English as a lingua franca: mutual intelligibility of Chinese, Dutch and American speakers of English
Wang, H.

Chapter six

Intelligibility of vowels[1]

6.1 Introduction

In this chapter we will present the results for intelligibility of vowels of the groups of listeners. These results are from 36 Dutch listeners in Leiden, 36 Chinese listeners in Changchun and 36 Americans in Los Angeles listening to the six selected, optimally representative speakers. We would like to find out, among other things, how the classification errors we found in the previous chapter by applying a Linear Discriminant Analysis to the acoustical properties of the vowel tokens are different from the actual human perception results. The correctness and classification matrixes presented in Chapter five will be the reference data for the present chapter.

As we predicted in Chapter three, the differences in the sound systems in the three native languages will lead to a foreign accent for the Chinese and Dutch speakers of English which consist in deviations from the generally accepted pronunciation norm of English that are reminiscent of the native language of the learners, either Chinese or Dutch. The established structures of the Chinese/Dutch representation must be confronted with speech data from the target language, English. As a source of variability in speech, can Dutch/Chinese-accented English be detrimental to speech perception? When listeners are unable to recognize phonetic segments, words or larger units, will the result be partial or complete misidentification? If so, how well are English vowels identified by native American, Chinese and Dutch listeners? What is their confusion structure? Can we relate the confusions to specific interference patterns that reflect structural properties of the mother tongue of the non-native speaker and/or listener (Chapter three)? Will the confusion structure be different from the automatic classification results in Chapter five? This chapter may provide answers to these questions.

---

[1] Summaries of Chapters six to nine have appeared in English as H. Wang and V. J. Van Heuven (2005) Mutual intelligibility of American, Chinese and Dutch-accented speakers of English. *Proceedings of Interspeech 2005*, Lisbon: ISCA, 2225–2228 and in Dutch as V.J. Van Heuven and H. Wang (2006) Onderlinge verstaanbaarheid van Chinese, Nederlandse en Amerikaanse sprekers van het Engels. In T. Koole, J. Nortier, B. Tahitu (eds.) *Artikelen van de vijfde sociolinguistische conferentie*, Delft: Eburon, 257–266.

6.2 Results

Our research focuses on English as the target language and Dutch and Chinese as the source languages. We compare the intelligibility of Chinese-accented English, Dutch-accented English and native American English in an attempt to clarify how well these people understand each other and themselves when they are speaking English with their respective accents.

We hypothesize that foreign-accented English must be more difficult for English listeners as the source language deviates more from English, but native listeners still have strategies which non-native listeners lack for coping with all sorts of non-optimal speech, including foreign accents. Generally, then, native English listeners will be at an advantage over foreigners when listening to non-native English. There may just be one exception to this rule: non-native listeners may understand their own accented English better than native English listeners do. Since the foreign listeners are acquainted with the interfering native language, they may be sensitive to cues in the source language that native English listeners fail to pick up. This is what was called the interlanguage benefit by Bent and Bradlow (2003). Provisional data showing that this effect does apply to the present problem were presented earlier by Wang and Van Heuven (2003, 2004).

6.2.1 Overall results

The overall results for vowel intelligibility are presented in Figure 1, broken down by nationality of listeners and broken down further by nationality of speakers.

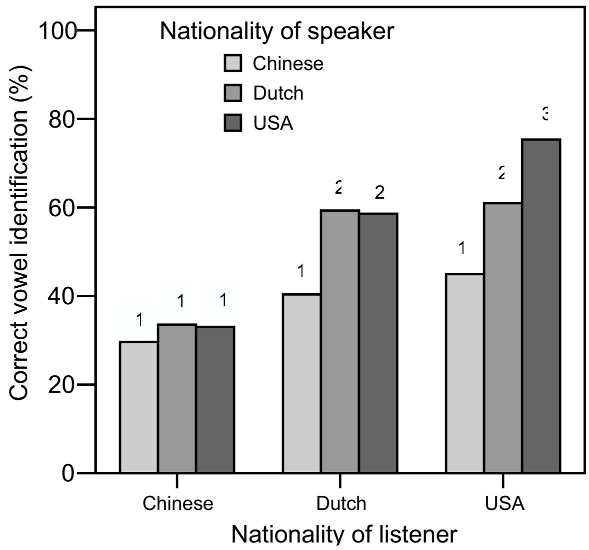

Figure 6.1. Percent correctly identified vowels for Chinese, Dutch and American listeners broken down by accent of speaker. Numbers above the bars indicate subgroup membership as determined by the Scheffé procedure (see text). Means and standard deviations are numerically specified in Appendix A6.1.

The data were submitted to an Analysis of Variance (ANOVA) run on the mean percent correct scores for each listener with nationality of speaker and nationality of listener as fixed factors.

Across speaker groups, the Chinese listeners have the lowest vowel identification scores (29–34% correct, mean = 32%). Dutch listeners perform at an intermediate level (40–59% correct, mean = 53%), and the American listeners are the best (45–75% correct, mean = 60%).[2] The effect of listener group was significant, $F(2, 315) = 204.9$ ($p < .001$). A post-hoc test (Scheffé procedure with $\alpha = .05$) indicates that all three speaker nationalities were different from each other.

Across listener groups Chinese speakers obtained the lowest vowel identification scores (38%). The Dutch and American speakers' vowels were identified with 51% and 56% correct, respectively. The effect of speaker nationality is significant, $F(2, 315) = 77.7$ ($p < .001$). The Chinese speakers are significantly poorer that the other two nationalities, which do not differ from each other. We may note that the effect of listener nationality is almost three times larger than the effect of speaker nationality.

Crucially, the interaction between listener and speaker nationality also reaches significance, $F(4, 315) = 17.0$ ($p < .001$). This implies that the mean scores obtained for specific combinations of listener and speaker nationalities cannot be computed by simply adding or subtracting a term for each factor level. Specifically, it can be shown that listeners obtain higher vowel identification scores when responding to materials produced by speakers of their own native language. This can be shown by computing the expected scores for each of the nine possible combinations of listener and speaker nationality and then comparing this expected score with the observed score. Mean percent correct vowel identification equals 50. When the listeners are Chinese, Dutch and American, the expected score is –18, +3 and +10 below or above the mean; for the three speaker nationalities the mean should be corrected with –12, +1 and +6, respectively. The expected and observed scores are listed in table 6.1 together with the difference between the two (delta or prediction error).

Generally, the observed scores are correctly predicted or even overestimated by the linear addition of the two main effects. Only in three combinations of factor levels is the observed score substantially better than the prediction. These are precisely the conditions in which the listeners are confronted with vowel tokens spoken by their fellow countrymen (shaded rows in Table 6.1). This native or interlanguage benefit is 5 to 10 percentage points better than the expected score.

---

[2] This result is different from a pilot test which showed that Dutch listeners performed best. In the pilot (Wang and Van Heuven, 2003) the Chinese listeners had the lowest vowel identification scores (50 to 60% correct). Dutch listeners performed best (65 to 80% correct), and the American listeners were intermediate (60 to 70% correct). Chinese-spoken vowels were most difficult for both Dutch and American listeners but not for Chinese speakers. Generally, listeners obtained the highest identification scores when responding to materials produced by speakers of their own native languages. This small advantage of Dutch-accented English for Chinese listeners may have been caused by the circumstance that our Chinese listeners had lived in the Netherlands for some six months, and therefore had had more exposure to Dutch-accented English than to L1 American English.

Table 6.1. Expected vowel identification scores (% correct) on the basis of grand mean = 50% and main effects for Listener and Speaker nationality for each combination of factor levels. Observed and error scores are indicated. Bolded delta's represent the interlanguage or native language benefit.

	Listener nationality		Speaker nationality		Expected	Observed	Δ
1.	Chinese	–18	Chinese	–12	20	30	**+10**
2.	Chinese	–18	Dutch	+1	33	34	+1
3.	Chinese	–18	American	+6	38	34	–4
4.	Dutch	+3	Chinese	–12	41	40	–1
5.	Dutch	+3	Dutch	+1	54	59	**+5**
6.	Dutch	+3	American	+6	59	59	0
7.	American	+10	Chinese	–12	48	45	–3
8.	American	+10	Dutch	+1	61	61	0
9.	American	+10	American	+6	66	75	**+9**

To conclude this part of the data presentation, we ran separate one-way ANOVAs in order to determine to what extent the three speaker nationalities differed within each of the three listener groups. Within the Chinese listeners, speaker was not a significant effect, $F(2, 105) = 1.4$ (ins.). In the Dutch listener group the Chinese speakers were more difficult to understand than either the Dutch or the American speakers, $F(2, 105) = 40.7$ ($p < .001$), who did not differ from each other (Scheffé). For American listeners the Chinese speakers were more difficult to understand than the Dutch speakers, who in turn were more difficult to understand than fellow Americans, $F(2, 105) = 69.9$ ($p < .001$), where all three speaker groups differed significantly. Significant differences between speaker groups have been indicated in figure 6.1 with superscript numbers over each bar.

6.2.2 Overview of the sound system

The experimental literature on foreign-language interference typically addresses one specific contrast at a time. For instance, there is a vast literature on the acquisition of the English /r ~ l/ contrast by speakers of Asian backgrounds (where the contrast is no part of the phonology). In the area of vowels much effort has been made to study the details of the acquisition of 'new' contrasts such as English /e ~ æ/ by Germans, or the English /iː ~ ɪ/ contrast by Hispanic learners (Flege, 1995). However, experimental studies targeting the confusion structure in an entire vowel inventory in a cross-linguistic setting are few and far between.

Before we present and analyze the confusion structure in the Chinese, Dutch and American tokens of English vowels, let us briefly recapitulate, in Table 6.2, the comparison of the three vowel systems as provided in Chapter three, in an attempt to derive specific predictions as to where confusions may arise in Chinese and Dutch-accented varieties of English.

Table 6.2. Summary of vowel systems of Mandarin Chinese, Dutch and English.

Chinese (source language)						
	Front		Central		Back	
High	i		y		u	
Mid	e		ɚ		o	
Low			ɑ			
Dutch (source language)						
	Front		Central		Back	
	Tense	Lax	Tense	Lax	Tense	Lax
High	i		y		u	
Hi-mid	e:	ɪ	ɸ:	ɵ	o:	
Lo-mid	ɛi	ɛ	œy			ɔ
Low			a:		ɑu	ɑ
English (target language)						
	Front		Central		Back	
	Tense	Lax	Tense	Lax	Tense	Lax
High	i:, ɪə^r				u:, ʊə^r	
Hi-mid	e:, ɛə^r	ɪ			o:, (ɔə^r)	ʊ
Lo-mid		ɛ	ɚ:^{r}	ʌ	ɔ:, ɔi	ɔ
Low	ai	æ			ɑ:, ɑu	

6.2.3 Correct vowel identification

In order to obtain an overview of which vowels are more difficult than others, for each combination of speaker and listener nationality, we present percentages of vowels correctly identified by Chinese, Dutch and American listeners in separate panels for Figure 6.2. In each panel the results have been broken down by nationality of the speakers. We have simplified the presentation rather drastically by listing the results only for those vowels that can be considered monophthongs. All full diphthongs, vowels followed by /r/ and the strongly confused /ɔ:/ have been omitted from the figures (for details on this data reduction, see below). In the panels the ten monophthongs have been ordered in descending order of correct identification when the speakers are American. Generally we would expect the results for the non-native speakers, i.e. by Chinese and Dutch speakers, to fall below the percentage correct of the American speakers. Only in exceptional cases do we expect the non-native vowels to be identified better than the native vowels.

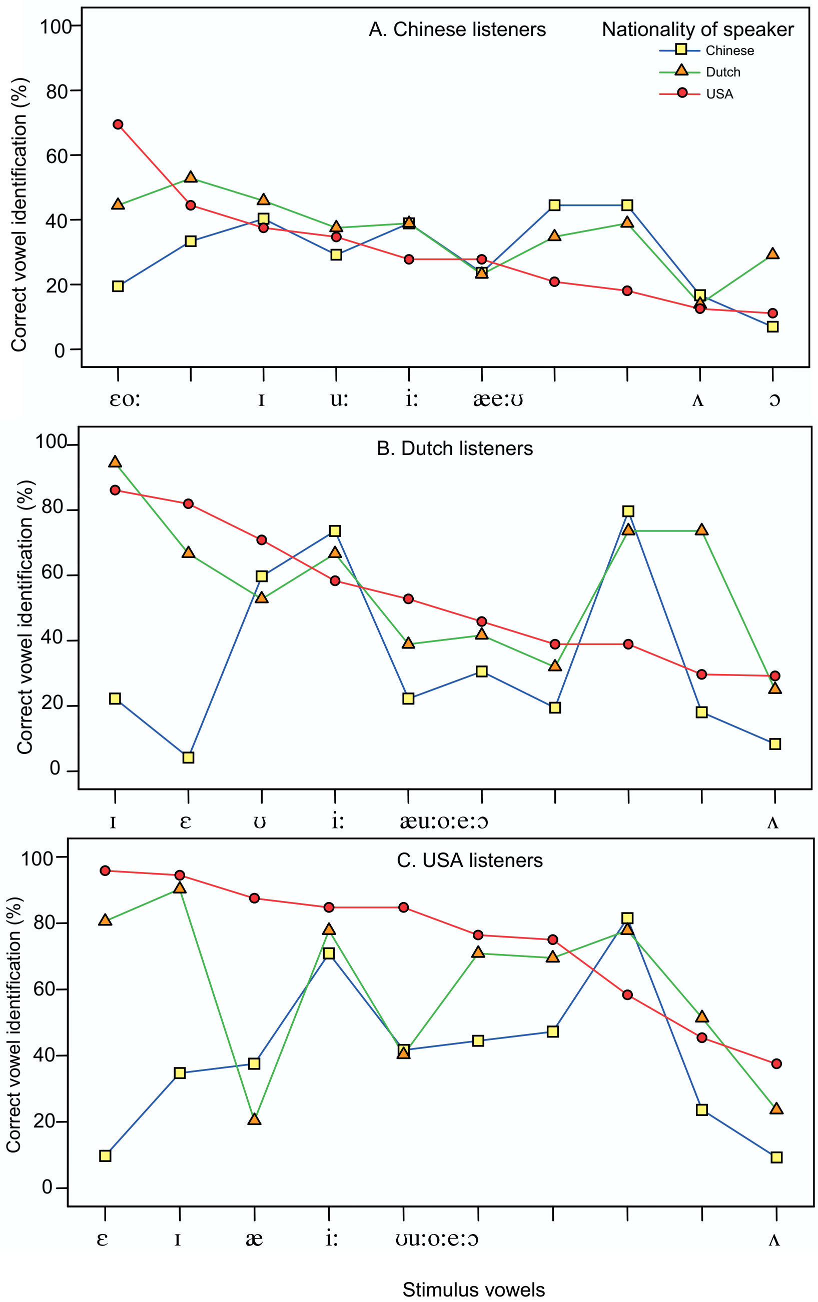

Figure 6.2 Correct vowel identification (%) for ten phonological English monophthongs produced by Chinese, Dutch and American speakers. Panels A, B and C present the results for Chinese, Dutch and American listeners, respectively.

Obviously, some vowels are more difficult than others. Moreover, there is hardly any correlation in the percentages correct identification of the vowels spoken by Chinese, Dutch and American speakers. That is to say, when a vowel spoken by an American speaker is easy to identify, it does not mean that the same vowel is also easy when it is produced by a Chinese or a Dutch speaker. Table 6.3 lists the correlation coefficients. In only one situation is the correlation significant, i.e., the identification of American and Dutch-spoken vowels is correlated for Chinese listeners.

Table 6.3. Pearson's correlation coefficients for identification of vowels produced by Chinese, Dutch and American speakers broken down by nationality of the listeners.

Listener nationality	Speaker nationalities		
	CN ~ NL	CN ~ US	NL ~ US
CN	r = 0.495	r = 0.032	r = 0.654*
NL	r = 0.277	r = –0.002	r = 0.513
US	r = 0.373	r = 0.118	r = 0.407

*: $p < 0.05$

Figure 6.2a shows that, relative to the American native speakers, the vowels /e:/ and /ʌ/ are easier to identify for **Chinese listeners** when the speakers are either Dutch or Chinese. The vowel /ɛ/, however, is clearly more difficult relative to the American pronunciation when it is spoken by a Dutch speaker, and even more so when the speaker is Chinese. In order to understand why the vowel /ɛ/ is a special source of difficulty we will have to examine its confusion structure, which we will defer to the next section.

When the **listeners are Dutch** (Figure 6.2b), we may observe that generally the American-spoken vowels are easier to identify correctly than the non-native tokens. Remarkably, the non-native tokens of /e:/, and for Dutch speakers also /o:/, are easier to identify than their American-accented counterparts. On the other hand, several non-native vowels are clearly more difficult than the native vowels. When the speakers are Chinese there are great difficulties with /ɪ/ and /ɛ/ as well as less severe problems with many of the other vowels: /ɛ, u:, o:, ɔ, ʌ/. When the speakers are Dutch themselves, there seem to be no specific difficulties.

American listeners (Figure 6.2c) are much better off listening to vowels spoken by fellow Americans than to foreign-accented vowels. Still, the non-native tokens of /e:/ are identified better by American listeners than their own tokens of /e:/. Also, Dutch-accented /o:/ is better than the American counterpart. Non-native /æ/ and /ʌ/ stand out as especially difficult vowels, as does lax /ɛ/ pronounced by Chinese speakers. Again, in order to better understand why certain vowels present special problems we need to know more about the specific confusion patterns in the vowel identifications, which is the topic of the next subsection.

6.2.4 Vowel confusion structures

6.2.4.1 Confusion matrices

Confusions in an identification task are customarily presented in a confusion matrix. Here the rows list the intended (stimulus) categories, while the columns represent the perceived categories. Correctly perceived stimuli appear in the cells along the main diagonal from top-left to bottom-right; errors are in the off-diagonal cells. As an illustration, Table 6.4 presents the confusion matrix for the 19 English vowels as produced by Chinese speakers and identified by American listeners.

In order to improve legibility, the cells in Table 6.4 have been shaded such that cells with larger numbers of observations in them have darker grey shades. Generally, the darkest cells find themselves along the main diagonal, indicating that very often the vowels as intended by the Chinese speakers were correctly identified by the American listeners. The values in the cells are percent correctly identified vowels relative to the row marginals, i.e. percentages should add up to 100 in each row of the matrix. Grey cells that are off the main diagonal represent substantial amounts of error or confusion. There are several concentrations of confusion in the table. For instance, tense /iː/ and lax /ɪ/ are strongly confused: /iː/ is misperceived by the American listeners (i.e. mispronounced by the Chinese speakers) in 25% whereas /ɪ/ is mistaken for /iː/ in more than half of the cases (53%). A similar confusion pattern can be seen further down the diagonal where there is a similar confusion pattern for tense /uː/ and lax /ʊ/, indicating that possibly all tense~lax contrasts are a source of error in the communication between Chinese speakers of English and American listeners. Interestingly, it also seems a recurring pattern that the tense vowel is dominant in the confusion pattern: the lax counterpart is confused more often with the tense vowel than vice versa. Such asymmetrical confusion patterns are often found in vowel perception studies.

There are several more concentrations of confusion in the table, which we will not analyze here. The point at issue here is that we need some method to extract and highlight the confusion structure in tables such as 6.4. Several methods have been proposed and applied in the literature. We will briefly review these, and then decide not to use any of these. Instead, we will propose a more practical analytical tool for our purpose, and then use this tool when analyzing the confusion structures in each of the nine combinations of speaker and listener nationalities. The full set of confusion matrices has been included in Appendix A6.2. In the main text of this chapter we will present a selection of the most obvious confusions in confusion graphs.

Table 6.4. Sample confusion matrix for 19 American English vowels produced by Chinese speakers and identified by American listeners.

		Response vowel (American)																		
		i:	ɪ	e:	ɛ	ɑ:	æ	u:	ʊ	ɔ:	ɔ	o	ʌ	ʌr	ai	ɔi	ɑu	iə	uə	ɛə
Stimulus vowel (Chinese)	i:	**71**	25	1		1					1									
	ɪ	53	**36**	1					6		3								1	
	e:	1		**85**	4		1				1				2	2	4		1	
	ɛ		3	8	**10**	3	7						1	1	63	1	1			
	ɑ:			1		**52**	10	1		7	13		3	9					3	
	æ			6	4		**39**	3		1					46			1		
	u:		1	1				**45**	39			3	1				3	1	3	
	ʊ				1	1	4	46	**42**								4			
	ɔ:		1	3		10	1	3		**18**		1		4	25	3	29			
	ɔ			2	2	6	2			18	**26**					2	42	2		
	o	3			1	1	3	3	29	1	4	**49**			1		3			
	ʌ		1	1	1	6	69	1		1	8		**9**		1	2				
	ʌr		1			14			4			1	1	**68**	1	1		1	6	
	ai		1	1		4	20			4	3		1		**61**		1	1		
	ɔi			1	3	1						4			37	**51**			3	
	ɑu	1			1		1	1	1	21	19	1		1	1	3	**46**			
	iə	1	1											21	1			**70**	1	3
	uə			1		3	1		1					7		6	1		**79**	
	ɛə					42	4			1				25	1	4		1		**20**

6.2.4.2 Extracting confusion patterns

Hierarchical cluster schemes (HCSs, Kruskal, 1964) have often been advanced as an analytic tool for extracting confusion structures from tables such as 6.5. The output of an HCS is a tree structure that visualizes which vowels constitute highly confusable subsets in the table. Alternatively, data reduction can be attempted by Multidimensional Scaling (MDS, Kruskal and Wish, 1978). We feel, however, that neither HCS not MDS do full justice to what actually goes on in the data. Both techniques presuppose a symmetrical confusion matrix, that is, the likelihood of vowel *x* to be confused with vowel *y* must be equal to that of *y* being confused with *x*. As Table 6.5 shows, this is not generally the case. Perceptual asymmetries such as those shown between tense and lax counterparts cannot be expressed in HCS or MDS; for instance, the asymmetrical confusions between /i:/ and /ɪ/ would average to a symmetrical 39%. For the sake of illustration we present just one HCS dendrogram and explain what features of the confusion structure are overlooked by the technique.

The dendrogram shows that the most confusable vowel pairs are /u:/ and /ʊ/. At approximately the same high level there is confusion between the pairs /ɛ/ and /ai/, and between /ɔ/ and /ɑu/. The tree also shows that there is just a little less confusion between the vowels /i:/ and /ɪ/ and between /ɑ:/ and /ɛə/. The /ɛ+ai/ cluster is joined at the next level by /æ/, indicating that /æ/ constitutes a more cohesive cluster with /ɛ+ai/ than any other vowel(s). In this way the tree structure seems to reveal the existence of roughly four more or less cohesive groups of vowels, plus a number of isolates. The groups would be the high back vowels (/uː, ʊ, oː/, the low front vowels /ɛ, ai, æ, ɔi/, and the low back vowels /ɔ, ɑu, ɔː/. The fourth group is not phonetically interpretable /ɑː, ɛə, ʌr/.

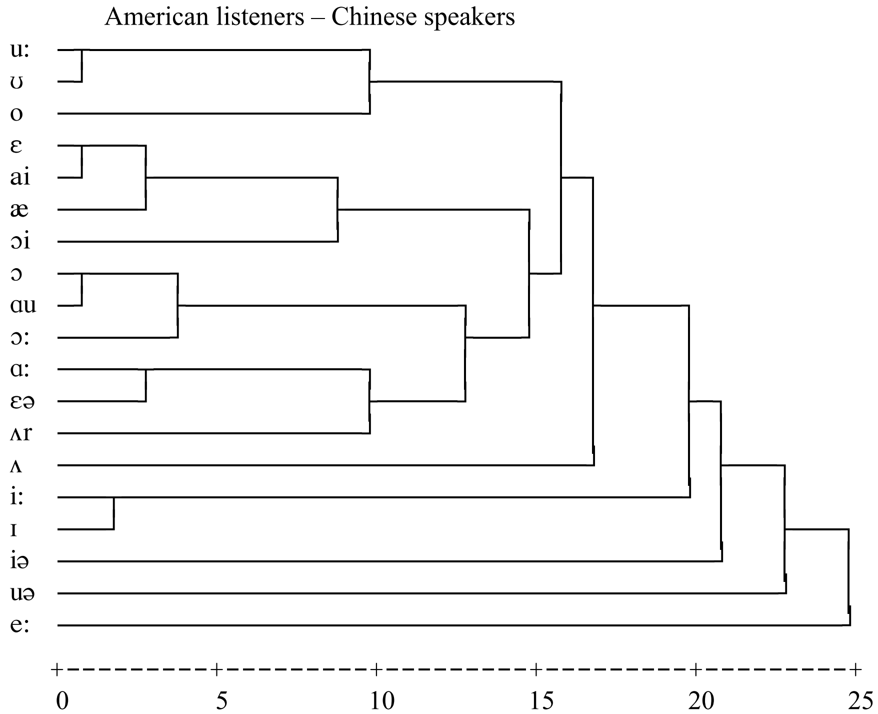

Figure 6.3. Hierarchical cluster scheme (average linkage between groups) for all 19 vowel stimuli in American responses to Chinese-accented English.

One reason why the groups are difficult to interpret phonetically is that both monophthongs and diphthongs are response categories, as are vowels followed by /r/. The vowel groups tend to be more coherent if only monophthongs are included in the trees. For the sake of completeness, all nine dendrograms (average linkage, ten phonological monophthongs only) for vowel confusions are presented in Appendix A6.3, but we will not discuss them in the text. Discussion will take place on the basis of the confusion graphs, which contain more information in a more insightful manner.

6.2.4.3 Design of the confusion graphs

We present confusion structures in the English vowels as produced by American, Chinese and Dutch speakers and as perceived by listeners of the same language backgrounds. Vowels are arranged according to the 4 (height) × 3 (backness) vowel quality grid, with a finer distinction between tense, lax and r-colored vowels by means of superposed ‘concentric’ rings. The tense vowels are located on the outer ring. Note that we have placed the vowel /æ/ more or less on the outer ring, indicating that is precisely at the boundary between tense and lax vowels. This arrangement would seem to do justice to the fact that this vowel behaves as a lax vowel in the phonological system of American English (where it is not allowed at the end of a word) and as a tense vowel from a phonetic viewpoint (long duration and extreme vowel quality). In the diagrams tense /ɔ:/ and lax /ɔ/ are kept separate, in order to demonstrate that these vowels are extremely confusable (to the point that they can be considered merged in the sound system of General American). Long /ɑ:/ has been located on the outer ring, even though in our stimulus word it was followed by /r/, because its quality appeared not to be centralized at all. Although we cannot be sure, we assume that this vowel would merge completely with /ɔ:/; the only reason why the listeners have kept the two vowels apart is because /ɑ:/ was followed by /r/ in the word *hard* which made it audibly quite distinct from *hawed*.

Confusion between any two vowels is indicated by an arrow from the intended to the non-intended vowel. The confusion percentage is indicated at the tip of the arrow. Arrows were drawn only for ‘problematic’ vowel pairs, defined as pairs that were confused in at least 20% of the responses. This is different from what did in the pilot test in which we defined problematic vowels as those vowels that were confused with some other vowel in more than 10% (Wang and Van Heuven, 2004). Since in this final test we had the Chinese subjects in China and the test was done with selected speakers, who were less proficient than those we used in the pilot test, the percentage of correctly identified vowels is lower than in the pilot test. Maintaining the more relaxed inclusion criterion of 10% vowel confusion would have yielded overly complicated and messy structures. That is why we now adopted the 20%-criterion as the definition of problematic vowels.

In the next sections I will present nine confusion graphs, one for each combination of speaker and listener nationality. The first three confusion graphs will contain the structures obtained for Chinese listeners, exposed to Chinese, Dutch and American speakers. Then I will repeat the set of three speaker nationalities for Dutch listeners, and I will conclude with the three sets obtained for American speakers.

In the confusion graphs the results for the stimulus type /ɔi/ has been omitted. This was done, firstly, to avoid visual clutter in the graphs. There seems to be no obvious place in the table where yet another (mid) low diphthong can be accommodated. A second reason why this particular stimulus category could be sacrificed, is that listeners either confuse this diphthong very strongly but in a non-systematic fashion (as happened in the case of Chinese listeners confronted with Chinese speakers) or they have no problems with this vowel at all. The problem of the Chinese speaker/listener group is not limited to this particular vowel but can be generalized to all three diphthongs /ai, ɔi, ɑu/. As a compromise, therefore, we will present the results in all the confusion graphs for the low diphthongs /ai, ɑu/ and decided to omit /ɔi/. For the full set of confusions I refer to Appendix A6.2.

6.2.4.4 Confusion structures of Chinese listeners

When Chinese listeners have to identify the English vowels spoken by fellow **Chinese speakers**, systematic confusion is found for no fewer than 15 stimulus vowels (in fact 16, if /ɔi/ had been included). The most frequent vowel confusions are /ʌr > ɑ:/ by 61%, /i: > ɪ/ by 50% and /ɔ > ɑ:/ by 46%. It shows /ɑ:/ as a big problem, not only because /ɑ:/ itself is confused with /ʌ/ in 37% but even more so because /ɑ:/ functions like a magnet, attracting massive confusions from neighboring vowels. As a result we may set up the larger group of (mid) low back vowels as a highly confusable vowel cluster. Highly problematic is also the front diphthong /ai/. It is strongly misperceived as either /i:/ or /ɪ/. It would appear, then, that the Chinese speakers emphasize the second part of this diphthong (and reduce the onset portion of the diphthong) so that it sounds rather like /i:/ or /ɪ/. Given also that Chinese listeners do not differentiate between the tense and the lax counterparts within this pair renders this a plausible confusion pattern. Interestingly, as we will see below, Dutch and American listeners do not have this problem with the Chinese /ai/. This suggests that the problem resides not so much with the Chinese speakers but with the Chinese listeners.

Although most of the confusion pairs are unidirectional (or 'asymmetrical'), there are some pairs which are confused in both directions. These are /i: ~ ɪ/, /ɛ ~ æ/ and /u: ~ ʊ/. This is what we can predict from the sound systems in the Chinese and English inventories, as the confusions are within spectrally adjacent members of tense~lax oppositions.

When Chinese listeners listen to **Dutch speakers**, they have 10 pairs of confusing vowels with no confusion pairs higher than 46%, which is lower than when they listen to their own speakers. The most frequent confusions are /ɛ ~ æ/ by 46% (the highest) and /ʌ ~ ɑ:/ by 40%. Symmetrical confusion pairs are /ɛ ~ æ/ and /u: ~ ʊ/.

When the **American speakers**' vowels are identified by Chinese listeners, there is much confusion as well. When Chinese listeners respond to American speakers there are asymmetrical confusions only. They confuse /i:/ with /ɪ/ by 50% and /ɔ/ with /ɑ:/ by 46%. The /u: ~ ʊ/ confusion is also unidirectional at 32%. The unidirectional confusion structure for the high vowels may be the result of the fact that Chinese speakers pronounce tense /i:/ shorter than the American and Dutch speakers do. Assuming that Chinese listeners attend to the duration cue rather than to spectral cues in the tense~lax contrast, there will be a bias towards perceiving the lax counterpart in these contrasts.

Chinese listeners have problems in front vowels in both Dutch and American speakers and their own speakers as well. Back vowels are difficult for Chinese listeners. The vowel /ɑ:/ remains a problem for all three groups of speakers.

Note, finally, that there are virtually no confusions that cross the boundary between front vowels and back vowels. That is to say, when a front vowel is confused, it is always with some other front vowel, and back vowels are confused with back vowels only. This also means that confusions take place mainly along the dimensions of vowel height and tenseness.

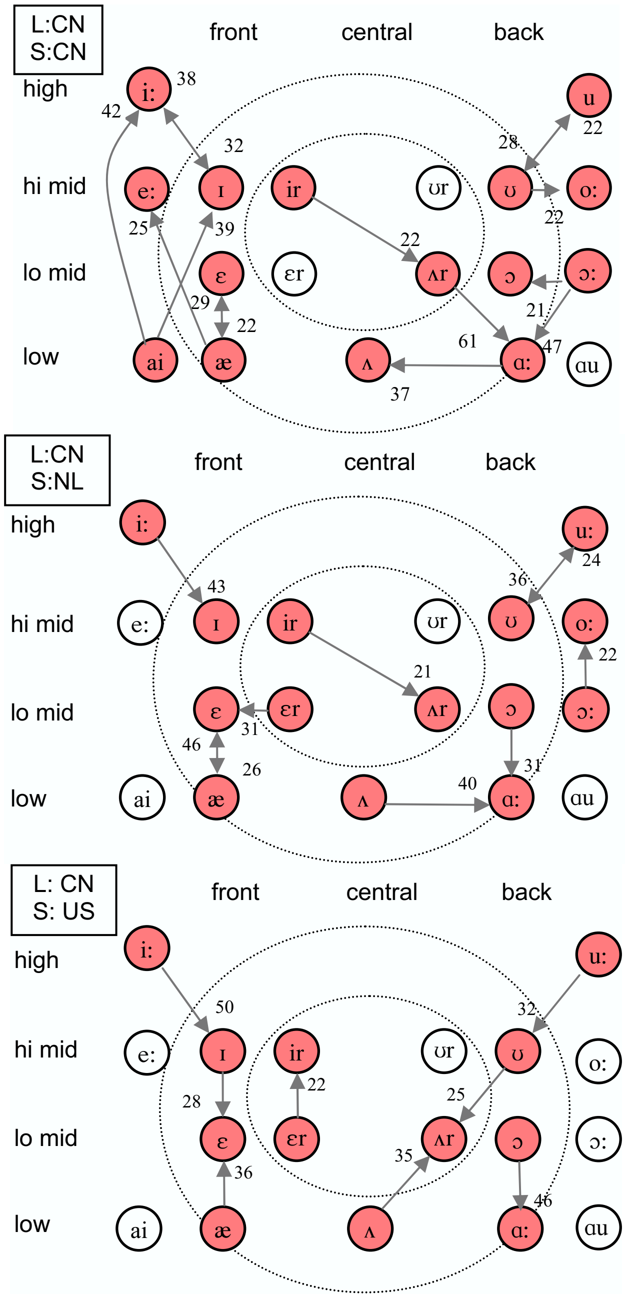

Figure 6.4. Confusion graphs for Chinese listeners, exposed to Chinese (CN), Dutch (NL) and American (US) speakers (from top to bottom). Only confusions ≥ 20% are indicated by arrows. L = listeners, S = speakers.

6.2.4.5 Confusion structures of Dutch listeners

Figure 6.5 presents the confusion structure in the vowels heard by Dutch listeners. When the speakers are **Chinese**, massive vowel confusion is observed. Lax /ɪ/ is strongly confused with tense /i:/; there is bidirectional confusion between and /u:~ ʊ/. The open front vowel /æ/ is unidirectionally confused with /ɛ/, and also with the diphthong /ai/, as is /ɛ/. In the ear of the Dutch non-native listener these three vowels spoken by Chinese learners are very poorly distinguished. Also Chinese /ʌ/ is unidirectionally mistaken for /æ/. Finally, the four low back vowels are strongly confused – and therefore poorly differentiated in the ear of the Dutch listeners. There is one confusion across front and back vowels: /ɔ:/ > /ai/.

When Dutch listeners respond to Dutch speakers of English the confusions are restricted to just four pairs, which seem rather predictable from a contrastive analysis of the Dutch and English vowel systems; these are the pairs /ɛ ~ æ/ and /u: ~ ʊ/. In addition to these there is unidirectional confusion of /ɔ/ to /o/ and of central /ʌ/ to /æ/. This latter confusion is unexpected but does in fact mirror the same confusion when the speakers are Chinese. It would indicate that both the Chinese and the Dutch speakers realize some of the /ʌ/ tokens rather too front and too low.

When the speakers are **Americans** the number of confusions is minimal. Dutch listeners confuse both open /æ/ and the long half-closed vowel /e:/ for /ɛ/, the first because they have no clear category boundary between /æ/ and /ɛ/, the latter possibly because the American speakers make their tense /e:/ too short for the Dutch norm of a tense vowel, or because the American onset of /e:/ is lower than what is expected by a Dutch listener. American /ʌ/ is confused with lax /ʊ/. This would indicate that Americans pronounce /ʌ/ more back than the Dutch and Chinese speakers do, and that the Dutch target of this vowel has a more front and low position. Long /u:/ is unidirectionally confused with lax /ʊ/, and, finally, /ɔ:/ in *hawed* is strongly confused for /ɑ:/ as in *hard*. This is quite likely due to the fact that American speakers tend to neutralize the contrast between these vowels.

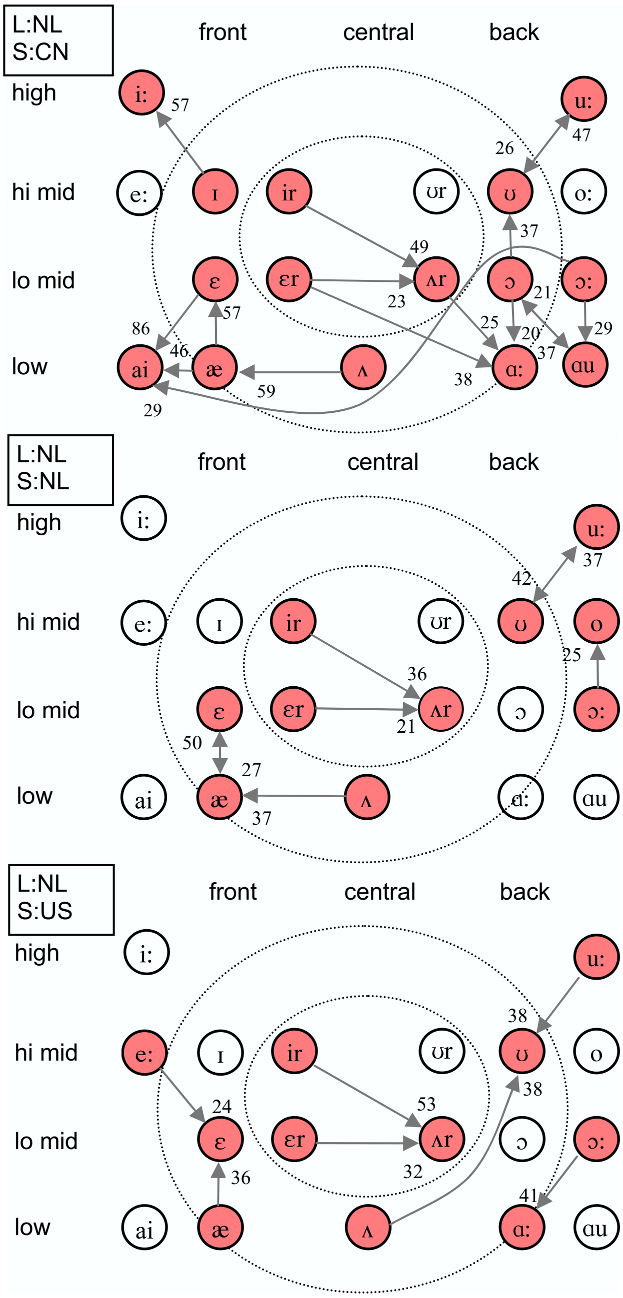

Figure 6.5. Confusion graphs for Dutch listeners, exposed to Chinese (CN), Dutch (NL) and American (US) speakers (from top to bottom). Only confusions ≥ 20% are indicated by arrows. L = listeners, S = speakers.

6.2.4.6 Confusion structures of American listeners

When American listeners listen to **American speakers** there is relatively little confusion in the vowels. The literature shows that even in such situations vowel identification is far from perfect, with scores ranging between 54 and 88 % correct (Peterson and Barney, 1952; Strange, Verbrugge, Shankweiler and Edman, 1976 and references therein). Our results are no exception since the identification performance ranges between 38 and 96%, depending on the vowel type. Still, only two vowel pairs were confused in more than 20%. These are the confusion of lax /ʌ/ > /ʊ/ and the partial merger of tense and lax /ɔ/.[3] The latter confusion was also the most frequent one in the classical study by Peterson and Barney (1952) but the /ʌ/ > /ʊ/ confusion, although it did occur in the classical data, only ranked third there (after /ʌ/ > /ɔ:/ and after /ʌ/ > /ɔ/). It would seem, therefore, that our American speakers pronounced their /ʌ/ somewhat differently than the Peterson and Barney speakers did. The most confusing pairs for both Chinese and Dutch listeners, /ɛ ~ æ/ and /u:~ ʊ/ (see below), do not constitute any problem for American listeners when they listen to their own speakers, so that, clearly, the Dutch and Chinese speakers fail to observe a contrast here.

The vowels produced by the **Dutch speakers** show some confusions that are also mentioned in the literature, i.e. /æ/ > /ɛ/ and /ʊ/ > /u:/ (see Table 3.5). These confusions are unidirectional: clearly the Dutch way of pronouncing /æ/ is not open enough, hence the unidirectional confusion with /ɛ/. The data also suggest that Dutch /u/ resembles American tense /u:/ more than its lax counterpart. There is considerable confusion of /ʌ/ > /ɔ/. Dutch-accented /ʌ/ tends to be too far back, causing unidirectional confusion with /ɔ/. This confusion was foreshadowed in Table 3.5, where the observation was made that (advanced) Dutch learners often substitute their /ɑ/ for /ʌ/. Not expected from Table 3.5 would be the remaining confusion, i.e. /ɔ:/ > /o:/ This confusion would indicate that the Dutch speakers tend to make the English /ɔ:/ too close.

When Americans listen to **Chinese speakers** there is confusion of height and tenseness in the high-front as well as in the high-back vowels. The American listeners have problems in identifying high and low front vowels and high and low back vowels with 15 confusion pairs (with 69% confusion for the most problematic vowel pair). There are four bidirectional pairs /i: ~ ɪ/, /u: ~ ʊ/, /i ~ æ/, and /ɔ: ~ au/. The former two confusions were also listed as pronunciation problems for Chinese learners of English in Table 3.6; the latter two confusions have not been noted in the pedagogical literature.

This configuration is also isomorphic to the pattern found for Dutch listeners exposed to Chinese speakers. This conformity shows that the source of the problem resides in the pronunciation of the Chinese speakers rather than in the perception of the Dutch listeners.

---

[3] See the following quote from Peterson and Barney (1952: 178): "The very low scores on [ɔ:] and [ɔ] … undoubtedly result primarily from the fact that some members of the speaking group and many members of the listening group speak one of the forms of American dialects in which [ɔ:] and [ɔ] are not differentiated ."

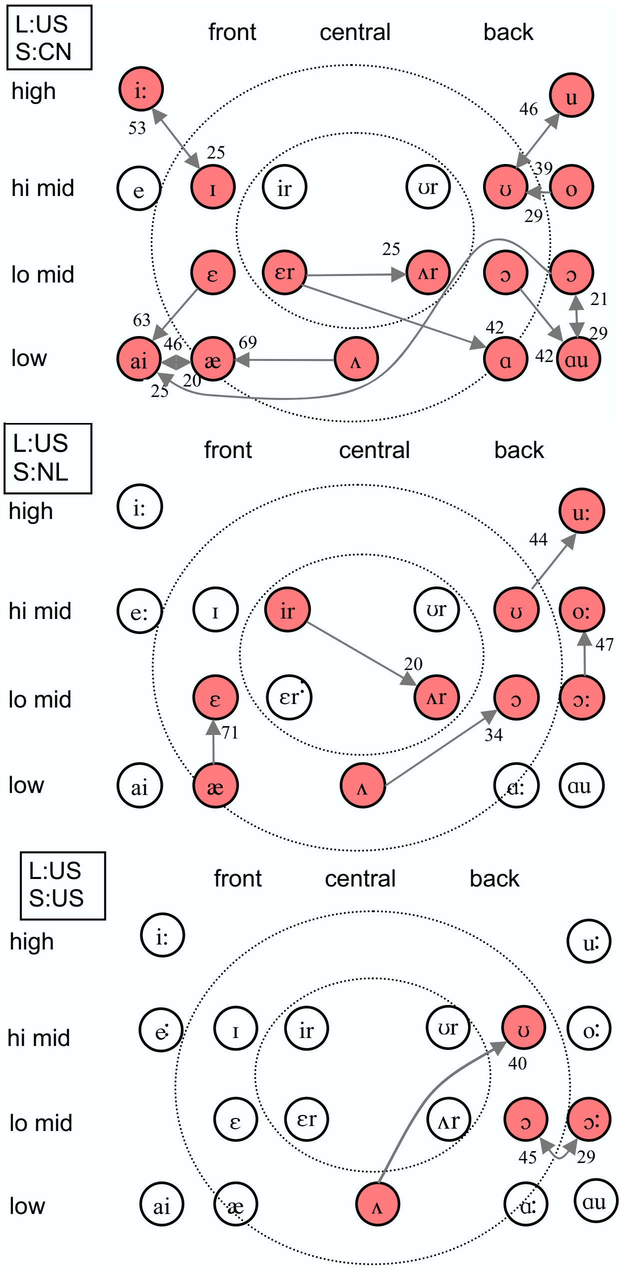

Figure 6.6. Confusion graphs for American listeners, exposed to Chinese (CN), Dutch (NL) and American (US) speakers (from top to bottom). Only confusions ≥ 20% are indicated by arrows. L= listeners, S = speakers.

6.3 Summary

By way of summary Table 6.5 lists the numbers of problematic vowels in the data. Here a problematic vowel is more or less arbitrarily defined as a vowel which in any speaker-hearer combination is identified correctly in less than 75%. The numbers are broken down for the nine combinations of speaker and listener language background.

Table 6.5. Number of problematic vowels (see text) broken down by nationality of speaker and of listener.

speaker	listener			
	Chinese	Dutch	USA	Total
Chinese	**19**	15	17	51
Dutch	19	**13**	11	43
USA	19	12	**4**	35
Total number	57	40	32	139

Table 6.4 shows that, overall, American native listeners have fewer problems with the English vowels than L2 listeners. Dutch listeners are a good second, and Chinese listeners clearly have problems. More generally, the language background of the listener exerts a stronger influence on the number of confused vowel pairs than the L1 of the speaker.

6.4 Conclusion and discussion

Our first hypothesis was that English vowels will be more difficult to identify as the foreign speaker's native language is more unlike English by the interference of their L1s. We predict, then, that Chinese-accented English vowels will be more difficult to identify for native English listeners than, for instance, Dutch-accented vowels. Conversely, English vowels produced by native English speakers should then be more intelligible to Dutch listeners than to Chinese listeners. Both predictions were clearly borne out by the experimental results. Although these results can indeed be seen as experimental support for our typological distance hypothesis, it should be pointed out that cultural and educational differences between the People's Republic of China (with little exposure to English) and the Netherlands (with an abundance of English) may also have contributed to the difference in intelligibility.

The confusion structure in the foreign-accented Englishes can partly be accounted for by a contrastive analysis of the vowel inventories of the target and source languages involved. For Dutch-accented English, we predicted problems with the non-high lax front vowels /ɪ ~ e ~ æ/ and with the /uː ~ ʊ/ contrast. The results show that these were, indeed, the most frequent confusion types, not only when L1 English listeners identified Dutch-accented vowels, but also when Dutch L2 listeners identified native English vowel tokens. Moreover, our contrastive analysis predicted that Chinese-accented English would have all the problems of Dutch English but would additionally suffer from massive tense~lax vowel confusion, both in production and in perception. The experimental results show that this prediction is correct.

On the other hand, we found a number of problematic vowel contrasts that are not easily predicted from a contrastive analysis, e.g. the /uː > oː/ and /ɔː > oː/ confusions for Dutch speakers identified by American listeners. We did not encounter any cases where predicted problems did not arise. Our results, then, provide partial support for the transfer hypothesis in foreign language learning, which claims that L2 learners will not distinguish between contrasts in the target language that do not occur in their native tongue. At the same time, a weaker version of the transfer hypothesis seems called for, in that, although it makes no false predictions, it predicts only a subset of the L2 vowel learning problems.

Many of the confusions found in this chapter were mentioned as learning problems in the pedagogic literature on the learning of English as a second language for Dutch and Chinese learners, in Tables 3.5 and 3.6, respectively. However, we noted some pronunciation problems that were not mentioned in the tables, indicating that occasionally pronunciation problems escape the trained ears of foreign-language teachers. I would claim that such problems can only be brought to light by experimental methods such as those used in the present study.