Available online at www.sciencedirect.com
ScienceDirect
Speech Communication 52 (2010) 980–995
SPEECH COMMUNICATION
www.elsevier.com/locate/specom

Native and non-native listeners' perception of English consonants in different types of noise

Mirjam Broersma [a,b,*], Odette Scharenborg [c]

[a] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
[b] Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[c] Centre for Language and Speech Technology, Radboud University Nijmegen, P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

Abstract

This paper shows that the effect of different types of noise on recognition of different phonemes by native versus non-native listeners is highly variable, even within classes of phonemes with the same manner or place of articulation. In a phoneme identification experiment, English and Dutch listeners heard all 24 English consonants in VCV stimuli in quiet and in three types of noise: competing talker, speech-shaped noise, and modulated speech-shaped noise (all with SNRs of −6 dB). Differential effects of noise type for English and Dutch listeners were found for eight consonants (/p t k g m n ŋ r/) but not for the other 16 consonants. For those eight consonants, effects were again highly variable: each noise type hindered non-native listeners more than native listeners for some of the target sounds, but none of the noise types did so for all of the target sounds, not even for phonemes with the same manner or place of articulation. The results imply that the noise types employed will strongly affect the outcomes of any study of native and non-native speech perception in noise.

Keywords: Speech Perception; Consonants; Identification; Noise; Non-native Language

1. Introduction

Noise comes in various kinds, and people communicate in all of them. Apart from differences in loudness of the noise (e.g., rustling leaves versus a rocket launch), particular types of noise might also disrupt communication more than others; e.g., factory noise might hinder speech perception more than the din of voices in a shopping mall does (cf. Cooke and Scharenborg, 2008). Furthermore, how noise affects perception also depends on the language background of the listener: noise hinders perception of a second language (L2) more than perception of the native language (L1) (e.g., Garcia Lecumberri and Cooke, 2006; Mayo et al., 1997; Van Wijngaarden et al., 2002). But are native and non-native listeners affected differently when they converse in a factory versus a shopping mall? Are particular speech sounds perceived more accurately in a factory than in a shopping mall? And crucially, are some speech sounds even perceived more accurately in a factory by native listeners, and more accurately in a shopping mall by non-native listeners? In this paper, we investigate to what extent different types of noise affect native and non-native listeners differently for the perception of different speech sounds.

Many studies of speech perception in noise, especially those that do not specifically aim to investigate the effect of particular noise types, use only a single type of noise. Most studies of native and non-native listeners' speech perception in noise so far have also used one type of noise, in some cases at varying signal-to-noise ratios (SNRs). Which type of noise was used varies between studies, including, for example, white noise (Bradlow and Bent, 2002), speech-shaped noise (Bradlow and Alexander, 2007; Hazan and Simpson, 2000), and multi-talker babble (Cutler et al., 2008, 2004).

* Corresponding author at: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.
E-mail addresses: mirjam@mirjambroersma.nl (M. Broersma), o.scharenborg@let.ru.nl (O. Scharenborg).

0167-6393/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.specom.2010.08.010

Such studies have typically shown that non-native listeners' perception suffers more from masking noise than native listeners' perception, in a range of tasks at various levels of processing, from sentence intelligibility (Bradlow and Bent, 2002; Cooke et al., 2008; Mayo et al., 1997; Van Wijngaarden, 2001; Van Wijngaarden et al., 2002) to word identification (Nábělek and Donahue, 1984) and phoneme identification (Cutler et al., 2008; Garcia Lecumberri and Cooke, 2006; Garcia Lecumberri et al., 2008; Hazan and Simpson, 2000). This native listening advantage is at least partially due to a more efficient use of higher level information, to compensate for the loss of intelligibility at lower levels of processing. Cutler and colleagues, e.g., found no difference between native and non-native listeners' perception of phonemes in noise (Cutler et al., 2004, 2005) unless there was even the smallest amount of predictability in the form of a constant noise lead duration that native listeners could benefit from (Cutler et al., 2008). Further, for native but not for non-native listeners, noise hindered the recognition of words less when they were preceded by a semantically related word than when they were preceded by a semantically unrelated word (Golestani et al., 2009). Similarly, for word recognition in sentences in noise, the largest differences between native and non-native listeners occurred where words were predictable from the semantic context (Mayo et al., 1997). Further, non-native listeners require a clearer signal, with less severe noise, than native listeners do in order to be able to exploit contextual information for the recognition of words in sentences (Bradlow and Alexander, 2007).

We know of only few studies that compare native and non-native listeners' perception of speech in different types of noise. Garcia Lecumberri and Cooke (2006, 2008) compared native and non-native listeners' perception of English intervocalic consonants in three and six types of noise, respectively. In both studies the target language was English; in the former study, non-native listeners were Spanish, and in the latter, eight groups of non-native listeners were tested. Both studies showed that in a consonant identification task, non-native listeners suffered more from noise than native listeners did in all noise types and that the order of the difficulty of the noise types (collapsing the results over all target consonants) was the same for all listener groups. Cooke et al. (2008) investigated native and non-native listeners' recognition of words in sentences in speech-shaped noise and competing talker noise. They found that non-native listeners performed consistently less accurately than native listeners did, and that the characteristics of the competing talker noise modified this difference: the difference between L1 and L2 listeners' performance was larger when the target signal and the noise were from same gender speakers than when they were from different gender speakers.

A large body of research confirms that different types of noise affect speech perception differently, even at the same SNR. Speech-shaped noise, for example, hinders speech perception more than speech from a competing talker at the same SNR, with a difference in speech-reception thresholds of around 6–8 dB (Festen and Plomp, 1990). Factory noise, in turn, hinders speech perception more than speech-shaped noise (Cooke and Scharenborg, 2008). Even within the seemingly homogeneous category of multi-speaker babble, large differences are found in the way different numbers of competing speakers affect speech perception. Simpson and Cooke (2005) showed that perception of the target speech decreased when the number of speakers in a multi-speaker babble masker increased from one to eight, stayed stable for 8–128 speakers, and then recovered when the number of speakers further increased up to 512 speakers. Further, speech perception is less accurate when the multi-speaker babble is variable than when one sample of babble is presented repeatedly (Felty et al., 2009). Van Engen and Bradlow (2007) further showed that for two-speaker but not for six-speaker babble, babble in the same language as the target speech hindered sentence comprehension more than babble in a different language which was unknown to the listeners.

As different types of noise vary in their spectrotemporal make-up, they vary in the "glimpses" that they provide of the target speech, where the energy of the target speech exceeds that of the masker (Cooke, 2006). Spectral and temporal energy modulations in the noise are thus important for its masking effect, and vary for different noise types. Importantly, therefore, different noise types may also affect different phonemes differently. Phatak et al. (2008) showed that white noise hindered speech perception more than speech-weighted noise, but individual phonemes differed considerably in the extent to which perception decreased from one noise type to the other.

There are thus clear indications that different noise types affect perception of different phonemes differently, and that noise affects native versus non-native listeners differently. In the present study, we combine those two findings and ask whether different noise types affect recognition of different phonemes differently for native versus non-native listeners.

Each phoneme has its own phonetic characteristics and each phoneme distinction its own perceptual cues. Native and non-native listeners, in turn, differ in their use and weighting of perceptual cues for each phoneme distinction (see, e.g., the collected papers in Bohn and Munro (2007) and Strange (1995)). Different noise types, finally, vary in the extent to which they mask particular perceptual cues. Therefore, each noise type might affect recognition of each phoneme for listeners with each language background differently.

We hypothesize that the effects of different types of noise on recognition of different phonemes by different types of listeners are highly variable. We expect that there is no single type of noise that affects non-native listeners' perception more than native listeners' perception for all phonemes, but rather that each type of noise will affect non-native listeners' perception more than native listeners' perception for some phonemes, but not for others. We further hypothesize that the effects of different types of noise on recognition of different phonemes by different types of listeners also vary within classes of phonemes with the same manner or place of articulation. Due to the diversity of perceptual cues that play a role, also for perception of phonemes with the same manner or place of articulation, and due to the wide range of possible differences between native and non-native listeners' use of perceptual cues (see Bohn and Munro, 2007; Strange, 1995), we thus expect that (a) effects of different noise types on perception of different phonemes by native versus non-native listeners will show variable results and (b) grouping phonemes by manner or place of articulation will not result in less variable patterns. Rather, each combination of noise type, phoneme, and listener type might yield a unique outcome.

Such a finding would have clear implications for studies of speech perception in noise, and in particular studies involving both native and non-native listeners. It would mean that the outcome of those studies strongly depends on the type of noise employed and cannot be generalized to other noise types.

To test those hypotheses, we investigate consonant identification in three different types of noise at a fixed SNR, namely speech-shaped noise, modulated speech-shaped noise, and competing talker noise, and in a quiet baseline condition. Perception of English intervocalic consonants by native listeners of Dutch, who had a high level of proficiency in English as a second language, was compared to perception by native listeners of British English. All 24 English consonants were included in the study. We first assess for which consonants the different listening conditions led to different effects for native and non-native listeners. Our analyses then focus on those consonants that show such differential effects for native versus non-native listeners, and assess in detail for each consonant how each type of noise affects English and Dutch listeners' perception.

2. Method

2.1. Participants

Participants were 18 native Dutch-speaking students and staff members (age range: 21–62, average: 36.1) from the Radboud University Nijmegen, the Netherlands, and 21 native English-speaking students and staff members (age range: 19–48, average: 31.9) from the University of Sheffield, UK. None reported any hearing loss, visual loss, or reading disability. The Dutch participants all received minimally six years of English education in primary and secondary schooling, and were very regularly exposed to written and spoken English through the media and in their work or education. Prior to the test, the Dutch participants were asked to indicate their English competence level on a four point scale from '1 = basic' to '4 = fluent'. All participants reported a proficiency level of at least 2; the average level was 3.0. Participants were volunteers and were not paid for their participation, but in order to encourage them to do well in the test, the highest scoring participant on each test set (separate for the English and Dutch participants) was given a small prize. All participants were phonetically naive.

2.2. Materials

The materials used in this study were recorded for the Interspeech 2008 Consonant Challenge (Cooke and Scharenborg, 2008). The materials consisted of all 24 English consonants in nine intervocalic contexts (VCV) consisting of all possible combinations of the three vowels /iː/ (as in "beat"), /uː/ (as in "boot"), and /æ/ (as in "bat"). The 24 English consonants were /p/, /t/, /k/, /b/, /d/, /g/, /tʃ/, /ʤ/, /m/, /n/, /ŋ/, /f/, /v/, /θ/, /ð/, /s/, /z/, /ʃ/, /ʒ/, /h/, /r/, /j/, /w/, /l/. Each VCV was produced using both initial and final stress (e.g. ˈaba versus abˈa). The materials were recorded by four female and four male native speakers of British English originating from various regions of the UK. None had a strong regional accent. Recordings were made in an IAC single-walled acoustically isolated booth at the University of Sheffield. Speakers produced VCVs in isolation by reading out tokens presented on a computer screen at a normal speaking rate, and were given both verbal and written instructions on how to interpret token names, with a particular focus on /θ/, /ð/, /ʤ/, and /ʒ/.

Seven test sets were produced, four of which are reported in this paper: one quiet and three noise conditions. Each test set contained two instances of each of the 24 consonants from each of the eight speakers resulting in 384 VCV items per test set. The VCV items were randomly distributed over the test sets with the restriction that the frequency of the nine vowel contexts and the two stress conditions was equal in all test sets. Each VCV item occurred in only one test set.

The three noise types reported in this paper were competing talker (COMP), speech-shaped noise (SSN), and modulated speech-shaped noise (MODSSN). All had a SNR of −6 dB. (The other three noise types are not reported here as they had different SNRs, of −3, −2, and 0.) The three noise backgrounds provide different types of spectral and temporal masking. SSN is a stationary noise with fixed spectral dips and no significant temporal modulations. It was generated by passing white noise through a 50-coefficient filter derived from the LPC spectrum resulting from the sum of 200 British English sentences taken from the Lu (2010) corpus. This corpus had been recorded in the same facility and employed identical post-processing as those used for the VCV materials, ensuring matched target and masker long-term spectra. COMP and MODSSN are non-stationary and have a long-term spectrum equivalent to that of speech, with significant modulations in both frequency and time. For COMP, masker signals were randomly chosen segments from sentences from eight talkers (four male, four female) from the Lu (2010) corpus. MODSSN, finally, shares its spectral shape with SSN but has temporal envelope modulations derived from a speech signal. It contains no intelligible components. MODSSN differs from SSN in permitting occasional clear glimpses of the signal. To generate MODSSN noise maskers, envelopes from random segments of competing speech (from Lu (2010)) were multiplied sample-wise with fragments of speech-shaped noise.

For the noise conditions, VCV tokens were added to noise samples of 1200 ms in duration. In order to make the start of the VCV unpredictable in the noise, the onset time of the VCV relative to the noise was varied: the onset took one of eight values linearly-spaced in the range 0–400 ms. Each consonant occurred the same number of times at each of the eight onsets. For each VCV token, the noise signal was scaled to produce the required SNR in the region where the speech was present. Further information on the creation of the test material can be found in Cooke and Scharenborg (2008).

2.3. Experimental set-up

Fig. 1 shows the screen layout that was presented to the participants. All 24 consonants were represented by their most logical and frequent grapheme combination in English with an example word in English below it containing the sound. Graphemes rather than phonetic symbols were used since data was collected from phonetically naive subjects (cf. Garcia Lecumberri et al., 2008).

Participants were tested individually or in little groups of maximally six participants in a quiet room. They were instructed in their native language to listen to the VCV, decide on the identity of the consonant, and indicate their decision as fast and accurately as possible by clicking on the appropriate consonant on the screen layout (Fig. 1) using a mouse.

The VCV stimuli were presented over closed high-quality headphones, one at a time in random order, in seven blocks; one for each condition. The experiment was self-paced, but no pauses were permitted during a block. Between blocks, participants were encouraged to take a short break. The test consisted of two sessions of on average 1 h. At the start of the first session, participants underwent a short practice session with tokens in QUIET condition. For the actual test, listeners started with the QUIET condition, followed by two noise conditions. In the second session, the remaining four noise conditions were tested. The order of the noise conditions was randomized.

Fig. 1. Experimental screen layout.

3. Results and discussion

The Appendix shows the confusion matrices for both listener groups and the four listening conditions. First, it was determined for which of the 24 target sounds the four conditions affected Dutch and English listeners differentially. To that end, Analyses of Variance (ANOVAs) were done for all target sounds, with proportion of correct responses as dependent variable, Native Language as between subjects factor and Condition as within subjects factor. These were planned comparisons, but α was set to .01 rather than .05 in order to adjust for the rather large number of comparisons. For eight of the target sounds, the conditions affected the Dutch and English listeners differently, as shown by a significant interaction between Native Language and Condition (Table 1). These were the voiceless stops /p t k/, the voiced stop /g/, the nasals /m n ŋ/, and /r/. Those eight target sounds were further analysed.

For each of those eight target sounds, confusion patterns were assessed. The most common errors as well as the clearest differences between Dutch and English listeners' errors (i.e., differences larger than 5%) are discussed. Next, proportions of correct responses for those eight target sounds were analysed. To that end, ANOVAs were done like above (but with α set to .05), but now, in order to assess which noise types affected Dutch and English listeners differently, conditions were analysed in pairs as well as individually. The following is reported for each of the eight target sounds: First, to characterize the data, for each individual condition, the Dutch and English listeners' proportions of correct responses are compared. Second, and crucially, each noise condition is compared to the baseline condition and interactions between Native Language and Condition are reported; such interactions indicate that the noise condition affected the Dutch listeners more (in this study, it was never less) than the English listeners (relative to the baseline condition). If this was the case for more than one noise condition, those conditions are compared to one another, and interactions between Native Language and Condition are reported; such interactions

Table 1
Interactions between Native Language and Condition.

Consonant	$F(3,111)$	p
/p/	22.63	<.001
/t/	4.12	<.01
/k/	5.49	<.001
/g/	6.82	<.001
/m/	7.53	<.001
/n/	5.56	<.001
/ŋ/	9.57	<.001
/r/	4.92	<.01

show that the difference between the Dutch and the English listeners is larger in one noise condition than in the other. Finally, the conditions are ranked according to the proportion of correct responses, to assess the order of the difficulty of the noise types. (Dutch and English listeners' results are collapsed where possible, and reported for both Native Language groups separately only for those pairs of conditions that showed a significant interaction between Native Language and Condition.) Results of the ANOVAs are reported in Table 2; numbers in square brackets in the text refer to that table.

3.1. Stops

3.1.1. /p/

As the Appendix shows, for the target /p/, both for Dutch and English listeners, confusions were rather diverse, involving especially 'k', 'b', 'h', and 'g' responses. Large differences in Dutch and English listeners' confusions occurred in COMP and SSN conditions, where the Dutch listeners gave more 'h' (COMP and SSN) and 'ð' responses (SSN) than the English listeners did (with a difference >5% between the Dutch and English listeners). The Dutch listeners' common 'h' and 'ð' responses might be due to the difference in aspiration between Dutch and English /p/; whereas English /p/ is aspirated, Dutch /p/ is voiceless unaspirated (Gussenhoven, 1999), which might have made the Dutch listeners interpret the English aspirated stops as fricatives.

Fig. 2 shows that the Dutch listeners had a very high proportion correct in the baseline condition, but strikingly low proportions correct in the three noise conditions. There was no significant difference between Dutch and English listeners' responses in the baseline condition, where both groups performed at ceiling. In all of the noise conditions, however, Dutch listeners' proportion correct was significantly lower than that of the English listeners [Table 2; 1–4].

Indeed, comparing the baseline condition with each of the three noise conditions yielded significant interactions between Native Language and Condition [5–7]. The difference between the two Native Language groups was smaller for MODSSN than for COMP and SSN, shown by significant interactions between Native Language and Condition [9–10].

The proportion correct was higher in the baseline condition than in all other conditions both for Dutch and English listeners. Further, for the Dutch listeners, the proportion correct was higher in MODSSN than in COMP and than in SSN, but for the English listeners, MODSSN did not differ from COMP or from SSN. For both Native Language groups alike, there was no difference between COMP and SSN [11–16].

Thus, whereas Dutch listeners were as accurate as English listeners in the baseline condition, all noise conditions, and especially COMP and SSN, affected the Dutch listeners more than the English listeners. For the English listeners, the effect of noise was similar in the three noise conditions, but for the Dutch listeners, the effect of noise was larger in COMP and SSN, where the Dutch listeners gave relatively many 'h' and 'ð' responses, than in MODSSN.

3.1.2. /t/

For the target /t/, as the Appendix shows, Dutch and English listeners showed rather similar patterns of confusions. The most common confusion, both for Dutch and English listeners, involved interpretation of /t/ as 'θ'. Dutch listeners gave more 'θ' responses than the English listeners did (with a difference >5%) in MODSSN. Further, many 'd' responses were given.

As Fig. 2 shows, Dutch listeners had a significantly smaller proportion of correct responses than the English listeners in the baseline condition as well as in the three noise conditions [Table 2; 1–4].

All types of noise again affected the Dutch listeners' responses more than the English listeners' responses: comparing the baseline condition with each of the three noise conditions yielded significant interactions between Native Language and Condition [5–7].

The proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners, except for MODSSN which did not differ from the QUIET baseline condition for the English listeners. The proportion correct decreased from MODSSN to SSN, and from SSN to COMP, with significant differences between each of the conditions, for Dutch and English listeners alike [17,19].

In summary, noise affected the Dutch listeners more than the English listeners. The main confusions and the order of the difficulty of the noise types was the same for Dutch and English listeners, but the extent to which noise hindered recognition was larger for Dutch listeners than for English listeners in all noise conditions.

3.1.3. /k/

For the target /k/, Dutch and English listeners again showed rather similar patterns of confusions (Appendix). By far the most common confusion for both listener groups involved interpretation of /k/ as 'g'. Further, many 'h', 'v', and 'p' responses were given. In SSN condition, the Dutch listeners gave more 'g' responses than the English listeners did (with a difference >5% between the Dutch and English listeners). The steady state noise in that condition thus hindered the non-native listeners' evaluation of the presence of voicing more than the temporally modulated noise types.

Dutch listeners had a smaller proportion of correct responses than the English listeners in the baseline condition as well as in the noise conditions COMP and SSN (Fig. 2). There was no significant difference between Dutch and English listeners' proportion of correct responses for MODSSN [1–4].

Comparing the baseline condition with each of the noise conditions yielded significant interactions between Native Language and Condition only for QUIET versus SSN [5–7].

Table 2
Analyses of Variance for eight target sounds. (NS = not significant.)

		/p/	/t/	/k/	/g/	/m/	/n/	/ŋ/	/r/
Dutch versus English	1. quiet	NS	$F(1,38) = 5.77$ $p < .05$	$F(1,38) = 4.27$ $p < .05$	$F(1,38) = 4.47$ $p < .05$	NS	NS	NS	$F(1,38) = 34.15$ $p < .001$
	2. comp	$F(1,38) = 64.15$ $p < .001$	$F(1,38) = 13.64$ $p < .001$	$F(1,38) = 6.04$ $p < .05$	$F(1,38) = 24.26$ $p < .001$	$F(1,38) = 15.30$ $p < .001$	$F(1,38) = 6.24$ $p < .05$	$F(1,38) = 11.84$ $p < .001$	$F(1,38) = 13.89$ $p < .001$
	3. SSN	$F(1,38) = 33.25$ $p < .001$	$F(1,38) = 24.73$ $p < .001$	$F(1,38) = 13.80$ $p < .001$	NS	NS	$F(1,38) = 18.53$ $p < .001$	$F(1,38) = 22.40$ $p < .001$	$F(1,38) = 29.06$ $p < .001$
	4. modSSN	$F(1,38) = 26.19$ $p < .001$	$F(1,38) = 14.18$ $p < .001$	NS	NS	$F(1,38) = 16.93$ $p < .001$	$F(1,38) = 12.83$ $p < .001$	NS	$F(1,38) = 27.11$ $p < .001$
Native Language * Condition	5. quiet–comp	$F(1,37) = 66.73$ $p < .001$	$F(1,37) = 8.22$ $p < .01$	NS	$F(1,37) = 12.55$ $p < .001$	$F(1,37) = 12.12$ $p < .001$	NS	$F(1,37) = 9.72$ $p < .01$	$F(1,37) = 5.42$ $p < .05$
	6. quiet–SSN	$F(1,37) = 32.06$ $p < .001$	$F(1,37) = 9.24$ $p < .01$	$F(1,37) = 11.48$ $p < .01$	NS	NS	$F(1,37) = 13.85$ $p < .001$	$F(1,37) = 27.00$ $p < .001$	$F(1,37) = 16.53$ $p < .001$
	7. quiet–modSSN	$F(1,37) = 21.77$ $p < .001$	$F(1,37) = 12.53$ $p < .001$	NS	NS	$F(1,37) = 15.68$ $p < .001$	$F(1,37) = 7.04$ $p < .05$	NS	$F(1,37) = 11.99$ $p < .001$
	8. comp–SSN	NS	NS	NS	$F(1,37) = 12.93$ $p < .001$	$F(1,37) = 10.01$ $p < .01$	$F(1,37) = 4.86$ $p < .05$	NS	NS
	9. comp–modSSN	$F(1,37) = 17.90$ $p < .001$	NS	NS	$F(1,37) = 14.26$ $p < .001$	NS	NS	$F(1,37) = 4.41$ $p < .05$	NS
	10. SSN–modSSN	$F(1,37) = 6.93$ $p < .05$	NS	NS	NS	$F(1,37) = 10.15$ $p < .01$	$F(1,37) = 4.70$ $p < .05$	NS	NS
Condition (Dutch)	11. comp–SSN	NS			NS	$F(1,17) = 40.11$ $p < .001$	$F(1,17) = 30.95$ $p < .001$		
	12. comp–modSSN	$F(1,17) = 24.03$ $p < .001$			$F(1,17) = 10.60$ $p < .01$			$F(1,17) = 25.41$ $p < .001$	
	13. SSN–modSSN	$F(1,17) = 4.57$ $p < .05$				$F(1,17) = 45.47$ $p < .001$			
Condition (English)	14. comp–SSN	NS			$F(1,20) = 45.76$ $p < .001$	$F(1,20) = 135.71$ $p < .001$	$F(1,20) = 35.38$ $p < .001$		
	15. comp–modSSN	NS			NS			$F(1,20) = 8.84$ $p < .01$	
	16. SSN–modSSN	NS				$F(1,20) = 104.64$ $p < .001$			
Condition (Dutch and English)	17. comp–SSN		$F(1,37) = 7.65$ $p < .01$	$F(1,37) = 65.79$ $p < .001$				$F(1,37) = 37.55$ $p < .001$	NS
	18. comp–modSSN			$F(1,37) = 13.16$ $p < .001$		NS	$F(1,37) = 6.69$ $p < .05$		$F(1,37) = 17.34$ $p < .001$
	19. SSN–modSSN		$F(1,37) = 25.79$ $p < .001$		$F(1,37) = 29.70$ $p < .001$				$F(1,37) = 47.21$ $p < .001$

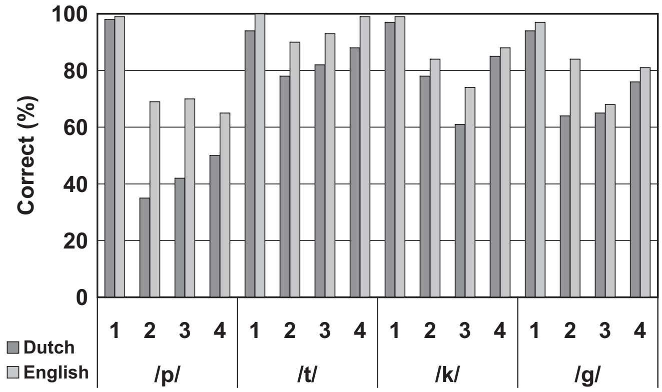

Fig. 2. Dutch and English listeners' percentage correctly identified /p, t, k, g/ in the conditions (1) QUIET; (2) competing talker (COMP); (3) speech-shaped noise (SSN); and (4) modulated speech-shaped noise (MODSSN).

Again, the proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners, and further decreased from MODSSN to COMP, and from COMP to SSN, with significant differences between each of the conditions [17,18].

To summarize, for the /k/ target, Dutch listeners had a smaller proportion of correct responses than English listeners in the baseline as well as two of the noise conditions, but noise only affected the Dutch listeners significantly more than the English listeners for SSN, where the Dutch listeners gave more 'g' responses than the English listeners did. In the other conditions, noise affected Dutch and English listeners to the same extent.

3.1.4. /g/

The most common confusions for the target /g/ were 'k' for the Dutch listeners and 'ŋ' and 'k' and to a lesser extent 'h', 'b', and 'd' for the English listeners. As the Appendix shows, in all noise conditions, the Dutch listeners gave more 'k' responses than the English listeners did (with a difference >5% between the Dutch and English listeners). The Dutch phoneme inventory contains a /k/, but /g/ only occurs in loanwords; therefore, Dutch listeners may have had difficulty interpreting the voicing of the /g/ target, especially when noise masked the perceptual cues to voicing. English listeners, on the other hand, responded 'ŋ' more than the Dutch listeners did (with a difference >5%) in SSN.

As Fig. 2 shows, Dutch listeners had a significantly smaller proportion of correct responses than the English listeners in the baseline condition and in the noise condition COMP. There was no significant difference between Dutch and English listeners' proportion of correct responses in SSN and MODSSN [1–4].

Comparing the baseline condition with each of the noise conditions showed that there was only an interaction between Native Language and Condition for QUIET versus COMP [5–7]. Thus, the noise in COMP hindered Dutch listeners more than English listeners.

Again, the proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners, and further decreased from MODSSN to SSN [19] for both Native Listener groups alike. In COMP, however, the English listeners' proportion correct was similar to that in MODSSN and higher than in SSN (QUIET > MODSSN–COMP > SSN) [14,15], but the Dutch listeners' proportion correct was similar to that in SSN and lower than in MODSSN (QUIET > MODSSN > COMP–SSN) [11,12]. (Separate analyses for Dutch and English listeners were warranted by significant interactions between Native Language and Condition when comparing COMP with SSN and COMP with MODSSN [8,9].)

Thus, for the /g/ target, the noise in COMP affected the Dutch listeners more than the English listeners, such that the ranking of the conditions according to the proportion correct was different for the Dutch and the English listeners, with COMP ending up in a shared second place for the English listeners and in a shared last place for the Dutch listeners. The noise in SSN and MODSSN affected Dutch and English listeners to the same extent. In all noise conditions, Dutch listeners gave more 'k' responses than English listeners did.

3.1.5. Summary

To summarize, for the target /p/, all types of noise hindered Dutch listeners' performance more than English listeners' performance, but the noise in COMP and SSN more than that in MODSSN. For the target /t/, all types of noise indiscriminately hindered Dutch listeners' performance more than English listeners' performance. For the target /k/, only the SSN noise, and for the target /g/, only the COMP noise hindered Dutch listeners' performance more than English listeners' performance. Thus, the effect of the different types of noise for the different stops was extremely varied.

3.2. Nasals

3.2.1. /m/

The most frequent errors for Dutch and English listeners alike involved 'n' responses (Appendix). Other common confusions were 'w', 'v', and to a lesser extent 'l'. Unlike in the other conditions, in SSN, 'b', 'l' and 'r' were the most frequent errors. A clear difference between the two listener groups was found in MODSSN, where the Dutch listeners gave more 'w' responses than the English listeners did (with a difference >5% between the Dutch and English listeners; Appendix).

As Fig. 3 shows, the Dutch listeners had a very high proportion correct in the baseline condition, but performed relatively poorly in all noise conditions. Note that both Dutch and English listeners had very high error rates in SSN. There was no difference between Dutch and English listeners' responses in the baseline condition, where both groups performed at ceiling, and in SSN, where both groups had a very low proportion correct. In the other

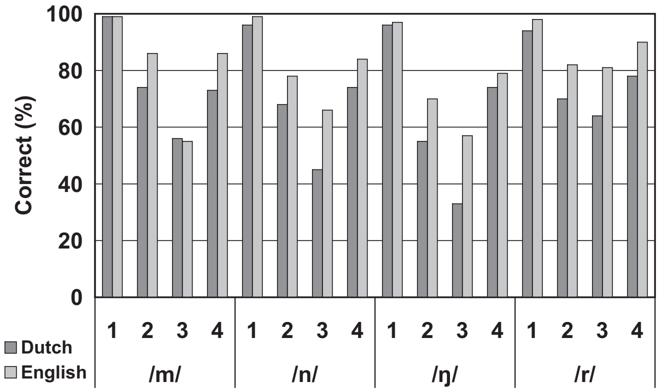

Fig. 3. Dutch and English listeners' percentage correctly identified /m, n, ŋ, r/ in the conditions (1) QUIET; (2) competing talker (COMP); (3) speech-shaped noise (SSN); and (4) modulated speech-shaped noise (MODSSN).

two noise conditions, however, Dutch listeners' proportion correct was significantly lower than that of the English listeners [1–4].

Indeed, comparing the baseline condition with each of the noise conditions yielded significant interactions between Native Language and Condition for QUIET versus COMP and QUIET versus MODSSN [5–7].

Again, the proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners. Also, both for Dutch and for English listeners, the proportion correct was lower in SSN than in COMP (with a significant interaction between Native Language and Condition) [11,14,8] and lower in SSN than in MODSSN (again with a significant interaction between Native Language and Condition) [13,16,10]. There was no difference between COMP versus MODSSN [18].

To sum up, for the /m/ target, the noise in SSN hindered Dutch and English listeners to the same extent, and very strongly. The noise in COMP and MODSSN (in the latter of which the Dutch listeners gave relatively many 'w' responses) affected their recognition of /m/ less than that in SSN, but hindered the Dutch listeners more than the English listeners.

3.2.2. /n/

For both groups of listeners, the most frequent errors were 'm' and 'l' responses, and to a lesser extent 'd', 'ŋ', and 'w' (Appendix). In SSN condition, Dutch listeners gave more 'd' and 'l' responses than English listeners did (with a difference >5%). The steady state noise and the resulting lack of clear glimpses of the signal in that condition might have obscured the perceptual cues to manner of articulation for the non-native listeners, such that they interpreted the alveolar nasal as an alveolar stop or approximant relatively often.

As Fig. 3 shows, the pattern of results for the target sound /n/ was rather similar to that for the /m/. Again, the Dutch listeners had a high proportion correct in the baseline condition, but performed relatively poorly in all noise conditions. However, whereas the Dutch listeners' performance was very poor in SSN again, the English listeners' performance in that condition was now not so bad. There was no significant difference between the Dutch and English listeners' proportion correct in the baseline condition, but in all of the noise conditions Dutch listeners' proportion correct was lower than that of the English listeners [1–4].

Comparing the baseline condition with the noise conditions yielded significant interactions between Native Language and Condition for QUIET versus SSN and for QUIET versus MODSSN [5–7]. The difference between the two Native Language groups was larger for SSN than for MODSSN, shown by a significant interaction between Native Language and Condition [10].

Again, the proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners. The proportion correct further decreased from MODSSN to COMP [18] and from COMP to SSN with significant differences between each of the conditions [11,14] (with a significant interaction between Native Language and COMP versus SSN [8]).

Thus, SSN and MODSSN noise hindered the Dutch listeners' recognition of the /n/ more than the English listeners' recognition. This differential effect of noise was larger in SSN, where Dutch listeners gave relatively many 'd' and 'l' responses, than in MODSSN. COMP noise hindered Dutch and English listeners to a similar extent.

3.2.3. /ŋ/

The most common confusion for both groups of listeners was 'g', followed by 'd' and 'n'. In SSN, Dutch listeners gave more 'g' responses than English listeners did (with a difference >5%; Appendix). Similar as for the /n/ target, this error might be due to the steady state noise in SSN condition obscuring the perceptual cues to manner of articulation for the non-native listeners. The generally large proportion of 'g' responses might be due to the fact that several speakers from the Sheffield region pronounced the target /ŋ/ as a combination of a nasal plus a stop.

For the target sound /ŋ/, noise affected both Dutch and English listeners relatively strongly, in particular in SSN again (Fig. 3). The Dutch listeners' proportion correct was similar to that of the English listeners in the baseline condition QUIET and in MODSSN, but significantly lower than that of the English listeners in COMP and SSN [1–4].

Indeed, comparing the baseline condition with each of the noise conditions showed that there was an interaction between Native Language and Condition for QUIET versus COMP and QUIET versus SSN [5–7]. Thus, noise affected Dutch listeners' responses more than English listeners' responses in COMP and SSN.

Again, the proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners, and it further decreased from MODSSN to COMP (with a significant interaction between Native Language and Condition [12,15,9]); and from COMP to SSN [17], with significant differences between all conditions.

Table 3
Noise conditions that hindered Dutch listeners more than English listeners (relative to the baseline condition QUIET), grouped by manner of articulation.

Consonant		Noise condition
Stops	/p/	COMP, SSN, MODSSN (COMP and SSN more than MODSSN)
	/t/	COMP, SSN, MODSSN
	/k/	SSN
	/g/	COMP
Nasals	/m/	COMP, MODSSN
	/n/	SSN, MODSSN (SSN more than MODSSN)
	/ŋ/	COMP, SSN
Approximant	/r/	COMP, SSN, MODSSN

Thus, the noise in COMP and SSN (in the latter of which Dutch listeners gave more 'g' responses than English listeners did) affected Dutch listeners' recognition of /ŋ/ more than English listeners' recognition, whereas the noise in MODSSN hindered Dutch and English listeners to a similar extent.

3.2.4. Summary

To summarize the results for the nasals, for each of the target sounds, two of the noise types hindered the Dutch listeners more than the English listeners; which two did, differed for each of the three target sounds. For /m/, COMP and MODSSN hindered Dutch listeners more than English listeners. For /n/, SSN and MODSSN hindered Dutch listeners more than English listeners, and this effect was larger in SSN than in MODSSN. For /ŋ/, COMP and SSN hindered Dutch listeners more than English listeners. Thus, although the general pattern of results was somewhat similar for all three nasals, the effect of the different types of noise was again highly variable.

3.3. Approximant /r/

As the Appendix shows, confusion patterns differed for the two listener groups. For the Dutch listeners, 'w' was the most frequent error in all conditions, followed by 'b', 'l', and 'v'. For the English listeners, 'w' and 'v' were the most frequent errors, followed by 'l' and 'm'. In all noise conditions, Dutch listeners gave more 'w' responses than English listeners did (with a difference >5%). Dutch /r/ can be pronounced as an alveolar approximant similar to the English target /r/, but is more commonly produced as an alveolar trill or tap, or a uvular trill or fricative; therefore, Dutch listeners might have been inclined to interpret the English approximant /r/ as the approximant /w/ rather than as /r/.

As Fig. 3 shows, the pattern of results for the target sound /r/ was largely similar to that for /t/. Like for the /t/, the Dutch listeners had a smaller proportion of correct responses than the English listeners in the baseline condition as well as in the three noise conditions [1–4].

All noise types affected the Dutch listeners' responses more than the English listeners' responses: comparing the baseline condition with each of the three noise conditions yielded significant interactions between Native Language and Condition for all noise conditions [5–7].

The proportion correct was higher in the baseline condition than in all other conditions for both Dutch and English listeners. Unlike for the /t/, the proportion correct now decreased from MODSSN to COMP and SSN, while the latter two conditions did not significantly differ from one another [17–19].

To summarize, for the /r/, all noise types affected the Dutch listeners more than the English listeners. The order of the difficulty of the noise types was the same for Dutch and English listeners, but the extent to which noise hindered recognition of the /r/ was larger for Dutch listeners than for English listeners in all noise conditions. Dutch listeners responded especially often with the approximant 'w'.

3.4. Overall result patterns

For the eight target sounds that showed differential effects of the four conditions for Dutch and English listeners, the pattern of results was highly variable. As we hypothesized, there was no single type of noise that affected the non-native listeners' perception more than the native listeners' perception for all phonemes. Further, also as expected, the effects of different types of noise on recognition of different phonemes by native versus non-native listeners varied strongly within classes of phonemes with the same manner of articulation. This already became clear from the discussion above, where the sounds were grouped according to manner of articulation. The results are summarized in Table 3. For each manner of articulation, each noise condition hindered the Dutch listeners more than the English listeners for some of the target sounds, but none of the noise conditions hindered the Dutch listeners more than the English listeners for all of the target sounds with the same manner of articulation.

Table 4 shows that regrouping the target sounds according to place of articulation does not result in a clearer picture. Again, for each place of articulation, each noise condition hindered the Dutch listeners more than the English listeners for some of the target sounds, but none of the noise conditions hindered the Dutch listeners more than the English listeners for all of the target sounds with the same place of articulation. There is one exception: MODSSN did not hinder Dutch and English listeners differentially for any of the velar targets. However, that hardly changes the picture that the differential effect for the two Native Language groups of the different types of noise is highly variable, also within place of articulation.

So far, the discussion of the results has focused on the differential effect of the noise conditions for the two Native Language groups. Stepping away from the differences

Table 4
Noise conditions that hindered Dutch listeners more than English listeners (relative to the baseline condition QUIET), grouped by place of articulation.

Consonant		Noise condition
Bilabials	/p/	COMP, SSN, MODSSN (COMP and SSN more than MODSSN)
	/m/	COMP, MODSSN
Alveolars	/t/	COMP, SSN, MODSSN
	/n/	SSN, MODSSN (SSN more than MODSSN)
	/r/	COMP, SSN, MODSSN
Velars	/k/	SSN
	/g/	COMP
	/ŋ/	COMP, SSN

between the Dutch and the English listeners, the effects of the noise conditions are slightly clearer. Table 5 summarizes the ranking of the noise conditions according to the proportion of correct responses. Dutch and English listeners' results are collapsed where possible. As the table makes clear, MODSSN always had the highest proportion of correct responses or a shared first position. This is in line with the results from seven other groups of L2 listeners taking part in the Interspeech 2008 Consonant Challenge (Cooke and Scharenborg, 2008; Garcia Lecumberri et al., 2008). (Note that the proportion of correct responses in QUIET, which is not indicated in Table 5, was always larger than that in MODSSN, except for the target sound /t/ for the English listeners, were the proportion correct was similar in QUIET and MODSSN.) Further, COMP has a higher proportion of correct responses than SSN more often than vice versa (i.e., only for /t/). Finally, for the nasals, SSN always had the lowest proportion of correct responses. Apart from these three observations, however, there are no clear regularities in the ranking of the noise categories.

4. General discussion

As predicted, the results showed that the effect of different types of noise on native and non-native listeners' recognition of different phonemes was highly variable. First, listening conditions affected native and non-native listeners differentially for eight consonants, but not for the other 16 consonants. Second, for those eight consonants for which listening conditions did affect native and non-native listeners differentially, the effects of different noise types on both listener groups were again highly variable. There was no single type of noise that affected native and non-native listeners' perception differently for all of those eight consonants.

Table 5
Ranking of the noise conditions according to the proportion of correct responses, collapsing Dutch and English listeners' results where they did not differ. (>: statistically significantly larger and –: not statistically different.)

Consonant		Noise condition
Stops	/p/	Dutch: MODSSN > (SSN–COMP); English: (SSN–COMP–MODSSN)
	/t/	MODSSN > SSN > COMP
	/k/	MODSSN > COMP > SSN
	/g/	Dutch: MODSSN > (SSN–COMP); English: (COMP–MODSSN) > SSN
Nasals	/m/	(COMP–MODSSN) > SSN
	/n/	MODSSN > COMP > SSN
	/ŋ/	MODSSN > COMP > SSN
Approximant	/r/	MODSSN > (COMP–SSN)

Further, also as predicted, within classes of phonemes with the same manner or place of articulation, the effects of different types of noise on native and non-native listeners' identification of different consonants varied strongly as well. Each type of noise affected the Dutch listeners' identification more than the English listeners' identification for some of the consonants, but none of the noise types did so for all of the consonants with the same manner or place of articulation (with the one exception that MODSSN did not affect the two groups differentially for any of the velars).

Taking the Dutch and English listeners' results together, some regularity in the difficulty of each noise type for the separate phonemes could be discovered. The highest percentage of correct responses for each phoneme was obtained in the QUIET condition, followed by MODSSN, that took a second or shared second position (in line with Cooke and Scharenborg (2008) and Garcia Lecumberri et al. (2008)). COMP had a higher percentage correct than SSN more often than vice versa. For the nasals, SSN always had the lowest proportion of correct responses.

Assessment of patterns of confusions, and especially the differences between Dutch and English listeners' patterns of confusions, for individual consonants provided some clues as to why particular types of noise might have affected Dutch and English listeners differentially. The steady state noise in the SSN condition seemed to obscure the voicing information for /k/ targets (leading to frequent 'g' responses for the Dutch listeners) as well as the manner information for nasal targets (leading to frequent 'd' and 'l' responses to /n/ targets, and 'g' responses to /ŋ/ targets for the Dutch listeners). Possibly, the temporal manipulations in the other two noise types allowed for glimpses of the signal that provided enough information for the non-native listeners to make more accurate decisions about voicing and manner, respectively. Further, the COMP and SSN conditions seemed to have obscured the temporal information in the /p/ signal, such that the Dutch listeners (who are familiar with voiceless unaspirated stops from their native language) interpreted the aspiration in the stop as a fricative instead. In the MODSSN condition, Dutch listeners relatively often misinterpreted /m/ as 'w'. Apparently those listeners relied on specific acoustic information to distinguish /m/ from /w/ that became especially obscured in MODSSN condition. Possibly, for all those sounds, native listeners were more flexible than non-native listeners in using different sources of information when a particular type of noise obscured perceptual information they would have normally used for the recognition of that sound.

For other target sounds, the L2 listeners' most common errors did not seem to depend much on the type of noise. For /g/, Dutch listeners gave relatively many 'k' responses in all noise conditions, most likely as a result of the absence of a /g/ in the Dutch phoneme inventory (except in loanwords); as Dutch listeners are not familiar with a velar stop voicing contrast in their native language, the voicing distinction may have been difficult for them to perceive, in any type of noise. Similarly, for /r/, Dutch listeners gave relatively many 'w' responses in all noise conditions. This was also argued to be due to differences between the Dutch and English phonemes; possibly, as /r/ is only sometimes produced as an approximant in Dutch, and more frequently as a trill, tap, or fricative, Dutch listeners interpreted the approximant /r/ as another approximant, i.e., /w/. Finally, for the target /t/, the most important confusions ('θ' and 'd') remained the same in all conditions for both groups of listeners. Thus, differences between Dutch and English listeners' confusions again suggest that there are no systematic patterns that are similar for all consonants with a particular manner or place of articulation, but rather that results can only be meaningfully interpreted on an individual basis for each consonant separately.

Interestingly, the eight phonemes that were affected differently by the different noise types for English and Dutch listeners do not seem to be a random set but include all the voiceless and one voiced stop, all the nasals, and one approximant. The phonemes that did not show such a differential effect were, thus, two out of three voiced stops (/b/ and /d/), three out of four approximants (/j/, /w/, and /l/), and all 11 fricatives and affricates.

The relatively low accuracy for some fricatives and affricates in QUIET for the native listeners may have left less room for interactions between listener groups and listening conditions. Indeed, fricatives and affricates, in particular dentals and labiodentals, are perceptually difficult for L1 English listeners in quiet and in noise (see, e.g., Maniwa et al., 2008; Wang and Bilger, 1973), and the present results form no exception (Appendix). A more important reason for the lack of an interaction between listener groups and listening conditions, however, seems that, whereas Dutch listeners made more errors than English listeners did, the differences between those listener groups were not notably larger (and sometimes even smaller) in noise than in QUIET (Appendix). Cutler et al. (2008) found similar results, and argue that this can be explained by the finding that English listeners rely more than Dutch listeners do on transitional cues, which do not survive well under noise, for fricative identification (Wagner et al., 2006).

Indeed, in general, the identification performance in QUIET does not seem to determine which phonemes did and which ones did not show a differential effect of noise for L1 and L2 listeners. For example /b/ and to a lesser extent /d/, /j/, /w/, and /l/ had a high percentage of correct responses in QUIET and a lower percentage correct in the noise conditions (Appendix), but showed no differences between the effects of the noise types for both groups of listeners.

Confusion patterns and differences in native and non-native listeners' confusions were also very diverse for the 16 consonants that did not show a differential effect of listening conditions for native and non-native listeners. For example, for the target /f/, Dutch listeners gave relatively many 'θ' responses compared to the English listeners but, importantly, they did so in all conditions. For the targets /θ, h, s, ʃ, z, ʒ, j/, on the other hand, there were no errors that were clearly more frequent for the Dutch listeners than for the English listeners, in any of the conditions. A more complicated pattern was found for the /v/ target: In each condition, there was one error that the Dutch listeners made more frequently (with a difference >5%) than the English listeners. This error differed per condition; in QUIET it was 'θ', in COMP 'b', in SSN and MODSSN 'ð'. Further, for that target, Dutch listeners also responded 'w' at least 5% more often than the English listeners did in all of those conditions. Overall, the error scores in the different conditions did not vary to the extent that there was an interaction between Native Language and Condition. Thus, there were highly variable patterns of results underlying the lack of an interaction between listener groups and noise types for those 16 target sounds.

Listeners have extensive experience with the speech of their native language and presumably have learned which perceptual cues survive in different noise conditions. Non-native listeners, on the other hand, use and weight perceptual cues differently (e.g., Broersma, 2005, 2008, 2010) and have far less exposure to the second language, and even more so to the second language in adverse listening conditions than native listeners do. Due to this lesser exposure to L2 in noisy conditions, non-native listeners might sometimes be less capable of adapting their use of perceptual cues to the listening conditions than native listeners, who might be able to use any available information to compensate for the loss of the information they might have preferentially used. Indeed, several studies have shown that non-native listeners suffer more from adverse listening conditions than native listeners do (e.g., at the phoneme level, Cutler et al. (2008) and Garcia Lecumberri and Cooke (2006)).

The present results show that the noise types employed might strongly affect the outcomes of any study of speech perception in noise. Different outcomes might in particular be expected for native versus non-native listeners, and those outcomes might further differ for individual phonemes, as both listener groups and phonemes are affected differently by different types of noise. In the present study, as predicted, we found highly variable effects of the three different noise types on the identification of the different target phonemes for the native and non-native listeners. Also as predicted, those effects of the different noise types on the different listener groups were also highly variably for phonemes with the same manner or place of articulation.

The diversity of perceptual cues that listeners have at their disposal, both under beneficial and adverse listening conditions, and the differences between native and non-native listeners' use and weighting of those cues preclude predictions about the effect of a particular type of noise on the perception of a particular phoneme by a particular listener group. Studies addressing the perception of speech in noise by native and non-native listeners have predominantly (exceptions being Cooke et al. (2008); Garcia Lecumberri and Cooke (2006); Garcia Lecumberri et al. (2008)) used a single type of noise. While we acknowledge the practical considerations that might keep one from including more than one type of noise in a perception experiment, the present results imply that the outcomes of any study into the effects of noise on native and non-native phoneme perception will crucially depend on the types of noise employed.

Acknowledgements

Each of the authors was supported by an individual (separate) Veni grant from the Netherlands Organisation for Scientific Research (NWO). We would like to thank Martin Cooke (University of the Basque Country) for kindly providing the English VCV data and two anonymous reviewers for helpful comments.

Appendix. Confusion matrices for 24 English consonants by Dutch and English listeners in four listening conditions

	Response																							
	b	tʃ	d	f	g	h	ʤ	k	l	m	n	ŋ	p	r	s	ʃ	t	θ	ð	v	w	y	z	ʒ
Dutch listeners, QUIET																								
b	94	1	·	·	·	·	·	·	·	·	·	·	5	·	·	·	·	·	·	·	·	·	·	·
tʃ	·	86	·	·	·	·	5	·	·	·	·	·	·	·	·	2	2	·	3	·	·	·	·	1
d	·	·	83	·	2	·	·	·	·	·	·	·	·	·	·	·	5	2	7	·	·	·	·	·
f	·	·	·	65	·	·	·	·	·	·	·	·	1	·	1	·	·	28	5	·	·	·	·	·
g	·	·	1	·	93	·	·	5	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·
h	·	·	·	·	1	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
ʤ	·	10	·	·	1	·	77	·	·	·	·	·	·	·	·	·	·	·	2	·	·	·	·	9
k	·	·	·	·	1	·	·	97	1	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·
l	·	·	·	·	·	·	·	1	100	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
m	·	·	·	·	·	·	·	·	·	99	1	·	·	·	·	·	·	·	·	·	·	·	·	·
n	·	·	·	·	·	·	·	·	1	3	96	·	·	·	·	·	·	·	·	·	·	·	·	·
ŋ	·	·	·	·	2	·	·	·	·	·	1	96	·	·	·	·	·	·	·	·	·	·	·	·
p	1	·	·	·	·	·	·	·	·	·	·	·	98	·	·	·	·	·	·	·	·	·	·	·
r	·	·	·	·	·	·	·	·	·	·	·	·	·	94	·	·	·	·	·	·	6	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	·	97	·	·	1	1	·	·	·	·	·
ʃ	·	1	·	·	·	·	·	·	·	·	·	·	·	·	1	98	·	·	·	·	·	·	·	·
t	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	94	3	1	·	·	·	·	1
θ	·	·	·	6	·	·	·	·	·	·	·	·	·	·	3	·	·	77	12	·	·	·	·	·
ð	·	·	1	·	·	·	·	·	1	·	·	·	·	·	·	·	·	27	59	10	·	·	2	·
v	·	·	·	6	·	1	·	·	·	·	·	·	·	1	·	·	·	8	11	61	10	·	·	·
w	·	·	·	·	·	2	·	·	·	·	·	·	·	1	·	·	·	·	·	·	94	2	·	·
y	·	·	·	·	·	2	3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	94	·	·
z	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	2	·	·	·	95	2
ʒ	·	1	·	·	1	·	9	·	·	·	·	·	·	·	·	2	·	·	·	·	·	1	2	85
English listeners, QUIET																								
b	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
tʃ	·	97	·	·	·	·	·	·	·	·	·	·	·	·	·	·	3	·	·	·	·	·	·	·
d	·	·	97	·	2	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·
f	·	·	·	88	·	·	·	·	·	·	·	·	·	·	·	·	·	8	3	1	·	·	·	·
g	·	·	·	·	97	1	·	1	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·
h	·	·	·	·	·	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	1	·	·
ʤ	·	2	1	·	2	·	87	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	7
k	·	·	·	·	1	·	·	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
l	·	·	·	·	·	·	·	·	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
m	·	·	·	·	·	·	·	·	·	99	·	·	·	·	·	·	·	·	·	·	·	·	·	·

(*continued on next page*)

Appendix (*continued*)

	Response																							
	b	tʃ	d	f	g	h	ʤ	k	l	m	n	ŋ	p	r	s	ʃ	t	θ	ð	v	w	y	z	ʒ
n	·	·	·	·	·	·	·	·	1	·	99	·	·	·	·	·	·	·	·	·	·	·	·	·
ŋ	·	·	·	·	3	·	·	·	·	·	·	97	·	·	·	·	·	·	·	·	·	·	·	·
p	·	·	·	·	·	·	·	·	·	·	·	·	99	·	·	·	·	·	·	·	·	·	·	·
r	·	·	·	·	·	·	·	·	·	·	·	·	·	98	·	·	·	·	·	1	1	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100	·	·	·	·	·	·	·	·	·
ʃ	·	1	·	·	·	·	·	·	·	·	·	·	·	·	1	97	·	·	·	·	·	·	·	1
t	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100	·	·	·	·	·	·	·
θ	·	·	·	6	·	·	·	·	·	·	·	·	·	·	4	·	·	79	10	·	·	·	·	·
ð	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	22	60	18	·	·	1	·
v	·	·	·	4	·	1	·	·	·	·	·	·	·	·	·	·	·	2	7	86	·	·	·	·
w	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	97	2	·	·
y	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100	·	·
z	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	1	96	2
ʒ	·	·	·	·	·	·	5	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	94
Dutch listeners, COMP																								
b	65	1	2	1	2	2	·	·	·	1	1	·	5	1	1	·	2	1	3	6	4	1	·	1
tʃ	·	69	1	·	2	·	9	2	1	·	1	·	·	·	·	3	5	·	1	·	·	·	·	4
d	2	·	67	·	4	1	·	2	·	·	3	1	·	2	3	·	4	2	6	·	1	·	·	1
f	1	·	·	47	·	2	·	·	1	·	·	·	1	·	8	1	·	35	2	1	·	·	·	·
g	4	1	2	1	64	4	·	13	1	·	1	2	1	·	·	·	1	·	1	·	1	1	·	·
h	·	·	1	2	2	80	·	2	2	1	2	·	1	1	1	·	1	1	·	1	1	1	1	·
ʤ	·	11	1	·	3	·	67	·	·	·	·	·	·	1	·	1	3	1	·	·	·	1	·	10
k	1	·	·	2	6	2	·	77	1	1	·	1	1	·	·	·	2	1	1	3	·	·	·	·
l	1	·	·	·	·	·	·	·	89	1	2	·	·	3	·	·	1	·	1	·	·	1	·	·
m	2	·	1	1	·	1	·	·	2	74	7	1	·	2	·	·	·	·	·	3	5	·	·	·
n	1	·	1	1	·	1	·	1	9	9	68	1	·	1	·	·	·	·	·	1	3	2	·	·
ŋ	1	·	8	1	17	·	·	2	1	2	6	55	·	1	·	·	1	1	2	·	·	2	·	·
p	11	1	·	3	8	10	·	12	1	2	2	1	35	2	1	·	4	2	1	1	1	·	·	·
r	6	·	·	·	1	1	·	2	3	1	1	·	·	70	·	·	·	·	1	3	9	·	·	·
s	1	·	·	3	·	·	·	1	·	·	1	·	·	·	75	·	1	7	2	·	·	·	6	·
ʃ	·	1	1	2	·	1	·	1	·	·	·	1	·	·	2	79	·	2	1	·	·	·	1	7
t	·	1	2	1	1	·	1	1	1	1	1	·	·	·	1	·	78	6	2	·	·	1	1	·
θ	·	·	·	21	·	·	·	1	·	·	·	·	·	·	4	·	2	60	8	1	1	·	1	·
ð	4	·	7	3	1	1	·	·	3	·	1	·	1	3	1	·	1	16	34	12	3	1	7	·
v	7	·	2	7	1	5	·	·	·	1	1	1	·	1	·	·	2	7	12	45	6	·	1	·
w	2	·	1	1	1	3	·	1	2	1	·	·	·	5	·	·	1	·	1	1	77	1	·	·
y	·	·	·	·	2	2	1	·	3	1	1	1	1	·	·	·	1	·	1	·	1	84	·	·
z	2	·	·	1	·	·	·	·	·	·	·	·	·	·	1	·	·	3	5	·	·	·	84	·
ʒ	·	·	1	·	1	1	10	·	·	·	·	·	·	1	·	4	·	·	·	1	·	2	3	75
English listeners, COMP																								
b	76	1	·	·	3	·	·	1	·	·	·	·	1	·	·	·	·	1	2	14	·	·	·	·
tʃ	·	83	·	·	·	·	7	2	·	·	·	·	1	·	·	2	3	·	·	·	·	·	·	1
d	1	·	80	·	4	·	·	·	·	·	3	·	·	·	1	·	2	2	4	·	·	·	1	·
f	·	·	·	70	·	·	·	·	·	·	·	·	1	·	3	·	·	21	2	2	·	·	·	·
g	3	·	1	·	84	1	·	3	·	1	·	3	2	·	·	·	1	·	·	·	·	·	·	·
h	·	·	·	1	1	85	·	2	·	·	·	1	2	2	·	·	·	·	1	3	1	1	·	·
ʤ	·	6	1	·	3	·	77	·	·	·	2	·	·	·	·	·	1	·	·	·	·	·	·	9
k	1	1	·	·	4	1	·	84	·	·	·	1	2	·	·	·	2	·	·	4	·	·	·	·
l	1	·	·	·	·	·	·	·	95	·	·	·	·	4	·	·	·	·	·	·	·	·	·	·
m	·	·	·	·	·	·	·	·	1	85	5	1	·	1	·	·	·	·	·	5	1	·	·	·
n	·	·	·	·	1	·	·	·	5	12	78	2	·	1	·	·	·	·	·	·	·	·	·	·

Appendix (*continued*)

	Response																							
	b	tʃ	d	f	g	h	ʤ	k	l	m	n	ŋ	p	r	s	ʃ	t	θ	ð	v	w	y	z	ʒ
ŋ	1	·	4	·	13	·	·	2	·	1	7	70	·	1	·	·	·	·	·	·	·	·	·	·
p	6	1	·	1	3	3	·	8	1	·	1	1	69	1	1	·	1	1	·	2	2	·	·	·
r	4	·	1	·	1	1	·	·	2	1	·	·	·	82	·	·	·	1	1	7	1	·	·	·
s	·	·	·	4	·	·	·	·	·	·	·	·	·	·	85	·	·	4	1	·	·	·	6	·
ʃ	·	1	·	1	·	·	1	·	·	·	·	·	·	·	·	88	·	2	1	·	·	1	·	5
t	·	·	1	1	1	·	·	1	·	·	1	·	·	·	1	·	90	3	1	·	·	·	·	·
θ	·	·	·	18	·	·	·	1	·	·	·	·	·	·	1	·	·	68	9	2	·	·	1	·
ð	1	·	2	1	1	·	·	·	2	·	·	·	1	2	·	·	1	16	42	29	·	1	1	1
v	1	·	·	3	·	1	·	·	·	·	·	·	·	1	·	·	·	3	10	77	·	·	·	·
w	1	·	·	·	·	·	·	·	·	1	·	·	·	6	·	·	·	·	1	2	87	1	·	·
y	·	·	·	·	1	·	·	·	2	1	1	1	·	1	·	·	·	·	·	1	1	91	·	·
z	1	·	·	·	·	·	·	·	·	·	·	·	·	1	1	·	·	1	4	5	·	·	85	2
ʒ	·	·	1	·	·	·	9	·	·	·	·	·	·	·	·	·	·	·	·	1	·	1	·	87
Dutch listeners, SSN																								
b	51	·	2	1	5	4	·	1	3	3	1	·	7	1	·	·	·	2	7	6	5	·	·	·
tʃ	·	67	·	·	1	·	22	·	·	·	·	·	·	·	·	5	1	·	·	·	·	·	·	3
d	1	·	59	·	5	·	1	2	3	·	1	·	·	·	·	·	13	3	10	·	·	1	·	·
f	·	·	·	44	·	1	·	·	·	·	·	·	1	·	·	·	1	26	11	12	·	·	2	·
g	1	·	8	·	65	3	·	14	·	·	·	3	·	1	·	·	1	1	1	·	·	·	·	·
h	5	·	·	1	3	69	·	·	3	1	1	·	·	5	·	·	·	1	1	1	2	4	·	1
ʤ	·	19	·	·	3	·	65	·	·	·	·	·	·	·	·	·	1	·	1	·	·	1	·	8
k	1	·	·	·	22	4	1	61	·	·	·	2	1	1	·	·	3	1	·	·	·	·	·	·
l	1	·	1	1	1	·	·	·	66	3	3	1	·	5	·	·	·	1	2	5	10	·	·	·
m	7	·	·	1	·	5	·	1	7	56	6	1	2	7	·	·	·	1	1	3	2	·	·	·
n	3	·	13	·	2	1	·	·	14	8	45	2	·	2	·	·	·	·	2	·	2	1	2	·
ŋ	2	·	3	·	27	3	1	6	7	2	5	33	·	·	·	·	1	·	·	·	1	7	·	1
p	2	·	1	6	2	8	·	11	1	2	2	·	42	1	·	·	2	6	8	3	1	·	·	·
r	3	·	1	·	1	2	·	1	4	2	1	·	1	65	·	·	·	·	·	2	18	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	1	92	·	·	1	1	·	·	·	5	·
ʃ	·	5	·	·	·	·	·	·	·	·	·	·	·	·	1	91	·	·	·	·	·	·	·	3
t	·	1	6	·	·	1	·	1	·	·	·	·	·	·	·	·	82	5	2	·	·	1	·	·
θ	2	·	3	8	1	1	·	1	1	·	·	·	1	·	3	·	2	52	17	1	·	·	5	·
ð	2	·	9	1	·	1	·	·	6	1	5	·	·	·	·	·	1	13	43	10	2	·	6	1
v	9	·	3	3	·	1	·	·	3	1	·	·	1	3	·	·	2	7	25	27	12	·	2	·
w	5	·	1	·	3	7	·	·	5	2	·	1	·	11	·	·	·	·	1	1	59	3	·	·
y	·	·	1	·	7	1	5	·	3	1	3	2	·	1	·	·	·	·	1	·	·	73	·	·
z	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	1	3	2	·	·	91	1
ʒ	·	1	·	·	1	2	15	·	·	·	·	·	·	·	·	3	·	·	·	·	·	1	3	72
English listeners, SSN																								
b	56	·	1	·	6	1	·	·	2	3	·	2	9	1	·	·	·	1	2	14	1	·	·	·
tʃ	·	80	·	·	·	·	13	·	·	·	·	·	·	·	·	1	2	·	·	·	·	·	·	4
d	1	·	71	·	7	·	·	1	3	·	2	4	·	·	·	·	8	1	3	·	·	·	·	·
f	·	·	·	65	·	·	·	·	·	·	·	·	·	·	·	·	·	11	6	17	·	·	·	·
g	1	·	7	·	68	·	·	7	1	·	·	10	1	1	·	·	·	·	·	·	1	·	·	·
h	5	·	1	1	2	67	·	·	1	·	1	·	3	3	·	·	·	1	·	7	1	6	·	1
ʤ	·	9	·	·	3	·	78	·	·	·	·	1	·	·	·	·	·	·	·	·	·	1	·	7
k	1	·	·	·	13	1	1	74	1	·	·	2	4	·	·	·	2	·	·	·	·	·	·	·
l	1	·	1	·	1	1	·	·	74	3	3	·	·	5	·	·	·	·	1	5	4	2	·	·
m	10	·	·	·	1	1	·	1	6	55	3	1	1	9	·	·	·	·	1	6	3	·	·	·
n	1	·	3	·	1	1	·	·	8	7	66	4	1	3	·	·	·	·	1	1	1	1	·	·

(*continued on next page*)

Appendix (*continued*)

	Response																							
	b	tʃ	d	f	g	h	dʒ	k	l	m	n	ŋ	p	r	s	ʃ	t	θ	ð	v	w	y	z	ʒ
ŋ	1	·	2	·	21	2	·	1	2	2	1	57	1	3	·	·	1	·	·	1	1	3	·	·
p	4	·	·	5	1	1	·	7	·	3	·	·	70	1	·	·	·	2	·	4	·	·	·	·
r	1	·	·	·	·	·	·	·	3	2	1	1	·	81	·	·	·	·	·	2	9	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	·	97	·	·	·	·	·	·	·	3	·
ʃ	·	1	·	·	·	·	1	·	·	·	·	·	·	·	1	94	·	·	·	·	·	·	·	3
t	·	·	2	·	·	·	·	2	·	·	·	·	·	·	·	·	93	1	·	·	·	·	·	·
θ	1	·	1	15	·	·	·	·	·	·	·	·	1	·	3	·	1	56	14	6	·	·	2	·
ð	3	·	7	·	1	·	·	·	6	·	3	·	1	·	·	·	·	9	41	24	·	·	4	·
v	9	·	3	1	·	·	·	·	·	·	·	1	1	3	·	·	·	4	7	67	1	·	1	·
w	5	·	·	·	2	·	·	·	2	3	·	·	·	10	·	·	·	·	·	2	71	3	1	·
y	·	·	·	·	6	2	1	·	1	·	2	7	·	·	·	·	·	·	·	·	·	79	·	·
z	1	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	2	4	·	·	92	1
ʒ	·	·	·	·	·	2	10	·	·	·	·	·	·	·	·	2	·	·	·	·	·	1	·	85
Dutch listeners, modSSN																								
b	51	1	4	3	1	1	·	·	1	2	2	·	5	2	·	·	·	1	10	7	8	·	·	·
tʃ	·	81	1	·	·	·	14	1	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	1
d	·	·	77	·	·	·	1	·	2	·	2	·	·	·	·	·	6	2	6	·	·	3	·	·
f	2	·	·	45	·	1	·	·	1	·	·	·	3	·	1	·	1	27	9	7	·	·	·	·
g	·	·	2	·	76	1	·	9	1	·	·	6	·	·	·	·	1	·	·	·	1	2	·	·
h	1	·	·	1	2	89	·	1	1	·	1	·	·	·	·	1	·	1	·	1	·	·	·	·
dʒ	·	7	5	·	8	·	69	1	·	·	·	1	·	·	·	·	2	·	·	·	·	1	·	6
k	·	·	·	·	6	5	·	85	·	·	·	1	1	·	·	·	1	·	·	·	·	·	·	·
l	1	·	2	·	1	3	·	·	77	1	2	·	·	1	·	·	·	1	5	1	3	1	1	·
m	4	·	·	·	·	·	·	·	1	73	8	1	1	·	·	·	·	·	·	3	8	·	·	·
n	1	·	6	·	·	1	·	·	6	8	74	1	·	·	·	·	·	1	1	1	·	·	·	·
ŋ	2	·	1	·	9	·	·	3	2	2	4	74	1	·	·	·	·	·	·	·	1	·	·	·
p	8	·	·	2	6	10	·	9	·	1	·	1	50	1	·	·	·	2	1	3	2	1	·	·
r	1	·	·	·	·	3	·	·	·	1	·	·	·	77	·	·	·	·	1	3	13	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	·	83	1	1	6	·	·	·	·	9	·
ʃ	·	4	·	·	·	1	1	·	·	·	·	·	·	·	1	81	·	·	·	·	·	·	1	11
t	·	·	3	·	·	·	2	·	·	·	·	·	·	·	·	·	88	6	1	·	·	·	·	·
θ	·	1	1	14	·	·	1	·	1	·	·	·	2	1	3	·	6	52	13	1	·	·	1	1
ð	5	·	14	1	2	2	·	·	2	·	·	1	1	·	·	·	·	11	42	13	2	·	2	·
v	12	·	1	5	1	·	·	1	·	·	·	·	5	·	·	·	·	6	19	39	9	·	1	·
w	5	·	1	·	·	2	1	·	1	1	·	·	·	1	·	·	·	·	·	1	83	1	·	·
y	·	·	1	·	3	1	2	1	2	·	1	2	·	·	·	·	·	·	·	·	1	86	·	·
z	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	2	2	·	·	·	94	1
ʒ	·	1	·	·	·	·	8	·	·	·	·	·	·	·	·	6	·	·	1	·	·	1	3	80
English listeners, modSSN																								
b	61	·	1	·	1	·	·	·	·	2	1	·	4	3	·	·	·	·	2	21	4	·	·	·
tʃ	·	90	·	·	·	·	7	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·
d	·	·	88	·	·	·	·	·	2	·	4	·	·	·	·	·	2	·	1	·	·	1	·	·
f	2	·	·	65	·	·	·	·	·	·	·	·	3	·	·	·	·	15	4	11	·	·	·	·
g	·	·	·	·	81	1	·	2	1	·	1	9	·	·	·	·	·	·	·	2	1	2	·	·
h	2	·	·	1	·	86	·	·	1	·	1	·	2	2	·	·	·	·	·	·	3	·	·	1
dʒ	·	4	2	·	6	2	74	1	·	·	·	1	·	·	·	·	5	·	1	·	·	·	·	3
k	·	·	·	·	5	4	·	88	·	·	·	1	1	·	·	·	·	·	·	·	·	·	·	·
l	1	·	1	·	1	1	·	·	81	1	1	·	1	3	·	·	·	·	2	5	1	1	1	·
m	2	·	·	·	·	·	·	·	1	86	3	·	3	·	·	·	·	·	·	3	1	·	·	·
n	·	·	4	·	·	·	·	·	4	5	84	·	·	·	·	·	1	·	1	·	·	·	·	·
ŋ	2	·	·	·	10	·	·	1	1	1	2	79	·	1	·	·	·	·	·	1	1	·	·	·
p	6	·	·	1	3	7	·	5	1	·	·	1	65	·	·	·	1	·	1	6	1	1	·	·

Appendix (*continued*)

	Response																							
	b	tʃ	d	f	g	h	dʒ	k	l	m	n	ŋ	p	r	s	ʃ	t	θ	ð	v	w	y	z	ʒ
r	1	·	·	·	·	·	·	·	·	2	·	·	·	90	·	·	·	·	·	4	2	·	·	·
s	·	·	·	·	·	·	·	·	·	·	·	·	·	·	90	·	1	4	·	·	·	·	5	·
ʃ	·	1	·	·	·	1	·	·	·	·	·	·	·	·	·	86	·	·	·	·	·	·	·	11
t	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	99	·	·	·	·	·	·	·
θ	1	·	·	19	·	·	·	·	·	·	·	·	2	·	3	·	3	54	14	1	·	·	·	·
ð	4	·	9	1	4	1	·	·	1	·	·	·	1	1	·	·	·	12	35	31	·	·	1	·
v	7	·	·	7	·	·	·	·	·	·	·	·	3	·	·	·	·	3	8	70	1	·	·	·
w	4	·	·	·	·	1	·	·	2	1	·	·	·	2	·	·	·	·	·	3	86	1	·	·
y	·	·	·	·	3	·	·	·	1	·	·	1	·	·	·	·	·	·	·	·	1	93	·	·
z	·	·	·	·	·	·	·	·	1	·	·	·	·	·	1	·	·	·	4	1	·	·	92	·
ʒ	·	·	·	·	·	·	6	·	1	·	·	·	·	·	·	1	·	·	·	·	·	·	·	91

References

Bohn, O.-S., Munro, M.J. (Eds.), 2007. Language Experience in Second Language Speech Learning: In Honor of James Emil Flege. John Benjamins, Amsterdam.

Bradlow, A.R., Alexander, J.A., 2007. Semantic and phonetic enhancement for speech-in-noise recognition by native and non-native listeners. Journal of the Acoustical Society of America 121, 2339–2349.

Bradlow, A.R., Bent, T., 2002. The clear speech effect for non-native listeners. Journal of the Acoustical Society of America 112, 272–284.

Broersma, M., 2005. Perception of familiar contrasts in unfamiliar positions. Journal of the Acoustical Society of America 117, 3890–3901.

Broersma, M., 2008. Flexible cue use in nonnative phonetic categorization. Journal of the Acoustical Society of America 124, 712–715.

Broersma, M., 2010. Perception of final fricative voicing: native and nonnative listeners' use of vowel duration. Journal of the Acoustical Society of America 127, 1636–1644.

Cooke, M., 2006. A glimpsing model of speech perception in noise. Journal of the Acoustical Society of America 119, 1562–1573.

Cooke, M., Garcia Lecumberri, M.L., Barker, J., 2008. The foreign language cocktail party problem: energetic and informational masking effects in non-native speech perception. Journal of the Acoustical Society of America 123, 414–427.

Cooke, M., Scharenborg, O., 2008. The Interspeech 2008 consonant challenge. In: Proceedings of the Interspeech 2008, Brisbane, Australia.

Cutler, A., Garcia Lecumberri, M.L., Cooke, M., 2008. Consonant identification in noise by native and non-native listeners: effects of local context. Journal of the Acoustical Society of America 124, 1264–1268.

Cutler, A., Smits, R., Cooper, N., 2005. Vowel perception: effects of non-native language vs. non-native dialect. Speech Communication 47, 32–42.

Cutler, A., Weber, A., Smits, R., Cooper, N., 2004. Patterns of English phoneme confusion by native and non-native listeners. Journal of the Acoustical Society of America 116, 3668–3678.

Hazan, V., Simpson, A., 2000. The effect of cue-enhancement on consonant intelligibility in noise: speaker and listener effects. Language and Speech 43, 273–294.

Felty, R.A., Buchwald, A., Pisoni, D.B., 2009. Adaptation to frozen babble in spoken word recognition. Journal of the Acoustical Society of America 125, 93–97.

Festen, J.M., Plomp, R., 1990. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. Journal of the Acoustical Society of America 88, 1725–1736.

Garcia Lecumberri, M.L., Cooke, M., 2006. Effect of masker type on native and non-native consonant perception in noise. Journal of the Acoustical Society of America 119, 2445–2454.

Garcia Lecumberri, M.L., Cooke, M., Cutugno, F., Giurgiu, M., Meyer, B.T., Scharenborg, O., Van Dommelen, W., Volin, J., 2008. The non-native consonant challenge for European languages. In: Proceedings of the Interspeech 2008, Brisbane, Australia.

Golestani, N., Rosen, S., Scott, S.K., 2009. Native-language benefit for understanding speech-in-noise: the contribution of semantics. Bilingualism: Language and Cognition 12, 385–392.

Gussenhoven, C., 1999. Dutch. In: Handbook of the International Phonetic Association. CUP, Cambridge, pp. 74–77.

Lu, Y., 2010. Production and Perceptual Analysis of Speech Produced in Noise. Doctoral Dissertation, University of Sheffield, UK.

Maniwa, K., Jongman, A., Wade, T., 2008. Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners. Journal of the Acoustical Society of America 123, 1114–1125.

Mayo, L.H., Florentine, M., Buus, S., 1997. Age of second-language acquisition and perception of speech in noise. Journal of Speech and Hearing Research 40, 686–693.

Nábělek, A.K., Donahue, A.M., 1984. Perception of consonants in reverberation by native and non-native listeners. Journal of the Acoustical Society of America 75, 632–634.

Phatak, S.A., Lovitt, A., Allen, J.B., 2008. Consonant confusions in white noise. Journal of the Acoustical Society of America 124, 1220–1233.

Simpson, S.A., Cooke, M., 2005. Consonant identification in N-talker babble is a nonmonotonic function of N. Journal of the Acoustical Society of America 118, 2775–2778.

Strange, W. (Ed.), 1995. Speech Perception and Linguistic Experience: Issues in Cross-language Research. York Press, Baltimore.

Van Engen, K.J., Bradlow, A.R., 2007. Sentence recognition in native- and foreign-language multi-talker background noise. Journal of the Acoustical Society of America 121, 519–526.

Van Wijngaarden, S.J., 2001. Intelligibility of native and non-native Dutch speech. Speech Communication 35, 103–113.

Van Wijngaarden, S.J., Steeneken, H.J.M., Houtgast, T., 2002. Quantifying the intelligibility of speech in noise for non-native listeners. Journal of the Acoustical Society of America 111, 1906–1916.

Wagner, A., Ernestus, M., Cutler, A., 2006. Formant transitions in fricative identification: the role of native fricative inventory. Journal of the Acoustical Society of America 120, 2267–2277.

Wang, M.D., Bilger, R.C., 1973. Consonant confusions in noise: a study of perceptual features. Journal of the Acoustical Society of America 54, 1248–1266.