Chapter 4

Phonetic Precision in Listening

Anne Cutler and Mirjam Broersma
Max Planck Institute for Psycholinguistics
Nijmegen, The Netherlands

> Not many people know very much about ... how speech actually works .... It is the privilege and the pleasure of phoneticians, with their colleagues from related disciplines professionally concerned with speech, to engage themselves in the analysis of this most intricate of our communicative skills. (Laver, 1994, p.592)

Among the skills which great phoneticians possess is the ability to perceive many subtle articulatory distinctions. However, it is indeed true that not many people share this ability. How precise is the phonetic perception of the ordinary listener?

Paradoxically, listeners' phonetic processing is both precise and imprecise. It is imprecise in that what matters ultimately for the listener is not phonetic detail but categorisation. Discrimination of speech sounds is hugely better across phoneme category boundaries than within categories. Imprecision of listening effectively does not matter within categories; the ordinary listener would just be distracted by analysing precise details of each speaker's pronunciation of each speech sound. This is because, after all, the point of phonetic perception is to understand spoken messages. What matters is to make distinctions between words, and phonemic categories make (by definition) distinctions between words.

However, listeners are also capable of extremely precise distinctions in the service of interword decisions, in cases where subphonemic information is available and can be of assistance. In this sense, phonetic perception can be highly precise. The next section discusses some relevant evidence. Furthermore, listeners' phonemic category boundaries are highly sensitive and can be adjusted to deal with distributional variation in the input; section 3 discusses some relevant recent evidence on this issue. Given these listening abilities, it is all the more striking how difficult it is for most people to acquire fully the perceptual phonemic categories of a second language in adulthood; the remaining sections discuss the implications of this for the ordinary listener.

HOW WORD RECOGNITION WORKS

> Phonological units (including, but not only, consonant and vowel phonemes) have the sole linguistic function of being combinable and permutable, within narrowly defined structural constraints of sequence, to give distinctive shape to the very large number of grammatical units (words) whose identity and sequence in their turn make up the lexical and syntactic patterns of the language. (Laver, 1989, p. 40)

Precise phonetic category decisions are not a necessary component of word recognition—in fact, word recognition does not necessarily involve phonetic category decisions at all. The point of word recognition is to convert the incoming speech signal into known meaningful units, and phonetic units do not in themselves carry meaning; morphemes, whether stand-alone or combinable, are the smallest units that can do that. Of course, phonetic decisions are indirectly involved in word recognition in that any acceptance of a word (*phone*) involves rejection of other words differing from it by minimally one phoneme (*fine, shown, foal*). But whether spoken-word recognition involves a stage of explicit conversion of the input into a representation in terms of phonemes is one of the most disputed questions in psycholinguistics, which has received more than three decades of attention without being resolved. The discussion continues unabated (see, e.g., Pallier, Colomé, & Sebastián–Gallés, 2001; Pierrehumbert, 2002), but the present authors do not intend to contribute further to it here.

Instead, we highlight some characteristics of the word recognition process which play a significant role in the research issues involved which underlie the new findings we discuss in later sections. The most important factors are multiple activation and competition. As the first of these terms suggests, the word to be recognised may not be the only lexical form actively participating in the recognition process. The arrival of incoming speech information calls up an array of potential word candidates which form at least temporarily a partial match to the speech input. Although the temporal nature of speech processing prompted early proposals that lexical forms could be recognised in sequential order of arrival, this is in practice hardly possible for the listener given the extent to which words in any vocabulary are similar and overlapping. Vocabularies contain tens or hundreds of thousands of words, but these are constructed using only a handful of phonetic categories (on average around 30; Maddieson, 1984). Moreover, languages (and presumably language users) prefer short words to long ones. The inevitable result is large numbers of minimal pairs of shorter words, and longer words with shorter words embedded within them. In fact, only about 2% of English words do not contain some other word form (Cutler, McQueen, Jansonius, & Bayerl, 2002). And, because more of these embeddings occur at the beginning than in the middle or at the end of the matrix word, the first full word a listener hears may not be the intended word, but only a spuriously embedded form occurring within it. Thus *star* may not be *star* but may turn into *start* or *stark* or *starve* or *starling* as more speech input arrives; *start* may become *starch* or *startle; starch* may turn out to have been *star chart* after all.

Under these circumstances, efficiency appears to be served by making available all the potential options in parallel: *star* and *start* and *stark* and *starve* and *starling* and *starch* and *startle* may all be simultaneously activated given the input *star-*. There is abundant experimental evidence for this phenomenon by now (see McQueen, 2004, for a review). The superfluity of simultaneously available candidates is resolved by a process of competition between them. That is, the more one candidate is favoured, the more it can disadvantage its rivals. Thus, incoming information which matches one candidate but not others (a /k/, for instance, after *star-*) does not simply result in the preferred candidate (*stark*) accruing more points in its favour and increasing in activation, but it also leads to a decrease in activation of the alternative candidates (*starling, starve,* etc.). Again, evidence from laboratory experiments attests to this inhibitory effect. Thus, lexical decision responses to a visual presentation of TRAFICO are significantly faster after an immediately preceding spoken fragment *trafi-* than after a control prime, but after the fragment *tragi-*, (matching *tragico, tragic,* rather than *trafico, traffic*) responses are significantly slower than the responses in the control condition (Soto–Faraco, Sebastián–Gallés, & Cutler, 2001). Words embedded within other words may be similarly inhibited (such as *mess* in *domestic;* McQueen, Norris, & Cutler, 1994). Obviously, the competition process is primarily constrained by incoming information from the speech signal, but because speech is continuous, word boundaries may not be apparent to the listener, so that multiple candidates may temporarily enjoy full support. Direct competition between all words supported (fully or partially) by the input offers an efficient means of evaluating such candidates without processing delays (see McQueen, Cutler, Briscoe, & Norris, 1995, for discussion of this issue). However, the more competition arises, the slower words may be recognised (Norris, McQueen, & Cutler, 1995).

The modulation of competition by incoming speech information is rapid; here the precision of the listener's processing is remarkable. Matching or mismatching, such information is put to use immediately, and this includes early arriving coarticulatory cues which can resolve a competition process. For instance, suppose a listener hears the fragment *jo-* from the word *jog*. Velar cues in the vowel are already sufficient to boost the activation of *jog* and inhibit competitor words such as *job*. This can be seen from the fact that if by cross-splicing a version of *job* is created in which the *jo-* actually comes from *jog*, positive lexical decision responses to that form are significantly slower than to a cross-spliced *job* in which the *jo-* comes from *jod*, a nonword. Further, if the nonword *smob* is presented, and its *smo-* comes from the real word *smog*, negative lexical decision responses are slower than if the *smo-* were taken from *smod*, another nonword (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Marslen–Wilson & Warren, 1994; McQueen, Norris, & Cutler, 1999). In both cases, the coarticulatory information in the vowel about the place of articulation of the upcoming consonant was passed on to influence lexical activation, favouring *jog* and *smog*, respectively.

These data, and many further demonstrations of effects of subphonemic mismatch on lexical activation (e.g., Streeter & Nigro, 1979; Utman, Blumstein, & Burton, 2000; Van Alphen & McQueen, 2003; Whalen, 1984) and of listeners' efficient use of coarticulatory information in distinguishing both vowels and consonants (e.g., Martin & Bunnell, 1981, 1982; Strange, 1989), clearly show that spoken-word recognition involves precise processing of contextually induced phonetic variability. Note that although attempts have been made to argue, on the basis of such findings, that a phonemic level of representation has no place in models of speech understanding, the data are actually neutral with respect to this issue. In activation/competition models, information may be probabilistically weighted to reflect degree of match or mismatch to more than one alternative during prelexical processing, just as more than one candidate may be simultaneously active at the lexical level. The data could certainly be held to rule out discontinuous models in which correct phonemic category decisions would form a prerequisite for lexical access, but in fact no such models enjoy current support in the psycholinguistic literature.

WORD RECOGNITION AND ADJUSTMENT OF PHONETIC PRECISION

> Two speakers may share the same vowel and consonant systems, have the same structural possibilities, choose the same lexical selection of phonemes, and yet have slightly different accents. The fine detail of how a given speaker pronounces his sounds can act as a marker of group membership, but it can also function as an individuating marker. (Laver & Trudgill, 1979, p. 19).

Precision would be served by constancy in the definition of phoneme categories; the more exact the better, if the decision about a particular category of the native language were subject to the same constraints in each and every instance. But this is not the case. There is considerable variation in the realisation of tokens of any other phonemic category, and the amount of variation changes as a function of context, of rate of speech, and across speakers and dialects. That in itself is not problematic; the category definitions can be loosely set. But categories are also not constant across speakers of a given language or dialect, and they are not immutable across time. The realisation of categories, and by extension the range of the associated variation, changes over time for the language community as a whole (Labov, 1994) and for individual speakers (Bauer, 1985; Harrington, Palethorpe, & Watson, 2000; Yaeger–Dror, 1994).

How can listeners adapt the definitions of phoneme categories? The only reasonable assumption is that such adaptation occurs as a by-product of word recognition. Even if an individual token is somewhat different from other tokens of the same phoneme that we have previously heard, the word best matched by the input as a whole will still win the competition process, and the ensuing lexical recognition will motivate a record of the altered input for future use. Of course, in practice, a shift within a language community will occur gradually, as a shift of the clustering within an existing range of variation. But the adaptation needed to shift is arguably the same adjustment that occurs when we accommodate to a new, unfamiliar speaker with a variant pronunciation (foreign accent, speech impediment, etc.). It is a familiar experience that such adjustment can occur quite rapidly; a new speaker can sound quite strange and hard to understand at the beginning of a lecture, but within quite a short space of time the strangeness has disappeared and the accent is no longer hard to process.

If these processes are the same, then it should not be necessary for native categories to undergo a gradual process of shift of centre of gravity in a variation space; it should be possible to induce adjustment of phonemic categories in the laboratory very quickly by supplying lexical support for the adjustment. We tested this hypothesis in our laboratory in a series of experiments involving the fricative sounds [s] and [f] in Dutch (Norris, McQueen, & Cutler, 2003). The lexical support was provided by placing the sounds in question at the end of real words which would normally end in [s] or [f]. Some examples are the Dutch words for *carcase* (*karkas*, with final stress) and *carafe* (*karaf*). These words have no minimally paired companions ending with the other fricative, just as English *carcafe* and *carase* are not words.

The series of experiments began with a standard phonetic categorisation study. Dutch listeners made forced-choice [s]-[f] decisions about stimuli along a continuum from a good [s] to a good [f] through 12 intermediate tokens created by sampling the two natural endpoints in different proportions. The results of this study produced a maximally ambiguous token between [s] and [f]. This token was then grafted onto the end of spoken tokens of Dutch words of the *carcase-carafe* type, in place of the originally produced fricative.

New groups of Dutch listeners then took part in the crucial experiment. They carried out a lexical decision task: they heard spoken items and had to decide for each one whether it was a real Dutch word. There were 200 items in the experiment, and nothing in the instructions called attention to phonetic processing; only the lexical decision was emphasised. Nearly all the items contained no occurrences of [s] or [f] at all; the exceptions were 20 real words of the *carcase* type and 20 of the *carafe* type.

There were two experimental groups of listeners, and they differed in how they heard these words. One group heard the 20 [s] words in natural form, and the 20 [f] words with the final [f] replaced by the selected ambiguous fricative. The other group heard the 20 [s] words with the final [s] replaced by the ambiguous sound, and the 20 [f] words in natural form. The objective was to shift the boundary of the [s]-[f] distinction for these two groups, but to shift it in opposite directions. Would exposure to only 20 words be enough to bring this about?

In the lexical decision task, the listeners certainly treated the items ending with the ambiguous sounds as words, that is, they signalled a YES decision (although their response times to do so were slower than responses to unmanipulated word tokens). But the real test came afterward. Once the lexical decision task had been completed, the listeners were given a phonetic categorisation task, using a subset of the same continuum used in the pretest study. And, indeed, their decisions had been affected by the experience in the lexical decision experiment, as Fig. 4.1 shows. The listeners who had heard the *carcase*-type words with the ambiguous sound made more [s] decisions in the categorisation task, and the listeners who had heard the *carafe*-type words with the ambiguous sound made more [f] decisions. The responses of control groups of listeners fell in between.

The control groups were crucial to the demonstration that the adaptation depended on lexical information. There were control lis-

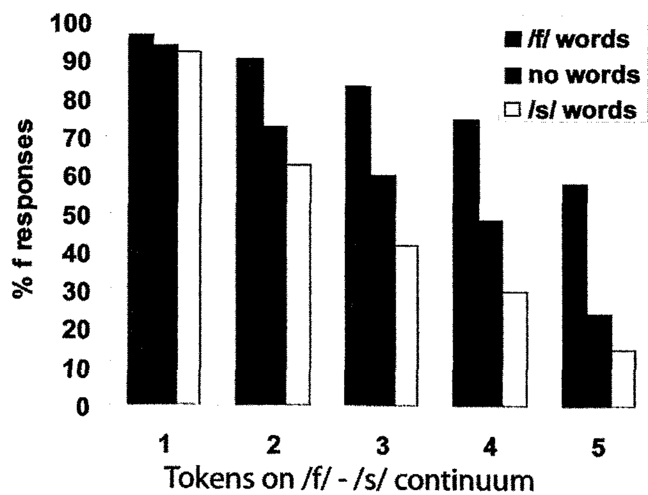

FIG. 4.1. Mean percentage of /f/ responses to ambiguous tokens lying on a continuum from Dutch /f/ to /s/, by listeners who had heard words containing token 3 in place of /f/ (mean of two groups), listeners who had heard no words containing the ambiguous tokens (mean of five different control groups), and listeners who had heard words containing token 3 in place of /s/ (mean of two groups; averaged data from Norris, McQueen, & Cutler, 2003).

teners who heard the ambiguous sound at the end of nonwords, but heard no words ending in [s] or [f]; or who heard the ambiguous sound at the end of nonwords and heard words ending with one or the other of the natural fricatives; or who heard words ending with one or other of the natural fricatives but no ambiguous sound at all. None of these groups showed the significant response shift that was observed with the experimental groups. Only the experimental groups had heard the ambiguous sound with lexical support signalling its interpretation, and only those groups showed an effect of this experience in their subsequent categorisation responses. The listeners who had received lexical support for interpreting the ambiguous sound as [s] shifted their category decisions toward [s], those who had received support for interpreting it as [f] shifted toward [f]. Moreover, this shift was not specific to the actual ambiguous sound that they had been exposed to, but embraced the other intermediate sounds which they received in the categorisation test. The whole continuum shifted for these listeners; just a short exposure, involving only 20 words, had been sufficient to adapt their phonemic category decisions. Thus, the precision of ordinary listeners' phonetic processing is sufficient to support considerable flexibility in phonetic category assignment.

WHEN PRECISION IS UNATTAINABLE

> Learning a language other than one's own native language is always a process in which the patterns of the first language interfere with the learning of the foreign language (Laver, 1994, pp. 78–79).

Given that the native phonetic processing is so precise when needed, imprecise when sufficient, and even flexible when appropriate, it is disturbing that mastering the categories of a second language acquired after childhood can be so extraordinarily difficult. Foreign accent is the most conspicuous symptom, and far more research attention has been given to phonetic imprecision in second-language production than perception. But of perception it is nevertheless well attested that responding to non-native category distinctions is already hard by the end of the first year of life (Werker & Lalonde, 1988; Werker & Tees, 1983), and adult listeners find it very difficult indeed to improve their performance (see Strange, 1995, for a review), although huge amounts of training can bring about some improvement (Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively, & Pisoni, 1991). Perception of native contrasts remains more categorical than perception of non-native contrasts (Burnham, Earnshaw, & Clark, 1991).

Adult listeners appear to be imprisoned within the categories of the native language. It is not the case that they cannot perceive non-native contrasts for some auditory reason (e.g., a loss of sensitivity due to lack of exposure). Rather, native phonology captures speech input and categorises it although this may mean forgoing much of the precision which the listener's auditory system could have delivered. Only when experimenters can contrive somehow to switch off phonological processing does the sensitivity of auditory processing have a chance of being revealed. Werker and Tees (1984) found that adult English speakers were able to perceive non-English phonemic distinctions (between Hindi retroflex versus dental alveolar voiceless stops, and Thompson—an Amerindian language—ejective uvular versus velar stops) if the distinctions were apparently nonlinguistic; that is, they could successfully discriminate vowel-less stops (which don't sound very much like speech), but failed to discriminate the same stops in syllable-initial position. Similarly, Japanese-speaking adults who cannot discriminate between [r] and [l] can make the (very similar) distinction between [r] and [w], sounds which are not contrasted in Japanese but are also not conflated (Best, MacKain, & Strange, 1982).

When non-native contrasts are similar to native contrasts it can be argued that auditory experience with the non-native sounds does occur in the native language—for instance, the sounds may occur as allophonic variations (e.g., unaspirated voiceless stops occur in English in postfricative position). Thus, the listener in effect has experience with assigning the sounds to native categories rather than to the categories which may be required for precise perception of a newly encountered language. Best, McRoberts, and Sithole (1988; see also Best, 1995) distinguished four principal possibilities for the mapping of a non-native phonemic contrast to a particular listener's native system:

1. The contrasting sounds are both assimilated to the same category in the listener's native language. This is the case with [r] and [l] for Japanese speakers, for instance. Such Single-Category Assimilations constitute the most difficult non-native contrast to perceive.
2. The contrasting sounds may be assimilated to the native language, but to different categories. This is the case with French unaspirated stops for English speakers. These Two-Category Assimilations are easy to perceive, because the sounds are assimilated to categories which also contrast in the native language.
3. One of the contrasting sounds assimilates well to a native category but the other does not. This is the case with Hindi dental versus retroflex stops for English speakers. These contrasts should appear as differences in relative "goodness of fit" to the native category. Best et al. (1988) predicted that such Category Goodness differences should be difficult but perceptible, especially if the goodness difference was extensive.
4. Finally, it might be the case that neither contrasting sound can be assimilated to the native phonemic space. Then the sounds may not even be heard as speech. Because the world's languages have selected their phonetic stock from a relatively limited range, such cases are rare; however, an example is found in the click contrasts characteristic of many African languages, which for speakers of nonclick languages are unlike any of their native sounds.

Best et al. (1988) argued that only the Nonassimilable case (as in number 4) would truly show whether adult listeners retain perceptual capabilities in the absence of any relevant auditory experience. Accordingly, they presented English-speaking listeners with Zulu click contrasts. Discrimination performance was excellent—as good, in fact, as that of adult Zulu speakers. Infants, needless to say, could also perform the same discrimination. Best et al. concluded that the perceptual narrowing for phonetic contrasts which is shown by adult listeners is not simply due to lack of relevant auditory experience, leading to sensori-neural atrophy or simply to loss of phonetic capabilities. These remain intact; what happens in the transition from prelinguistic infant to fully capable language user is a phonological reorganisation of speech sound percepts. Efficient recognition of speech requires rapid identification of (native) phonetic categories; it is therefore expedient for categorisation according to the native system to override superfluous phonetic precision.

Adult language users, therefore, have perceptual capacity which they do not always exploit. But although unused, the perceptual abilities are intact, and the discriminatory skills of highly expert phoneticians show that it is possible (at least for some speakers) to learn to ignore native phonological categorisations on occasion. Extended linguistic exposure can help even untrained speakers to improve their discrimination; Flege and Eefting (1987) found that Dutch native speakers with good English shifted the boundary of their categories [t] and [d] as a function of whether they thought they were listening to Dutch or English. However, some types of contrast are more difficult to acquire than others. Flege and Hillenbrand (1986) examined [s]-[z] discrimination in syllable-final position by speakers of English and French (in which this contrast occurs), and Swedish and Finnish (in which it does not occur). They separately manipulated vowel duration and fricative duration, and found that only English listeners fully exploited both cues; French listeners relied mainly on fricative duration (as per the French native cues), whereas the speakers of the languages without this contrast used vowel duration cues only. In this case, degree of exposure to English had no effect on the non-English-speakers' performance. Flege and Hillenbrand concluded that contrasts requiring integration of more than one cue are particularly difficult for adults to acquire.

Interestingly, cues which are indeed present in productions of a native contrast may not be exploited if they are not phonologically significant in the language. Gottfried and Beddor (1988) also found that French listeners did not make use of variations in vowel duration, this time as a cue to vowel identity. Their productions of the same vowel showed systematic variation in vowel duration as well as spectral information, but in perceptual identification they used the spectral information alone. English listeners presented with the same French stimuli, however, were able to make use of both temporal and spectral variation. Gottfried and Beddor explain these results by pointing to differences in the French and English vowel systems: in English, temporal and spectral information trades off to distinguish vowels (and indeed, the English listeners reported that they were assimilating the French stimuli to English categories); in French, however, duration is an unreliable cue to vowel identity, so French speakers have learned to ignore it.

To date, the model proposed by Best et al. (1988) makes the most explicit predictions concerning the effects of native phonetic categories on the perceptibility of second-language categories. Best (1995) refined and extended the original typology, allowing for the possibility of uncategorizable cases (second-language phonemes which map to a point in phonetic space distant from any native category). The model has been supported by data from Japanese perception of English (Best & Strange, 1992) and English perception of Zulu (Best, McRoberts, & Goodell, 2001; Best et al., 1988); it also accounts well for English perception of German vowels (Polka, 1995) and Japanese ratings of English consonants for "goodness of fit" to Japanese categories (Guion, Flege, Akahane-Yamada, & Pruitt, 2000).

It is true, however, that the model provides a better fit to data involving perception of a more complex system by listeners with a simpler native system—that is, where potentially larger native categories capture smaller second-language categories. The model effectively predicts the reverse case to be, in comparison, less problematic for the listener. For instance, perception of vowels in a language with a five-vowel system by listeners with a more densely populated native vowel space would be predicted by Best's (1995) model to involve assimilation to the nearest native peripheral vowel categories, all of them distinct, and hence misidentifications should be avoidable. As Escudero and Boersma (2002) point out, however, the existence of multiple native categories in the space occupied by tokens of a single second-language category may also cause problems. They presented Dutch listeners with tokens of Spanish /i/ and /e/ in labelling tasks; in one case, the listeners thought they were labelling Dutch tokens. Dutch has three vowels in the relevant area of vowel space occupied by the two Spanish vowels, and those listeners who more often used the three Dutch categories in labelling the vowel set also performed worse when identifying the same stimulus set in terms of the two Spanish categories.

As the next section describes, there are also other aspects in which the modelling of second language phonemic perception is in need of further elaboration. We consider the whole issue from the point of view which is also the most burning question for the average listener: what effect does insoluble imprecision have on success in understanding? That is, what are the consequences of inaccurate phonetic processing for the recognition of spoken second-language words?

THE WORST BEST CASE

> Variability in vowel perception ... a vowel located in one position on one occasion may well be located in a distinctly separate position one week, one day, or even ten minutes later (Laver, 1965, p. 95).

When two distinct categories of a second language map onto one native-language category, which corresponds better to neither of them, the perceptual situation is notoriously difficult. If the native category corresponds fairly well to one of the second-language categories, then discrimination may be assisted by a difference in category goodness between the two second-language sounds. But if both are possible approximations to the native category, the native category can potentially capture all tokens of either category in the second-language speech.

That is the case, for instance, with English /r/-/l/ for Japanese listeners or /æ/-/ɛ/ for Dutch listeners. In the former case, English distinguishes two kinds of voiced alveolar approximant, and Japanese has neither, but instead a voiced alveolar flap, which offers the only available category for capturing the English sounds—equally badly, but neither worse than the other. In the latter case, a similar situation arises, but with vowels: standard southern British English distinguishes two open midfront unrounded vowels, whereas Dutch has only one vowel in this part of the vowel space. The Dutch vowel is written with IPA's epsilon so that the transcription of English *neck* and Dutch *nek* is identical. However, in fact, the vowel in a *nek* is just that bit lower or more open than the vowel in a *neck*, bringing it closer to English /æ/ (as in *knack*), with the effect that, again, neither of the two southern British vowels maps well to the Dutch category but both are possible approximations. Dutch rhyme dictionaries list English loan words with both vowels as rhymes for Dutch words with /ɛ/ (e.g., besides *check* also *crack, snack* and *Jack* as rhymes for *nek*; Bakker, 1986, p. 163).

What happens to word recognition then? First, the listener is notoriously incapable of distinguishing between word pairs based on this distinction: ***knack*** and ***neck*** for the Dutch, ***right*** and ***light*** for the Japanese.

Second, phantom words could be recognised. There is no need for a minimal pair to satisfy the conditions for this, given the extent of spurious within-word embedding described earlier. Thus, hearing the word ***phantom*** may activate the pseudoembedding ***fan*** for any listener; and similarly, hearing ***chastise*** might activate ***chess*** for a Dutch listener and hearing ***regular*** might activate ***leg*** for a Japanese listener. There are no real English words ***chass*** or ***reg***, so this is a case where no comparable competition arises for the native listener.

Does this happen? Is ***chass*** perceived as a token of ***chess*** by a listener with Dutch as first language?

One way to find out is simply to ask non-native speakers if ***chass*** is a word. Unfair, perhaps, but psycholinguists actually spend a lot of their time presenting lists of nonwords and words and asking people to decide which is which. This procedure is called the lexical decision task, and was described earlier. It is probably the most widely used laboratory method in experimental psycholinguistics. Accordingly, we approached the present question via a lexical decision experiment (Broersma, 2002) in which Dutch and English listeners were presented with real English words (***share, wish***), clear cases of English nonwords (***plog, strisp***), and what we called "near words:" spoken forms such as ***chass***. Twenty-four native speakers of Dutch and 24 native speakers of British English (from the University of Birmingham) took part. The Dutch participants had a high level of proficiency in their second language English, whereas the English participants did not know any Dutch.

Altogether there were 32 monosyllabic English words involved in the /æ/-/ɛ/ comparison, 16 with /ɛ/ (***chess, desk***) and 16 with /æ/ (***fact, gang***). Further criteria for selection were that the word did not sound like an existing Dutch word, and that replacement of the target phoneme with its confusable counterpart did not result in an existing English or Dutch word. Near words were formed by replacing the target phoneme with its confusable counterpart (e.g., ***chess*** became ***chass*** and ***gang*** became ***geng***). The target words were divided into two lists, balanced for frequency of occurrence, and contained equal numbers of each of the two base vowels. For both the Dutch and the English groups, half the listeners heard the words from one list in their real-word form and those from the other list in near-word form, whereas the other half heard the reverse mapping. Thus, if *chass* was indeed accepted as a real word, we could compare how often this happened in comparison with the acceptance of the correct pronunciation *chess.*

With real English words and clear nonwords as filler items, the experiment contained 200 items in all. Participants read instructions about the task, then heard the items binaurally over headphones; after each item was heard, they pressed one response button if they thought the presented item was an English word and another if they thought it was not. Presentation of the next item started 500 msec after the response was given.

The results were very clear, as can be seen from Fig. 4.2. First, the Dutch were very good at performing lexical decisions on English input. Their rate of positive responses to real words and of negative responses to clear nonwords did not differ from that of native Eng-

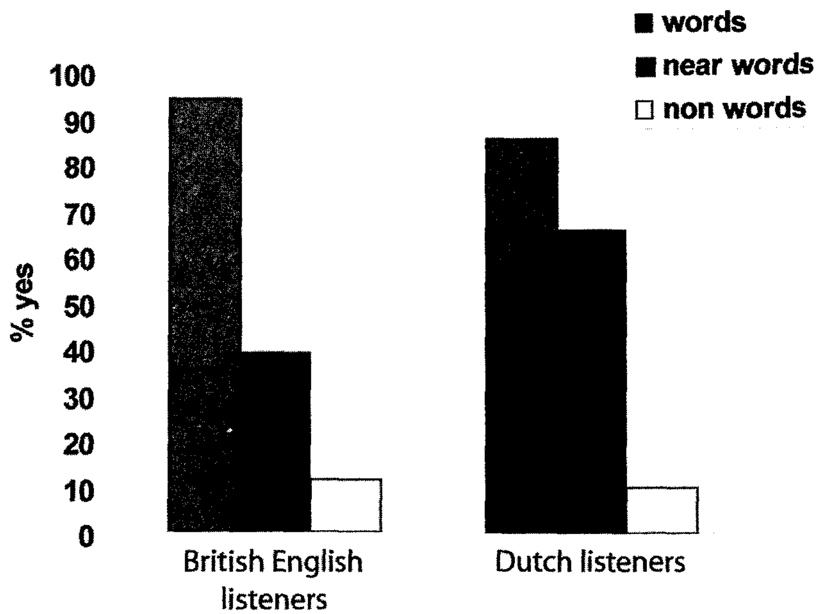

FIG. 4.2. Mean percentage of YES responses to real English words containing /æ/ or /ɛ/ (e.g., *chess, gang*), near words formed from them (e.g., *chass, geng*), and clear nonwords (e.g., *plog, strisp*) for native British English listeners and Dutch-native listeners (from Broersma, 2002).

lish speakers. Second, the near words were indeed rather confusing even for the native speakers, who were more likely to make a false positive response to a near word than to a clear nonword. Third, however, the Dutch listeners were significantly more likely to fall into this sneaky experimenter-set trap. In 66% of the cases, they simply responded yes to near words like *chass* and *geng* (significantly more often than did native listeners). That is, for these Dutch listeners, most of the tokens of *chass, geng,* and the like, were as good as canonical English words.

An extension of the same type of test to a competition situation was carried out in a second experiment within the same study (Broersma, 2002). This experiment involved minimal pairs based on the same vowel contrasts: *flash-flesh, mansion-mention,* and the like. Here a cross-modal priming paradigm was used: the task was again lexical decision, but this time with visually presented targets, and these were preceded by spoken primes. With visually presented targets, we assume that the Dutch listeners and the native listeners will not differ substantially in how they process the form of the targets, so that the dependent variable is, in this case, time to make a response rather than proportion of positive versus negative responses. Crucially, we compare the effects of a previously presented prime which is the same word as the target (e.g., FLESH preceded by spoken *flesh*) versus a prime which is the target's minimal pair (e.g., FLESH preceded by spoken *flash*), plus a control condition in which the prime is phonologically and semantically unrelated to the target (e.g., FLESH preceded by spoken *spite*). As described earlier, this kind of experiment usually shows that targets preceded by a minimally mismatching prime are responded to least rapidly, due to inhibitory effects (Soto–Faraco et al., 2001); note that vowel and consonant mismatches exercise equivalent inhibition (thus, Soto–Faraco et al.'s 2001 study included pairs like *sardina-sardana* as well as pairs like *tragico-trafico*).

Seventy-two native speakers of Dutch and 72 native speakers of British English, from the same populations used in the preceding experiment, took part. There were more participants because there were more conditions to counterbalance in this case: each target word could be preceded by an identical prime, a minimal-pair prime, or a control prime, and for each minimal pair of words either one could be the target word, but any one participant should only be presented with one member of the pair. This made for six presentation conditions in the experiment. Again, the potential target words were divided across conditions in such a way that the frequency of occurrence of the target words was balanced as closely as possible.

As in the auditory lexical decision experiment, none of the items were phonetically similar to any Dutch words, and in addition, none of the words that were used as visual targets in this set of materials were orthographically similar to Dutch words. The experiment contained in all 168 trials, including many filler trials with word and nonword targets. The listeners heard the prime words over headphones, then saw a target word on a computer screen; as before, they were asked to decide if the target word was an existing English word, and to indicate their response through button press.

The mean response times for each group in each condition can be seen in Fig. 4.3. As expected, the English native speakers showed

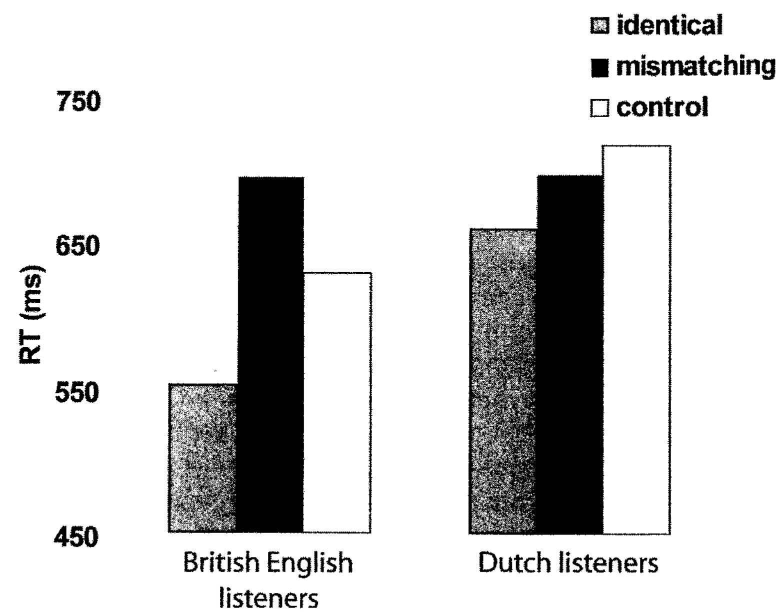

FIG. 4.3. Mean response time (msec) to visually presented English words containing /æ/ or /ɛ/ (e.g., *flesh*) preceded by an identical spoken prime (e.g., *flesh*), a minimally mismatching prime (e.g., *flash*), or a control prime (e.g., *spite*), respectively, for native British English listeners and Dutch-native listeners (from Broersma, 2002).

the well-known dual pattern of facilitation and inhibition: compared with the control condition, responses to targets were significantly faster after matching primes but significantly slower after minimally mismatching primes. This is exactly the pattern that was found in the experiments of Soto–Faraco et al. (2001) and others. The Dutch listeners, however, showed a different pattern of results. They did show significant facilitation in the matching condition in comparison with the control condition, but there was no trace of inhibition (and there was no facilitation) in the minimally mismatching condition. Their responses were, naturally, slower than those of the native speakers (this can be most easily seen in a comparison of the two control condition means); but the important aspect of the results is how each group's matching-prime and mismatching-prime conditions compare to the control condition, and in this the two groups clearly differ.

The Dutch listeners' pattern can be explained by assuming that competition between prime and target words remained unresolved. Neither one of the lexical competitors obtained sufficient activation to inhibit the other. However, it is not just the case that Dutch listeners treat English /æ/ and /ɛ/ as equally valid exemplars of a single category (and, accordingly, *flash* and *flesh* as homophones). In the auditory lexical decision experiment, as we saw, over one third of near words were correctly rejected by these listeners. In this experiment, too, responses in the minimally mismatching condition were longer than those in the matching condition, which implies that the primes must have activated identical targets more strongly than minimal-pair targets. This, in turn, must mean that the non-native listeners succeeded in differentiating to some extent between the English vowels, albeit not as clearly as did the native listeners. We suggest that the non-native listeners may maintain separate vowel categories for /æ/ and /ɛ/, but these will not be as distinct as those of the native listeners; indeed, they may even overlap, which would allow for (weaker) support to be given to the minimally mismatching words. Thus, *flesh* would activate *flesh* more than *flash* for native and non-native listeners alike; however, whereas for native listeners *flesh* would decisively mismatch *flash*, for non-native listeners *flash* would be partially matched by *flesh*. The mismatch of *flash* would then not be large enough to lead to inhibition. Phonetic imprecision would thus extend the competition process in non-native word recognition.

However, although this kind of explanation can be easily applied to the worst case in Best's (1995) classification, in which two second-language categories cover space containing only one native language category, there are other situations in which it seems less obviously applicable. The following section describes such a case.

ANOTHER NON-BEST CASE

> The brain obviously transmits different neural commands for a phoneme occurring in initial as opposed to final syllable position (Laver, 1970, p. 71).

In fact Best's classification, as Escudero and Boersma (2002) also observed, does not fully cover the range of possibilities. Consider a distinction which is made in the same way in the first and the second language, but is restricted in its distribution in the native language such that listeners of that language have experience with making the distinction only in certain positions or under certain conditions. An example here is the voicing distinction for obstruents. This is independent of syllable position in English: *dough, toe, ode,* and *oat* are all words and no two of them are homophonous. In German and Dutch, however, a syllable-final devoicing rule bars voicing contrast in final position: German *Rat* and *Rad* are homophones, as are Dutch *raat* and *raad.*[1] In syllable-initial position, on the other hand, the contrast is as effective as in English: German *Deich* and *Teich* and Dutch *dij* and *tij* are nonhomophonous minimal pairs.

This situation is clearly not a cut-and-dried mapping of one set of categories exactly or inexactly against another. German or Dutch users of English will be able to map English obstruents contrasting in voicing against their native categories for [t] versus [d], [p] versus [b], [f] versus [v], and so forth. They will have had plenty of native experience in making these distinctions. Yet this experience will have been limited to syllable-initial contrasts, and in their native language at least they will never have needed to attend to a voicing distinction in syllable-final position. As a result, although they will be familiar with the English distinction, they may simply overlook it in positions in which it never occurs in their native language. This may be because they cannot attend to the distinction in those positions, or because awareness at a phonological level is insufficient, and to be useful must be accompanied by relevant phonetic processing experience; English correlates of the voicing distinction differ in syllable-initial versus syllable-final position, and listeners unfamiliar with the cues used in final position, such as preceding vowel duration, may be unable to exploit them. (Of course, both inattention and insufficient experience may exercise simultaneous effect!)

Indirect evidence from loan words in rhyme dictionaries can be again adduced in support of the complete irrelevance of the voicing distinction in final position: not only *snack* but also *shag* is listed as a possible rhyme for *nek,* both *sweet* and *tweed* as rhymes for Dutch *friet* ('chips'), and both *cop* and *job* as rhymes for Dutch *drop* ('licorice;' Bakker, 1986, pp. 163, 238, 330).

Thus, voicing contrasts in syllable-final position may present our Dutch listeners with exactly the same problems as the English vowel contrasts discussed earlier; they may consider *glope, sice,* and *quode* to be words in the same way as they misclassified *chass* and *geng.* We tested this in the auditory lexical decision experiment reported earlier, which contained 32 near words differing vocalically from real English words, and 32 near words differing from real words in the voicing of a final obstruent.

Fig. 4.4 shows the results (with clear nonwords again included to facilitate comparison). In this case, the native listeners had no problem making the relevant distinctions and hardly ever produced a false positive response to a near word like *glope.* However, the Dutch listeners did. Again, a majority of these items received a false positive response, that is, were treated as words.

The cross-modal priming experiment which we reported in the preceding section also examined English syllable-final voicing distinctions. Minimal pairs such as *bride* versus *bright, phase* versus *face,* and *robe* versus *rope* formed part of the stimulus set, and again, either member of a pair could appear as a visual target, preceded by a matching, minimally mismatching, or control spoken prime (thus BRIDE could be preceded by spoken *bride, bright,* or *shave*). The results for this group of items are shown in Fig. 4.5. It can be seen that obstruent voicing contrasts produced the same pattern of results as vowel contrasts: for the native listeners, facilitation and inhibition appeared in tandem, again both significantly differing

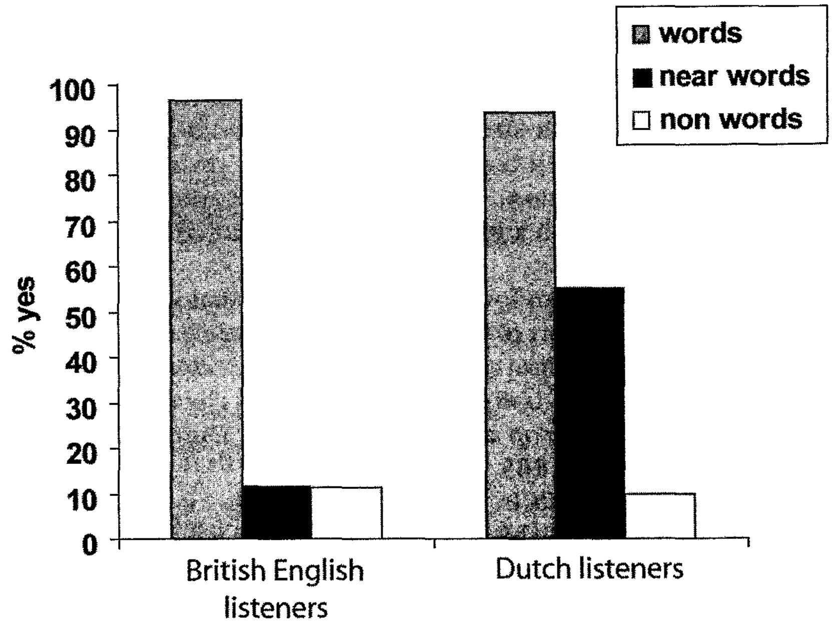

FIG. 4.4. Mean percentage of YES responses to real English words containing final obstruents (e.g., *globe, quote*), near words formed from them (e.g., *glope, quode*), and clear nonwords (e.g., *plog, strisp*) for native British English listeners and Dutch-native listeners (from Broersma, 2002).

from the control condition; but for the Dutch listeners, there was facilitation by matching primes but no inhibition from minimally mismatching primes.

Thus, the position-specific mismatch between the repertoires of English and Dutch voicing contrasts seems to cause the same problems for spoken-word recognition as the complete absence of a second-language contrast from the native repertoire: near words can treacherously pass for real words, and competition between minimal pairs can be unnecessarily extended. Our results suggest that comparisons across the phonemic repertoires of languages cannot be made on the basis of the category repertoire alone; Best's (1995) classification must be separately applied for syllable-initial versus syllable-final position (and possibly according to further relevant differences affecting contrast occurrence). The category repertoire may match in one area of

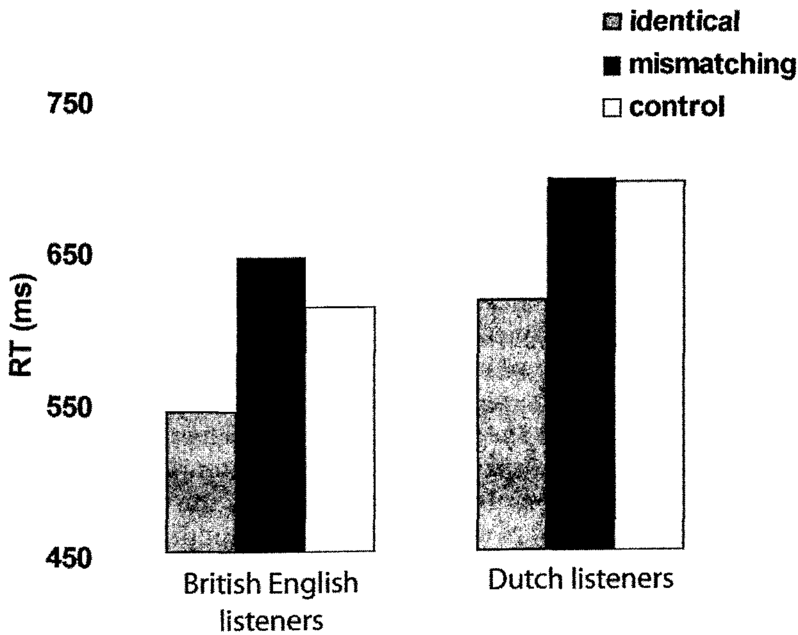

FIG. 4.5. Mean response time (msec) to visually presented English words containing final obstruents (e.g., BRIDE) preceded by an identical spoken prime (e.g., *bride*), a minimally mismatching prime (e.g., *bright*), or a control prime (e.g., *shave*), respectively, for native British English listeners and Dutch-native listeners (from Broersma, 2002).

application but mismatch in another, leading to overlapping representations in the latter situation only.

CONCLUSION

> The perceptual skills that allow us to participate successfully in face-to-face conversation are very complex (Laver, 1976, p. 345).

Phonetic processing in listening to speech is a balancing act in which the aim is to achieve the highest precision where it does good, and abandon precision where it is not necessary. This general conclusion holds good as well for listening to a second language as for listening to speech in the native tongue. Yet the degree to which listeners achieve the necessary balance in a second language falls far short of their success with the native language.

Where the second language makes distinctions which are not made in the native language, the native language can override the distinction; and where the native language makes distinctions which are not made in the second language, these distinctions can intrude despite their irrelevance. Over decades of research, speech perception investigators have amassed a formidable body of evidence of this listening imprecision. However, for the second language user, the consequences are conspicuous not so much in the phonetic realm as in the effects on communication. Imprecise phonetic processing delays communication.

Comprehension of speech involves rapid selection of the correct sequence of words from a panoply of potential word candidates which are fully or in part supported by the spoken input. The more candidates are active, the more competition arises, and the more competition there is (and the longer it lasts), the more difficult successful word recognition becomes. Imprecise phonetic decisions can result, as we have seen, in phonetic sequences which are in fact not words at all being recognised as words: a Dutch listener who hears *chass* or *glope* is likely to activate the English words *chess* and *globe*.

Of course, we hope that native speakers do not in general go around uttering nonwords to hapless non-native listeners. We (two authors who grapple daily with our imprecise perception of our respective second languages) would beg readers of this book never to take up such a tormenting pastime. But the continuity of speech and the construction of vocabularies make it highly likely that non-native listeners will hear relevant near words anyway. Thus, *chass* is embedded within *chastise*, as we pointed out; likewise, *eace* occurs in *easterly* and *sice* in *precise*. The native speaker should experience little unwarranted competition from *chess, ease,* and *size* in these contexts, but our results suggest that many non-native listeners will. Not only within but across words such sequences can be found: where /æ/ and /ɛ/ cannot be distinguished, *stamp* may be activated in *the last emperor,* and where voiceless final obstruents are not distinct from voiced, then *cheese* may be heard in *each Easter.* Combine the two cases and *friend* is activated in *frantic* and *sand* in *this entrance.*

On top of the near-word recognition problem there is also the failure to distinguish minimal pairs incorporating the distinctions in question. For Japanese listeners, as we pointed out earlier, English pairs requiring the /r/-/l/ distinction, such as *right* and *light* or *brew* and *blue*, may collapse to a single phonetic form. Likewise the Dutch listeners to English in our experiments can experience the same problem with *flash* and *flesh*, or with *hard* and *heart*, or *beck* and *beg*. Speakers of languages which do not maintain intervocalic voicing distinctions will confuse English pairs such as *waver* and *wafer*, or perhaps hear *Laver* in *souffle for dinner*. All of these consequences of phonetic imprecision in listening combine to exacerbate the problem of interword competition for the non-native listener. Speech, the "most intricate of our communicative skills," requires from its users quite formidable precision in determining when and when not to be precise.

NOTE

[1] Extensive research has in fact shown that in both languages there is evidence of incomplete neutralisation: Port and O'Dell, 1985, for German; Warner, Jongman, Sereno, and Kemps, 2004, for Dutch. However, in each case the incomplete neutralisation effects seem to be in part a reflection of orthographic awareness, and native speakers of each language certainly consider the relevant minimal pairs to be homophones, for example, for use as rhymes. For ease of exposition of the experiments reported here, which concerned the perceptibility of English syllable-final voicing contrasts only, we also pass over all other aspects of the obstruent phonology of German and Dutch, for example, restrictions on certain fricative voicing contrasts in syllable-initial position as well, availability of voicing assimilation in certain contexts, and absence of some English obstruent categories from the German and Dutch repertoires.

REFERENCES

Alphen, P. M. van, & McQueen, J. M. (2003). *The effect of voice onset time differences on lexical access.* Manuscript submitted for publication.

Bakker, J. (Ed.). (1986). *Nederlands rijmwoordenboek* [Dutch rhyming dictionary]. Amsterdam: Ooievaar Pockethouse.

Bauer, L. (1985). Tracking phonetic change in the received pronunciation of British English. *Journal of Phonetics, 13*, 61–81.

Best, C. T. (1995). A direct realist view of cross-language speech perception: Standing at the crossroads. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.

Best, C. T., MacKain, K. S., & Strange, W. (1982). A cross-language study of categorical perception for semi-vowel and liquid glide contrasts. *Journal of the Acoustical Society of America, 71* (Suppl. 1), S76.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109,* 775–794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance, 14,* 345–360.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20,* 305–330.

Broersma, M. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. *Proceedings of the 7th International Conference on Spoken Language Processing,* (pp. 261–264). Denver, Colorado: Center for Spoken Language Research.

Burnham, D. K., Earnshaw, L. J., & Clark, J. (1991). Development of categorical identification of native and non-native bilabial stops: Infants, children and adults. *Journal of Child Language, 18,* 231–260.

Cutler, A., McQueen, J. M., Jansonius, M., & Bayerl, S. (2002). The lexical statistics of competitor activation in spoken-word recognition. *Proceedings of the 9th Australian International Conference on Speech Science and Technology, Melbourne,* 40–45.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16,* 507–534.

Escudero, P., & Boersma, P. (2002). *The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish.* Paper presented at the 26th annual Boston University Conference on Language Development.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6,* 185–202.

Flege, J. E., & Hillenbrand, J. (1986). Differential use of temporal cues to the /s/–/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America, 79,* 508–517.

Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language & Speech, 31,* 57–75.

Guion, S. G., Flege, J. E., Akahane–Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107,* 2711–2724.

Harrington, J., Palethorpe, S., & Watson, C. I. (2000). Does the Queen speak the Queen's English? *Nature, 408,* 927–928.

Labov, W. (1994). *Principles of linguistic change.* Cambridge, MA: Blackwell.

Laver, J. (1965). Variability in vowel perception. *Language and Speech, 8,* 95–121.

Laver, J. (1970). The production of speech. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 53–75). Harmondsworth, Middlesex, England: Penguin.

Laver, J. (1976). Language and non-verbal communication. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception, Vol. 7, Language and Speech* (pp. 345–363). New York: Academic.

Laver, J. (1989). Cognitive science and speech: A framework for research. In H. Schnelle & N.–O. Bernsen (Eds.), *Logic and linguistics, vol. 2* (pp. 37–69). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Laver, J. (1994). *Principles of phonetics.* Cambridge, England: Cambridge University Press.

Laver, J., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. In K. R. Scherer & H. Giles (Eds.), *European studies in social psychology: Social markers in speech* (pp. 1–32). Cambridge, England: Cambridge University Press.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/ III: Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America, 96,* 2076–2087.

Logan, J. S., Lively, S. E., & Pisoni, D. B (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America, 89,* 874–886.

Maddieson, I. (1984). *Patterns of sounds.* Cambridge, England: Cambridge University Press.

Marslen–Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101,* 653–675.

Martin, J. G., & Bunnell, H. T. (1981). Perception of anticipatory coarticulation effects. *Journal of the Acoustical Society of America, 69,* 559–567.

Martin, J. G., & Bunnell, H. T. (1982). Perception of anticipatory coarticulation effects in vowel-stop consonant–vowel sequences. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 473–488.

McQueen, J. M. (2004). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The handbook of cognition* (pp. 255–275). London: Sage.

McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes, 10,* 309–331.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 621–638.

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision-making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1363–1389.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 1209–1228.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.

Pallier, C., Colomé, A., & Sebastián–Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12,* 445–449.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology* 7 (pp. 101–139). Berlin, Germany: Mouton de Gruyter.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America, 97,* 1286–1296.

Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics, 13,* 455–471.

Soto–Faraco, S., Sebastián–Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45,* 412–432.

Strange, W. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustical Society of America, 85,* 2135–2153.

Strange, W. (Ed.), (1995). *Speech perception and linguistic experience: Issues in cross-language research.* Baltimore: York Press.

Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America, 65,* 1533–1541.

Utman, J. A., Blumstein, S. E., & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics, 62,* 1297–1311.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics, 32,* 251–276.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672–683.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37,* 278–286.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75,* 1866–1878.

Whalen, D. H. (1984). Subcategorical mismatches slow phonetic judgments. *Perception & Psychophysics, 35,* 49–64.

Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics, 50,* 351–360.

Yaeger–Dror, M. (1994). Phonetic evidence for sound change in Quebec French. In P. A. Keating (Ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonology III* (pp. 267–292). Cambridge, England: Cambridge University Press.