragging about his salary to his Harvard friends.	96.9	8.5
Davidson and Brooks claim to be thorough ale-manufacturers.	77.7	4.9
Davidson and Brooks claim to be thorough rail-manufacturers.	98.4	12.9
The president of Russia eagerly awaits his caviar.	62.4	4.1
The president of Russia regally decorated his new office.	100.7	10.6
In this area anglers are still ice fishing.	78.6	1.5
In this area wranglers are experienced horseback tour guides.	86.2	5.4
In postwar America aches and pains are usually ignored by people without health insurance.	68.2	0.5
In postwar America rakes and planes are sold in hardware shops.	91.3	6.3
You should use extra old cheese in this recipe.	62.3	3.2
You should use extra rolled oats in this recipe.	83.9	8.2
I think that the extra ink cartridges were too expensive.	53.2	1.5
I think that the extra rink did not result in more ice skaters visiting it.	80.3	4.8

Best, C. T., McRoberts, G. W., and Goodell, E. (**2001**). "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," J. Acoust. Soc. Am. **109**, 775–794.
Bradlow, A. R., and Alexander, J. A. (**2007**). "Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners," J. Acoust. Soc. Am. **121**, 2339–2349.
Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. I. (**1997**). "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production," J. Acoust. Soc. Am. **101**, 2299–2310.
Broadbent, J. (**1991**). "Linking and intrusive r in English," UCL Working Papers Linguist. **3**, 281–302.
Broersma, M., and Cutler, A. (**2008**). "Phantom word recognition in L2," Syst.: Int. J. Educ. Technol. Appl. Linguist. **36**, 22–34.
Broersma, M., and Cutler, A. (**2011**). "Competition dynamics of second-language listening," Q. J. Exp. Psychol. **64**, 74–95.
Cho, T., McQueen, J. M., and Cox, E. A. (**2007**). "Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English," J. Phonetics **35**, 210–243.
Clopper, C. G., and Bradlow, A. R. (**2006**). "Effects of dialect variation on speech intelligibility in noise," J. Acoust. Soc. Am. **119**, 3424.
Cohen, J. (**1992**). "A power primer," Psychol. Bull. **112**, 155–159.
Collins, B., and Mees, I. M. (**1999**). *The Phonetics of English and Dutch* (Brill, Leiden), pp. 178–181.
Connine, C. M., Ranbom, L. J., and Patterson, D. J. (**2008**). "Processing variant forms in spoken word recognition: The role of variant frequency," Percept. Psychophys. **70**, 403–411.
Coren, S., Ward, L. M., and Enns, J. T. (**1994**). *Sensation and Perception*, 4th ed. (Harcourt Brace College Publishers, Fort Worth), pp. 234–235.
Cruttenden, A., and Gimson, A. C. (**1994**). *Gimson's Pronunciation of English* (Arnold, London), pp. 187, 262–264.
Cutler, A., Smits, R., and Cooper, N. (**2005**). "Vowel perception: Effects of non-native language vs. non-native dialect," Speech Commun. **47**, 32–42.
Cutler, A., Weber, A., and Otake, T. (**2006**). "Asymmetric mapping from phonetic to lexical representations in second-language listening," J. Phonetics **34**, 269–284.
Darcy, I., Peperkamp, S., and Dupoux, E. (**2007**). "Bilinguals play by the rules: Perceptual compensation for assimilation in late L2-learners," in *Papers in Laboratory Phonology 9*, edited by J. Cole and J. Hualde (Mouton de Gruyter, Berlin), pp. 411–442.
Dutoit, T., Pagel, V., Pierret, N., Bataille, F., and van der Vreken, O. (**1996**). "The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non–Commercial Purposes," in *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 1996)*, Philadelphia, PA, pp. 1393–1396.
Escudero, P., Hayes-Harb, R., and Mitterer, H. (**2008**). "Novel second-language words and asymmetric lexical access," J. Phonetics **36**, 345–360.
Evans, B. G., and Iverson, P. (**2004**). "Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences," J. Acoust. Soc. Am. **115**, 352–361.
Floccia, C., Goslin, J., Girard, F., and Konopczynski, G. (**2006**). "Does a regional accent perturb speech processing?," J. Exp. Psychol. Hum. **32**, 1276–1293.
Foulkes, P., and Docherty, G., eds. (**1999**). *Urban Voices: Accent Studies in the British Isles* (Arnold, London), pp. 51, 76, 111, 133, 147, 174.
Giegerich, H. J. (**1992**). *English Phonology: An Introduction* (Cambridge University Press, Cambridge, UK), pp. 66, 28–283.
Guion, S. G., Flege, J. E., Akahane-Yamada, R., and Pruitt, J. C. (**2000**). "An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants," J. Acoust. Soc. Am. **107**, 2711–2724.
Gussenhoven, C., and Broeders, A. (**1997**). *English Pronunciation for Student Teachers* (Wolters-Noordhoff, Groningen), pp. 16–17, 154–155.
Ingram, J. C. L., and Park, S.-G. (**1998**). "Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners," J. Acoust. Soc. Am. **103**, 1161–1174.
Irwin, P., and Nagy, N. (**2007**). "Bostonians /r/ speaking: A quantitative look at (R) in Boston," Penn Working Papers Linguist. **13**, 135–148.
Kirby, J. (**2010**) "Dialect experience in Vietnamese tone perception," J. Acoust. Soc. Am. **127**, 3749–3757.
Miller, J. L., and Grosjean, F. (**1997**). "Dialect effects in vowel perception: The role of temporal information in French," Lang. Speech **40**, 277–288.
Mitterer, H., Csépe, V., and Blomert, L. (**2006a**). "The role of perceptual integration in the perception of assimilation word forms," Q. J. Exp. Psychol. **59**, 1395–1424.
Mitterer, H., Csépe, V., Honbolygo, F., and Blomert, L. (**2006b**). "The recognition of phonologically assimilated words does not depend on specific language experience," Cogn. Sci. **30**, 451–479.
Mitterer, H., and Ernestus, M. (**2006**). "Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch," J. Phonetics **34**, 73–103.
Otake, T., and Cutler, A. (**1999**). "Perception of suprasegmental structure in a non-native dialect," J. Phonetics **27**, 229–253.
Norris, D., McQueen, J. M., and Cutler, A. (**2003**). "Perceptual learning in speech," Cogn. Psychol. **47**, 204–238.
Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler-Lussier, E. (**2007**). "Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu]" [Department of Psychology, OH State University (Distributor). Columbus, OH] (Last viewed 9 June 2010).
Polka, L. (**1995**). "Linguistic influences in adult perception of non-native vowel contrasts," J. Acoust. Soc. Am. **97**, 1286–1296.
Scott, D. R., and Cutler, A. (**1984**). "Segmental phonology and the perception of syntactic structure," J. Verbal Learn. Verbal Behav. **23**, 450–466.
Spinelli, E., Cutler, A., and McQueen, J. M. (**2002**). "Resolution of liaison for lexical access in French," Rev. Fr. Linguist. Appl. **7**, 83–96.
Spinelli, E., McQueen, J. M., and Cutler, A. (**2003**). "Processing resyllabified words in French," J. Mem. Lang. **48**, 233–254.
Sumner, M., and Samuel, A. G. (**2009**). "The effect of experience on the perception and representation of dialect variants," J. Mem. Lang. **60**, 487–501.
Trudgill, P. (**1986**). *Dialects in Contact* (Basil Blackwell, Oxford), pp. 76–77.
Tuinman, A., Mitterer, H., and Cutler, A. (**2007**). "Speakers differentiate English intrusive and onset /r/, but L2 listeners do not," in *Proceedings of the 16th International Congress of Phonetic Sciences*, edited by J. Trouvain and W. J. Barry (Pirrot, Dudweiler, Germany), pp. 1905–1908.
Underbakke, M., Polka, L., Gottfried, T. L., and Strange, W. (**1988**). "Trading relations in the perception of /r/-/l/ by Japanese learners of English," J. Acoust. Soc. Am. **84**, 90–100.
van Bezooijen, R. "Approximant /r/ in Dutch: Routes and feelings," (**2005**). Speech Commun. **47**, 15–31.
van de Velde, H. (**1996**). "Variatie en verandering in het gesproken Standaard-Nederlands (1935–1993) [Variation and change in spoken standard Dutch (1935–1993)]," Dissertation, Nijmegen University.
Weber, A., and Cutler, A. (**2004**). "Lexical competition in non-native spoken-word recognition," J. Mem. Lang. **50**, 1–25.
Wells, J. C. (**1982**). *Accents of English* (Cambridge University Press, Cambridge), p. 520.
Wolfram, W., and Schilling-Estes, N. (**2006**). *American English: Dialects and Variation* (Blackwell, Malden, MA), pp. 105–106.