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

Wang, H.

Chapter three

Contrastive analysis

A classic question in phonetic theory is: "What sounds can a language have?” It has been asked about vowel inventories (Lindblom, 1986) and consonant inventories (Lindblom and Maddieson, 1988). Every speech sound belongs to one or other of the two main classes known as vowels and consonants. Describing the vowel and consonant inventories is a start in describing the salient phonetic structure of a language. The standard view is that the sound inventory in a language is the result of two competing forces: one favors sounds that are easy to produce while the other force pulls the system towards more distinctiveness, i.e. maximal contrast between elements of the system (Lindblom, 1986). The native language (source language) of L2 learners plays the role of a ‘Phonological filter’ which deeply influences the L2 learners’ pronunciation so that it deviates from that of native speakers of the target language (Polivanov, 1931; Trubetzkoy, 1939/1969). As a result L2 learners have perceptual blind spots which lead to perceptual errors. These blind spots prevent the L2 listeners from identifying the foreign phonemes correctly. Instead, they substitute their own L1 sounds for the foreign phonemes.

In this chapter I will give the sound inventories of three languages, General American English (GA),[1] Standard Dutch (Algemeen Beschaafd Nederlands, ABN)[2] and Standard Mandarin Chinese (Putonghua).[3] These are the native languages of the three groups of speakers and listeners that will be studied in the present dissertation. More details on the choice of experimental subjects will be provided in Chapter four.

3.1 Vowels

A vowel is defined as a typically voiced sound in the production of which the air issues in a continuous stream through the pharynx and mouth, there being no obstruction and no narrowing such as would cause audible friction. Vowels are

[1] The American subjects for the final experiment in this research are from Los Angeles, California, which is generally regarded as a place where GA is used. Californian English is a dialect of the English language spoken in the U.S. state of California. As a variety of American English, Californian English is similar to most other forms of American speech in being a rhotic accent, which is historically a significant marker in differentiating different English varieties.

[2] The Dutch subjects come from the cities around Leiden called ‘city belt’ (Dutch: Randstad), where people speaking standard Dutch, ABN, can easily be found.

[3] The language for the Chinese subjects in our experiment is Standard Chinese, which is the present-day dialect of Beijing promulgated as a standard language in Mainland China, Taiwan and Singapore. Our Chinese subjects come from the northeast of China, Changchun, where people speak the Northeast dialect, which is very close to Standard Mandarin Chinese.

usually described in terms of quality and duration. Since vowels are distinguished from one another chiefly by whether they are produced in the front, centre, or back of the mouth, whether the tongue position is high, mid or low, and whether the lips are spread or rounded, the basic building blocks of most vowel systems are the three qualities, as many of the vowels of the world's languages can be described simply by the three traditional dimensions high-low, back-front, and rounded-unrounded.

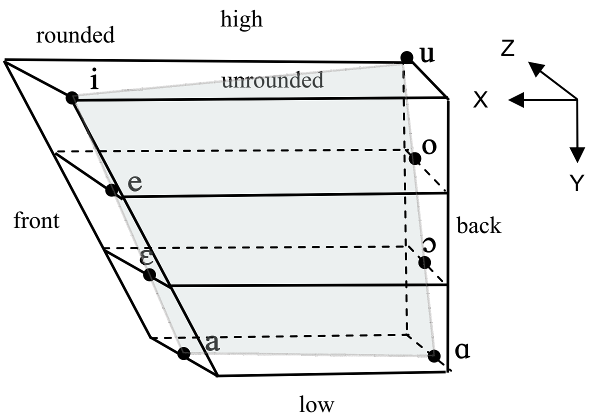

Figure 3.1. The location of the eight cardinal vowels in a three-dimensional articulatory vowel space defined by backness (X dimension), height (Y dimension), and rounding (Z dimension) [after Ladefoged (1971: 72), Ladefoged and Maddieson (1990: 94)].

This figure shows the location of a set of reference vowels, i.e., the cardinal vowels described by Jones (1956), within a space defined by these dimensions. What Jones effectively gave phonetics in his CV system (Cardinal Vowel system) was a mapping system which presented what is essentially auditory and acoustic information in a convenient visual form.[4] It is the only widely used system for vowel description. It gave phoneticians a yardstick for measuring the vowel quality which is invaluable in phonetic description.

Another element which is considered by some to be of importance in determining vowel quality is the state of the tongue and lips as regards muscular tension. Those who consider that vowels may be differentiated by degrees of muscular tension distinguish two classes, tense vowels and lax vowels. Tense vowels are supposed to require considerable muscular tension on the part of the tongue; in lax vowels the tongue is supposed to be held loosely. The difference in quality between the English vowel *seat* and *sit* is described as a difference in tenseness: the vowel in *seat* is considered tense and the vowel in *sit* lax (Jones, 1956).

[4] Jones's system has been be criticized by Collins and Mees (1981) as follows: '[Jones] took no account of the significance of the root of the tongue and its relationship to the pharynx wall. Indeed, he disregarded the pharynx cavity altogether, mentioning only tongue height in his theory. Later research has shown that it is the relative sizes of the oral and pharyngeal cavities which are the crucial factors in vowel quality'.

In some languages there are vowels which are distinguished by duration alone. For instance, in Danish, there is an opposition between long and short vowels, and in Estonian even between short, long and superlong (Lehiste 1970). In many languages, similar oppositions between sets of vowels are also marked by differences in vowel quality. Such combinations of duration and vowel quality are employed in Dutch and English.

3.1.1 Vowel inventories in the three languages

In the following sections I will compare the vowel inventories of English, Dutch and Mandarin Chinese. I will present literature data consisting of structural vowel tables published for the languages and of formant measurements. Formant measurements have been used since 1950 as an semi-objective way to determine vowel quality. The technique will be discussed at greater length in Chapter four; for the purpose of the present chapter it is sufficient to know that the centre frequency of the lowest resonance in the speech signal (first formant, F1) varies with the degree of mouth opening and that the second-lowest formant (F2) corresponds inversely with the degree of backness. Acoustic vowel charts plot F1 from top to bottom against F2 from right to left; in this way the configuration of vowel points assumes the same orientation as in a traditional articulatory vowel chart, with /i/ in the top left-hand corner, /u/ in the top right-hand corner, and /a/ at the bottom. In the charts we present below, we did not plot the formant frequencies in hertz but transformed the hertz-values to Bark units. Equal distances in the Bark space correspond to equal differences in perceived timbre (or: vowel quality). For details we refer to Chapter four.

3.1.1.1 English vowels

The vowel system of General American English (GA, as exemplified for instance in the American English pronouncing dictionary by Kenyon and Knott, 1944, see also Gussenhoven and Broeders, 1976: 186-195) is best described as composed of four vowel heights and three degrees of backness. Height is a four-level parameter with high, high-mid, low-mid and low as the phonetically relevant degrees. Backness has three degrees, viz. front, centre and back. English has a split in its vowel system such that most vowels are tense (long duration, peripheral articulation) but some are lax (short duration, more centralized pronunciation). The four degrees of height are defined on the tense vowel set; the back vowels require four degrees. When tense and lax vowels are kept apart, three degrees of height suffice for the front vowels (high/close for /iː/ (*heed*, *bead*), mid for /eː/ (*hayed*, *stayed*) and low/open for /æ/ (*had, mad*). For the back vowels, however, we have to distinguish between high/close /uː/ (*who'd*, *mood*), high-mid /oː/ (*hoed*, *showed*), low-mid /ɔː/ (*hawed*, *clawed*), and low/open /ɑː/ (*father*). In many American dialects and probably also in General American, /ɔː/ and /ɑː/ have merged (Wells, 1982; Labov, Ash and Boberg, 2006), simplifying the vowel system to three degrees of height, and restoring symmetry between front and back vowels. The lax vowel set comprises just two degrees of height (high vs. low). Textbooks on British English often mention so-called centring diphthongs as an extra set of vowel phonemes, as in *fear*, *fair*, *poor*.

These vowels could be claimed to be phonemes on the strength of such minimal triplets as *bead ~ beard ~ bid*. It seems to me, however, that these vowels can be treated as positional allophones of tense vowels followed by coda-/r/. The reason why the underlying vowel should be tense rather than lax has to do with the phonotactics of the centring diphthongs: they cannot be followed by any other consonants than the alveolars (/t, d, s, z/), which is the same environmental constraint that applies to other tense vowels; lax vowels can be followed by a larger variety of consonants and clusters, which are not possible as codas after murmur diphthongs. For instance, centring diphthongs cannot be followed by coda clusters except when the last consonant is one of the set /t, d, s, z/, i.e., the set covering the suffixes used to code plural, past tense, or third person singular after stems with either voiceless consonants (/-t, -s/) or with voiced sounds (/-d, -z/). Quite probably also [ə:] should be analysed as a surface phenomenon in non-rhotic varieties of (British) English. In rhotic varieties, and especially in General American, this vowel sound can be analyzed as /ʌ/ followed by coda-/r/.

GA has two diphthongs, /ai, ɑu/, which start at an open position and glide towards a close position along the front and back side of the vowel space, respectively. The third diphthong is /ɔi/, which runs from back to front in the mid part of the space. Table 3.1 summarizes the vowel inventory of GA. The unstressed neutral vowel schwa (/ə/) is not included in table 3.1.

Table 3.1. The General American vowel inventory. Vowels in parentheses are allophones in GA before /r/, but have surface-phonemic status in RP English.

	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	

Figure 3.2 presents the classical formant data collected for American English by Peterson and Barney (1952) drawn separately for male (squares) and female (circles) speakers, and broken down by the tense (solid lines) versus lax (dotted lines) subsystems.[5] The F1 and F2 frequencies have been transformed to Bark units, in

[5] Peterson and Barney (1952) identified the primary acoustic features of the American English vowels on the basis of /hVd/ productions by 28 women, 33 men, and 15 children (ages not specified). They found a general correspondence between vowel type and frequencies of the first and second formants (F1 and F2). Hillenbrand, Getty, Clark and Wheeler (1995) replicated and extended the Peterson and Barney study. Hillenbrand et al. sampled 45 men, 48 women, and 46 10- to 12-year-old children. Analysis of formant data by Hillenbrand et al. showed differences from the formant data in the Peterson and Barney study, both in terms of mean frequencies of F1 and F2, and the degree of overlap among adjacent vowels. However, the data were similar to Peterson and Barney regarding vowel-specific formant frequencies, as well as change in formant values according to vocal tract size and shape.

order to create a visual display in which equal distances between vowels represent auditorily equal differences in vowel quality (timbre). Note that the vowel /æ/, which is a lax vowel in terms of its distributional properties (may not occur in an open syllable at the end of a word) is considered a tense vowel (see § 5.2). Also, the vowel /ɔ/ is treated as tense (see § 5.3).

The American English vowel system consists of 11 distinct vowels (or monophthongs) /i, ɪ, e, ɛ, æ, ʌ, u, ʊ, o, ɔ, ɑ/ (Peterson and Barney, 1952). Categorization of vowels according to features of tongue articulation reveals a vocal tract vowel space which consists of four distinct corners corresponding to a quadrilateral shape. Vowels identified for each corner are /i/ (high-front), /æ / (low-front), /u/ (high-back), and /ɑ/ (low-back).

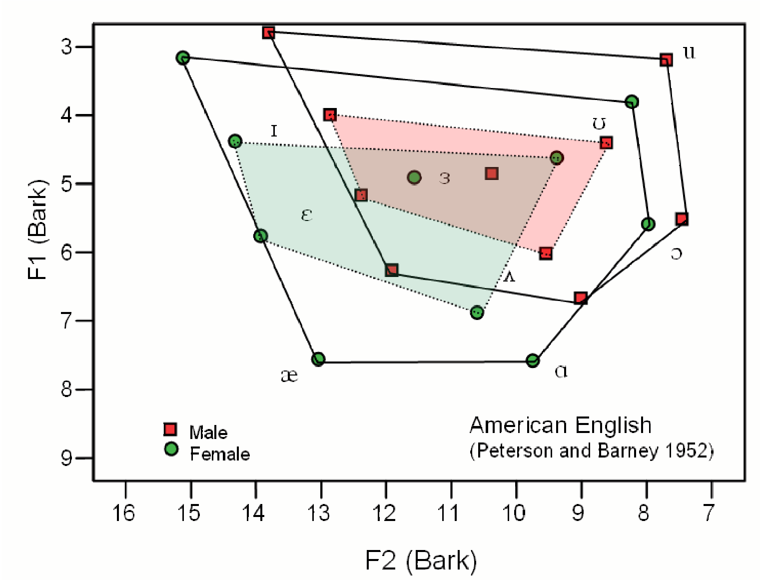

Figure 3.2. The tense (solid lines) and lax (dotted lines) vowels of General American plotted in an F1 (top to bottom) by F2 (right to left) display. Male (squares) and female (circles) vowels have been plotted separately. (After Peterson and Barney, 1952, with Bark-transformed frequency values for F1 and F2).

3.1.1.2 Dutch vowels

The Dutch vowel system (Table 3.2) is in many respects similar to English. It also has tense and lax vowels, and distinguishes four degrees of height and three degrees of backness. However, the central part of the vowel space is more densely filled as Dutch has (rounded) central high and high-mid vowels. In the lax front vowels Dutch distinguishes two degrees of height for the /ɪ ~ ɛ/ contrast, where English has three: /ɪ ~ e ~ æ/. Dutch is underdifferentiated relative to English in the high back vowels, where English has the tense ~ lax opposition /uː ~ ʊ/ while Dutch only has

/u:/. Dutch also has a number of vowels that are absent in the English system; such overdifferentiation is hardly ever a source of confusion (cf. Lado, 1957).

Dutch has three full diphthongs, /ɛi, œy, ɑu/, the first two of which have their starting point at a low-mid vowel height and the latter at a fully open position. Also, Dutch has some degree of diphthongization on the tense high-mid vowels, so that /e:/, /ɸ:/ and /o:/ are realized as [eⁱ], [ɸʸ] and [oᵘ], respectively. There is a sixteenth vowel, schwa, which is not included in the table. This neutral vowel cannot be stressed; if it is, it will change to /ɵ/, the rounded lax central vowel.

Table 3.2. The basic Dutch vowel inventory (Rietveld and Van Heuven 2001).

	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

Figure 3.3 gives the arrangement of the twelve monophthongs of Dutch (excluding schwa) in an acoustic vowel diagram. Dutch also has tense and lax vowels, but the lax subsystem seems reduced along the height dimension only, not also along the backness parameter, as it is in English.

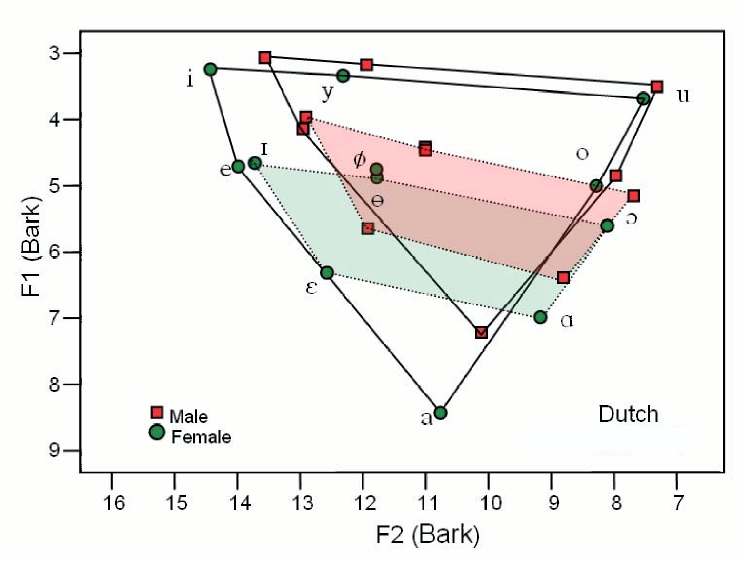

Figure 3.3. Dutch monophthongs plotted in an F1 by F2 plane (Barks). Male data (50 speakers) adapted from Pols et al. (1970), female data (25 speakers) from Van Nierop, Pols and Plomp (1973).

3.1.1.3 Chinese vowels

There has been a longstanding controversy in the literature on the number of underlying vowel categories in Mandarin, and the relationship of the myriad of surface vowel forms to these phonemic categories (e.g., Chao, 1934, 1968), R. Cheng, 1966; C. Cheng, 1973; Pulleyblank, 1984; Lin, 1989; Wang, 1993; Wu, 1994). The reason for this controversy is that most phonetic manifestations of vowels in Mandarin occur in a fairly narrow range of contexts, which suggests that they probably can be reduced to a smaller set of basic vowel categories. There is disagreement both on the number of surface (phonetic) vowels in Mandarin as well as on the number of underlying, abstract (phonological) vowels. Surface vowels can be as many as twelve or thirteen; the number of underlying vowels varies between four and six (Wan and Jaeger, 2003). The large majority of sources distinguish twelve surface vowels (see also Flege et al., 1997; Li and Thompson, 1981; Light, 1976; Maddieson, 1984; Wu 1964), which can be reduced to a smaller number of underlying vowels in different ways, yielding different numbers. We assume that positive and negative transfer of vowels from L1 to L2 is located towards the surface level rather than at some deep level of representation. Cheng (1966) relates the twelve surface vowels to their underlying forms as follows:

/i/	→	[i], [ɾ],[ɩ]
/y/	→	[y]
/u/	→	[u]
/ə/	→	[e], [ə], [o], [ɤ]
/a/	→	[a], [ɑ], [ɛ]

The twelve surface vowels can be represented in a structural way as exemplified in Table 3.3.

As the table shows, Mandarin has no length (i.e. no tense ~ lax) contrast; contrasts such as /iː ~ ɪ/, /uː ~ ʊ/, /ɔ ~ ɔː ~ oː/, and /e ~ æ/ do not occur.

Table 3.3. The Mandarin surface vowel inventory.

	front		central		back	
	–round	+round	–round	+round	–round	+round
high	i	y	ɩ			u
high-mid	e		ɚ		ɤ	o
low-mid	ɛ					ɔ
low			a			ɑ

Chinese is different from both Dutch and English because Chinese is a tone language, in which tones are lexically specified. In general, all full syllables carry a lexical tone, whereas weak syllables have the neutral tone (or are 'toneless'). As far as we know, however, the tones of Chinese do not interfere in any way with the production or perception of English sounds by Chinese learners. This does not rule out the possibility that tonal interference may be found in the learning of English (sentence) prosody, but this is outside the scope of the present research.

Studies examining acoustic characteristics of vowel production in Mandarin are limited. Wu (1964) examined vowels produced by Mandarin speakers (four male adults, four female adults, four children). His measurements included, among other properties, F1, F2, F3 frequencies of formants for six standard vowels /i, e, y, u, o, ɑ/, as well as of allophones of /i/ and /e/. A later study by Howie (1976) acoustically analysed these six vowels produced by two male speakers.

Figure 3.4 presents formant measurements of F1 and F2 plotted in the same way as we did for English and Dutch. These formant values were published by Li, Yu, Chen and Wang (2004) for five male and five female speakers of Mandarin producing seven monophthongs /i, y, u, e, ɚ, o, ɑ/. Five of these have a fairly unrestricted distribution; /e/ and /o/, however, may be considered allophones of /i/ and /u/, respectively, which surface in specific environments only.

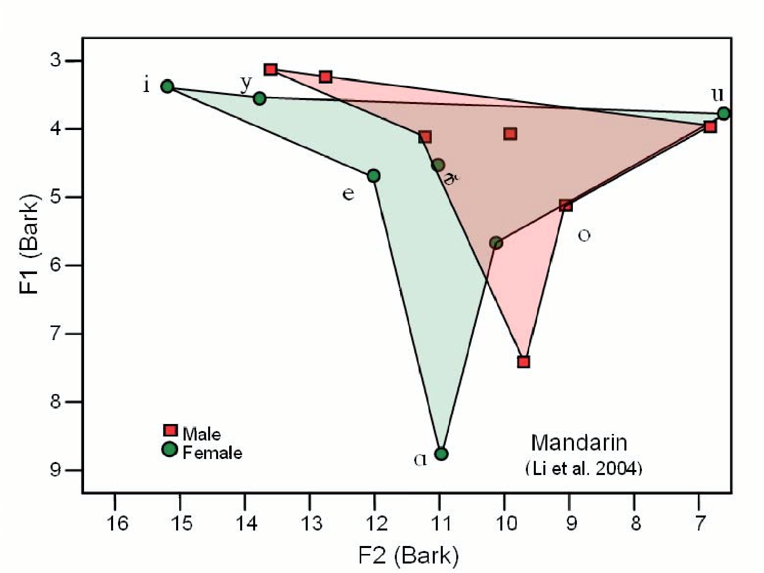

Figure 3.4. F1 versus F2 (Bark) for seven monophthongal vowels of Mandarin (Beijing dialect) spoken by five men and five women (adapted from Li et al. 2004). [6]

3.1.2 Prediction of pronunciation problems in vowels

In Table 3.4 below, I have attempted to present together the vowel inventories of Dutch, Mandarin and (American) English in a crude contrastive analysis. Here I use the principles that were advocated by Lado (1957), and which also underlie the

[6] F2 for male /u/ is specified by Li et al. (2004: 257) as 9.147862 Bark. I assume that the first digit is in error and corrected it to 7. This decision is supported by Figure 3 in Li et al.

categories of Flege's (1987) Speech Learning Model (SLM), in order to define three classes of speech sounds in a target language. The first is the category of identical sounds. These are sounds that are transcribed with the same narrow IPA symbol in source and target language; they should constitute no learning problem. In the table they have been left unmarked. The second category are sounds in source and target language that are written with the same IPA base symbol but differ in diacritic marks. These sounds are phonetically similar but not identical; such similar sounds are predicted to constitute long-term learning problems in second-language acquisition. In the table, similar sounds are indicated in grey cells. The third type are new sounds. Here a sound is needed in the target language which does not occur in the source language. The sound in the source language that is phonetically closest to the target is written with a different base symbol in the IPA. The prediction is that such new sounds constitute a learning problem in the initial stages of the acquisition process, but sooner or later the new category will emerge, and that it will be quite authentic. In the tables, new sounds are printed in white against a black background.

Table 3.4. Contrastive vowel analysis of Dutch and English (upper panel) and of Mandarin and English (lower panel). Grey cells in source languages denote source sounds that are not needed in the target language. White, grey and black cells in the target language represent identical, similar and new sounds, respectively.

V-height	Place of Constriction					
	Front		Central		Back	
Source: Dutch						
	Tense	Lax	Tense	Lax	Tense	Lax
High	i		y		u	
High-mid	e:	ɪ	ø:	ɵ	o:	
Low-mid		ɛ				ɔ
Low			a:			ɑ
Diphthong	ɛi		œy			ɑu
Target: English						
	Tense	Lax	Tense	Lax	Tense	Lax
High	i:				u:	
High-mid	e:	ɪ			o:	ʊ
Low-mid		ɛ				
Low	æ			ʌ		ɒ
Diphthong	ai		ɔi			ɑu

Table 3.4. Continued.

V-height (down)	Place of Constriction (across)					
	Front		Central		Back	
Source: Mandarin						
	−round	+round	−round	+round	−round	+round
High	i	y	ɩ		u	
High-mid	e		ɚ		ɤ	o
Low-mid	ɛ					ɔ
Low			a			
Diphthong	ai		ɔi			ɑu
Target: English						
	Tense	Lax	Tense	Lax	Tense	Lax
High	i:				u:	
High-mid	e:	ɪ			o:	ʊ
Low-mid		ɛ				
Low	æ			ʌ		ɒ
Diphthong	ai		ɔi			ɑu

It is rather unclear how realistic the predictions of SLM are. The vowels in the Dutch inventory have an unrestricted distribution, and can readily be employed in English. Some of the Mandarin vowels are highly context-sensitive allophones, which may or may not generalize to English. Moreover, Mandarin has no length (or tense~lax) contrast. The lax members of the opposition in English are transcribed with separate base symbols – and are therefore new sounds. The tense (long) members differ from the Mandarin counterparts in a diacritic only (length mark) and are therefore similar sounds.

In the next sections we will review comments made by experts on English pronunciation teaching to Dutch (§ 3.1.2.1) and Mandarin (§ 3.1.2.2) learners. These comments are not predictions based on an a priori comparison of source and target sound systems but summarize classroom experience.

3.1.2.1 Dutch ~ English

This section summarizes comments made in pronunciation text books at the university level for Dutch learners of English (e.g. Gussenhoven and Broeders; 1976, 1981; Collins, Hollander and Rodd, 1977; Collins and Mees, 1981). When in these comments Dutch and English are called similar, the term does not necessarily have the same status it has in Flege's SLM. The authors of the textbooks, who are accomplished phoneticians with a keen ear for minute phonetic differences between sounds, hardly ever call a pair of sounds in source and target language identical or the same. Therefore 'similar' sounds may refer to pairs of Dutch/English sounds that are written with the same base symbol and diacritics. The summary is presented in Table 3.5.

Table 3.5 Survey of pronunciation problems with vowels by Dutch learners of English, derived from Collins and Mees (1981) and Gussenhoven and Broeders (1976: 88). D = Dutch, E = English.

English target	Dutch substitutions/typical errors/comments
/iː/	Absent in D; similar sound D /i/ is the usual replacement. Too close, too front, especially too short. The articulation is also considerably tenser than E /iː/.
/ɪ/	Absent in D; D /ɪ/ is similar to this sound. Generally D learners have no problem with this sound.[7]
/eː/	Similar to D /eː/. Both E and D /eː/ are phonetically diphthongized, E /eː/ has a slightly lower onset and a stronger glide element. D /eː/ is within the range of acceptable pronunciations of E /eː/.
/ɛ/	D learners use similar D /ɛ/ as a replacement. D learners are generally unaware of the E /æ~e/ contrast so that perceptual confusion may result.
/æ/	Absent in D; most D learners will substitute D /ɛ/. Perceptual confusion is predicted between E /æ~e/.
/ɑː/	Absent in D; typically replaced by D /aː/, which varies considerably in quality. There is considerable overlap between D /aː/ and E /ɑː/.
/ɔː/	Absent in D; there appears to be no regular substitution from Dutch speakers. Some use an extended D /ɔ/, whilst others use the marginal vowel D /ɔː/ (as in French loan words). Others use the allophone of D /oː/ that occurs before D /r/.
/oː/	E /oː/ and D /oː/ are phonetically realised as diphthongs. E /oː/ has lower onset and somewhat stronger diphthongization but no perceptual confusion will arise with any other E vowel.
/ʊ/	Absent in D. E /ʊ/ is perhaps the most difficult vowel for D learners. There is no D vowel near E /ʊ/. Most D speakers confuse E /ʊ/ and E /uː/, hearing both in terms of D /u/.
/uː/	Absent in D, but similar to D /u/. Some speakers substitute D /u/. This vowel is closer to E /uː/, and the D sound is shorter (except before /–r/). The D articulation is also tenser. Most D speakers regularly confuse E /uː ~ ʊ/.
/ʌ/	Absent in D; D learners tend to substitute D /ɵ/ for E /ʌ/. More advanced students sometimes substitute D /ɑ/ for this sound.
/əː/	Absent in D; usually replaced by D /øːr/ or /ɵr/. Neither substitution is acceptable, having inappropriate lip-rounding, and too close a tongue position.

[7] Speakers from The Hague, Rotterdam, Amsterdam, Antwerp, may confuse E /iː~ ɪ/. Speakers from Dordrecht, Nijmegen, Noord-Brabant and Limburg may have a very open quality, which may give rise to confusion with English /e/.

Table 3.5. Continued

/ə/	The quality of D /ə/ is similar to that of E /ə/ in most contexts and transfers well into E. Difficulty may arise word-finally. D final allophone tends to be closer and is rounded, giving a markedly different effect from the very open word-final E. E /ə/ is similar to /ʌ/ in syllable final-position, D /ə/ is closer to D /ɵ/.
/ɒ/	Absent in D. The usual D substitution for E /ɒ/ is D /ɔ/, which is too close, too round and generally over-tense. Tenseness of D realisation of E /ɒ/ is especially noticeable before fortis plosives. A less common error is to pronounce /ɒ/ too front and unrounded, so losing contrast with /ʌ/. Mispronunciation of /ɒ/ is a very persistent error, often heard from otherwise proficient speakers. It appears to be difficult for D native speakers to detect.
/ai/	Absent in D; is often replaced by D VC sequence /a:j/, whose vowel part is too long, esp. before voiceless plosives, so that confusion may arise with voiced plosive (e.g. *tight ~ tide*).
/ɑu/	Similar sound D /ɑu/ is substituted; its onset may be too rounded but no perceptual confusions arise.
/ɔi/	Absent in D; D vowel+glide sequence /o:j/ is often substituted, whose onset is too close but does not lead to perceptual confusion

3.1.2.2 Chinese ~ English

Although many textbooks have been produced describing the differences between the sound systems of Dutch and English (see above) and giving detailed analyses of pronunciation errors of Dutch learners of English, such studies are virtually non-existent for Chinese learners of English. In fact, I know of just one pedagogical study by Zhao (1995), which makes a comparison between the sounds of Mandarin and of English and contains a discussion of pronunciation errors of Chinese learners of English. Much of what will be discussed in the following paragraphs has been taken from Zhao (1995); it should be borne in mind that her comments, too, relate to the sounds of British English, specifically RP. This is not a great concern as long as we are dealing with the consonants, since these do not differ very much between British and American English. It is a major concern when dealing with the vowel system.

The Chinese sound system that Zhao (1995) uses as her reference is that of Mandarin (also called Putonghua or Common Speech), which is comparable in status to RP in England. Like RP in English, there is also a standard form of pronunciation in modern Chinese. This pronunciation, which is being popularized throughout the P.R. China, is based on the northern dialect family, with Beijing speech sounds as the norm. In China, TV and radio announcers use the Common Speech. Teachers and students in school are required to use it, too. It is the main language spoken in China and one of the world's major languages, ranking among the official working languages at the United Nations and other international organizations.

According to Zhao (1995), experience shows that Chinese learners of English who speak the Common Speech have fewer difficulties in acquiring a good English pronunciation than those who speak with broad local accents, who often have many difficulties to overcome before they can pronounce English acceptably. This would be because there are more similarities between the pronunciation of Common Speech and that of English. These claims seem rather speculative and remain to be tested in future research; such testing is clearly beyond the scope of the present dissertation.

Figure 3.5 has been copied from Zhao (1995). It is a traditional cardinal vowel chart with the RP-English vowels drawn as solid black circles and the Chinese vowels as open circles.

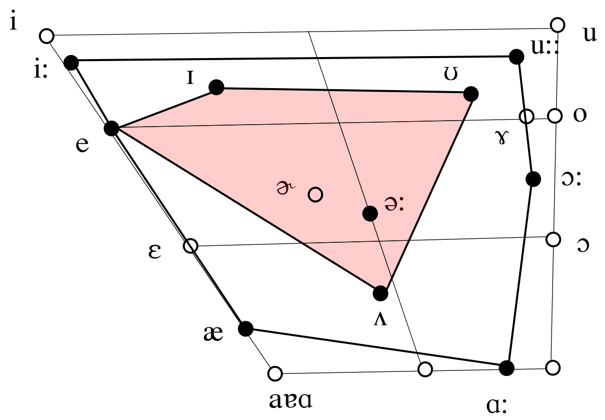

Figure 3.5. Comparison of RP English (solid markers) and Mandarin (open markers) vowels in a traditional Cardinal Vowel diagram (after Zhao, 1995).

For the purpose of the present dissertation Figure 3.5 has to be interpreted with some caution, as we will use General American as the pronunciation norm for English. As will be shown in more detail in Chapter five, the vowel system of English, whether British or American, can be conceived of as two subsystems, one of which is peripheral, with so-called tense (and long) vowels along the outer edge of the vowel diagram, and the other is more centralized, with four vowels configured along an inner circle. Zhao (1995) does not treat tense /e:/ and /o:/ as monophthongs. Rather, she deals with these vowels as half diphthongs, which is why they have not been included in Figure 3.5. The positions of the RP vowels seem quite reasonable; However, I would question the locations of the Chinese vowel sounds. Zhao seems to suggest that Chinese /i, a, ɑ, u/ are identical to cardinal vowels 1, 4, 5 and 8, respectively. It would seem rather unlikely that a language such as Mandarin,with a smaller vowel inventory than English, would have its vowels in more peripheral positions.[8] The high central vowel /y/, which is one of the vowels of Chinese, has been omitted from the chart, most likely as Zhao believes that this vowel is never a reasonable substitute for any vowel of English. The half-close unrounded vowel /ɤ/ is given in the figure; Zhao claims that it is used as a substitute for English /ə:/.

I will now present a list of vowel pronunciation errors as identified by Zhao (1995). Later, in Chapter six, we will have occasion to check the predictions in this list with the confusion data collected in our own experiments. We will then be able to either confirm or disconfirm whether such errors do indeed occur. Moreover, we will examine our data to see whether there are any systematic errors that were not predicted by Zhao. If such errors should be found, the added value our experimental approach would be shown: we predict that even a trained teacher of English as a foreign language may well miss systematic pronunciation errors in foreign-accented English (especially when the teacher is a native speaker of the same language as that of the learners), that can only be brought to light through experimental methods.

The following table is a summary of Zhao's treatment of the English vowel sounds by Chinese learners. It lists all the vowel phonemes of (RP) English in the left-hand column. In the right-hand column I first specify if the particular sound has no counterpart in Chinese. When no remark is made as to the absence of the vowel in Chinese, Zhao implicitly claims that there is some vowel in Chinese that Chinese learners of English will use as a reasonable substitute for the target sound in English. Sometimes the substitute is a good match for the target sounds, in which case no further comments are made. Most of the time, however, the substitute differs from the target; the table will then specify how the substitute differs, and what perceptual confusions are likely to arise as a result of the substitution. The perceptual consequences are sometimes explicitly mentioned by Zhao, but when she makes no explicit claims, I have derived the predictions myself.

Zhao (1995) makes two claims with respect to the diphthongs of English. The first is that Chinese learners tend to reduce the contrast between long and short vowels in English, which would follow from the fact that Chinese does not use length as a distinctive feature. She then goes on to say that diphthongs are like long vowels, implying that Chinese-accented diphthongs will be too short. Chinese has both falling and rising diphthongs. A falling diphthong has its most prominent element first and the less prominent (semivowel, glide) element last, while a rising diphthong has the more prominent element last. English has falling diphthongs only.[9] Zhao adds a warning that there are rising diphthongs in Chinese and the beginning of these rising diphthongs is less prominent than the end. She seems to

---

[8]At first glance one would be tempted to believe that the open circles in figure 3.5 are in fact the cardinal vowel positions, given as reference points. However, these are explicitly the articulatory positions indicated by Zhao (1995) for the vowels of Mandarin. It is unclear from her description how these positions were determined, nor did she supply any references.

[9] It would be possible, however, to analyze the realization of tense /u:/ as [ju:] after certain consonants as a rising diphthong. Examples would be: *puke* [pju:k], beauty [bju:ti], *mew* [mju:], *tune* [tju:n], *dune* [dju:n], *new* [nju:], *cue* [kju:], and many others. Since the glide [j] *only* occurs in combination with tense /u:/ (including its centring diphthong allophone [ʊə], there is no point in increasing the set of onset clusters with a large number of /Cj/ sequences.

imply, therefore, that Chinese learners tend to substitute rising diphthongs for English target diphthongs. However, given the large inventory of falling diphthongs in Chinese we do not think it very likely that Chinese learners of English will ever use a rising diphthong as an approximation to an English target – except perhaps for /Cju:/ (also see note 2), but then the substitution would be highly felicitous.

Table 3.6. Survey of pronunciation problems with vowels by Chinese learners of English, derived from Zhao (1995). M = Mandarin, E = English.

English target	Mandarin substitutions/typical errors
/i:/	M /i/, too short, not tense enough, not high enough; confusion with E /ɪ/
/ɪ/	absent in M, M /i/ substituted, too long, too tense, too high, confusion with E /i:/
/e/	Generally no problem, but Northern speakers may substitute [ai] or [ei], yielding confusion with E /ai/ and with E tense /e:/
/æ/	Absent in M, pronunciation will be too close, confusion with E /e:/
/ɑ:/	M has three allophones: [a] (open syll.), [ɐ] (closed syll.) and [ɑ] (before nasal coda). Realization not open enough, confusion with E /ʌ/.
/ɔ/	Sound does not exist in M. Diphthong [au] substituted with glide and not enough lip-rounding. Confusion with E /au/
/ɔ:/	M [o)] substituted, too open but quite similar to modern (closer) British E pronunciation for /ɔ:/
/ʊ/	M [u] substituted. Too long, confusion with E /u:/
/u:/	M [u] substituted. Too short, confusion with E /ʊ/
/ʌ/	Sound does not exist in M. M [ɐ] substituted. Too open, confusion with E /ɑ:/
/ə:/	[ɤ] and [o] substituted. Too short, too close, too backward, confusion with /ɔ/ and/or /ɔ:/ The central vowel /ɚ/ is actually a retroflex [ə]; this sound would be quite similar to the American E realization of /ə:/
/ə/	[ɤ] and [o] substituted. Too long, too close, too backward, confusion with /ɔ/ and/or /ɔ:/

3.2 Consonants

Consonants are made by causing a complete or partial obstruction in the mouth or pharynx, and are usually described in terms of where the obstruction is made in the mouth (or: place of articulation), how the sound is made (or: manner of articulation), and whether or not the vocal cords vibrate (or: voicing). Consonants, therefore, all differ from each other in at least one of these ways.

In terms of the size of the inventories, Chinese has the largest variety with 26 different onset consonants, but only two of these may occur in the coda (while some consonants can be considered variants of each other). English has a slightly less rich inventory of 24 consonants including three that are exclusively found in the coda. Dutch has the smallest inventory with 21 consonants, nine of which cannot occur in the coda.

3.2.1 Dutch consonants vs. English consonants

The classification of consonants involves at least three factors, the state of glottis, the place of articulation, and manner of articulation. The two charts 'Dutch vs. English' (Table 3.7) and 'Chinese vs. English' (Table 3.8) include all the consonant symbols in English, Dutch and Chinese. The horizontal axis shows the various places of articulation, the vertical axis the various manners of articulation, while the voiceless consonants are distinguished from voiced ones by placing the former on the left in any box and the latter on the right. Consonants that occur both in the Dutch and in the English inventory are in white cells. Dutch sounds in grey-shaded cells are absent in English, English (target) sounds in black cells are absent in Dutch. Grey cells in the English panel contain target sounds that occur also in Dutch but which have different phonetic realizations. These would be transcribed with the same broad phonemic symbol but differ from their Dutch counterparts in phonetic detail, i.e. in diacritic marks. These 'similar sounds', as they would be classified by Flege, are indicated in the bottom panel against a gray background. Here we simply count 24 consonants in the English inventory, six of which do not occur in the Dutch inventory and seven of which differ in phonetic detail from their Dutch counterparts. Specific predictions of learning problems will be discussed later.

Table 3.7. Consonant of Dutch (upper panel) versus English (lower panel) in a manner (down) by place (across) table. Further see text.

Manner	Place of Articulation							
	labial	labial-dental	dental	Alveolar	Alveolar-palatal	(retroflex)	palatal	velar	glottal
Source: Dutch									
Stop	p b			t d				k	
Nasal	m			n				ŋ	
Fricative		f v		s z	ʃ ʒ			χ ɣ	
Affricate									
Approx.		ʋ		r			j		ɦ
Lateral				l					
Target: English									
Stop	p^h b̥			t^h d̥				k^h g̊	
Nasal	m			n				ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ				
Affricate					tʃ dʒ				
Approx.	w			ɹ			j		h
Lateral				l					

3.2.2 Chinese consonants vs. English consonants

Table 3.8 presents a contrastive listing of the consonants of Chinese and English arranged by manner (down) and place (across). When a table cell contains two sounds, the one on the left represents the fortis (aspirated) and the one on the right the lenis (unaspirated voiceless) member of a pair of obstruents. Grey cells in the Chinese panel denote sounds that do not occur in English, black cells in the bottom panel represent sounds that occur in English but are absent in Chinese.

Table 3.8. Consonant sounds of Chinese (upper panel) versus English (lower panel) in a manner (down) by place (across) table. Further see text.

Manner	Place of Articulation								
	labial	labial-dental	dental	Alveolar	Alveolar-palatal	(retroflex)	palatal	velar	glottal
Source: Chinese									
Stop	p^h b̥			t^h d̥				k^h g̊	
Nasal	m			n				ŋ	
Fricative		f		s		ʂ	ɕ		
Affricate				ts^h ts		$tʂ^h$ tʂ	$tɕ^h$ tɕ		
Approx.	w				ɽ		j	χ	
Lateral				l					
Target: English									
Stop	p^h b̥			t^h d̥				k^h g̊	
Nasal	m			n				ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ				
Affricate					tʃ dʒ				
Approx.	w			ɹ			j		h
Lateral				l					

The table reveals that of the 24 English consonants ten do not occur in Chinese; however, the remaining 14 should be quite similar to their Chinese counterparts.

3.2.3 Prediction of pronunciation problems in consonants

As we did for vowels, we will now present tables containing the most likely errors in the production and perception of English consonants by Chinese and Dutch learners. The consonant data are largely based on Zhao (1995) for Chinese learners of English; the Dutch data are based on Collins and Mees (1981) and Gussenhoven and Broeders (1976). Again, since both textbooks deal with pronunciation difficulties of British English (RP) sounds we adapted some of the claims so as to be applicable for American English.

3.2.3.1 Dutch-English consonant transfer

Table 3.9 presents a summary of remarks and observations made by Collins and Mees (1981) on differences between the Dutch and English consonants. A pervading problem in the pronunciation of English consonants by Dutch learners is that Dutch does not allow voiced (lenis) obstruents in coda positions; in such positions the voiced ~ voiceless (or lenis ~ fortis) opposition is neutralized, and only the voiceless (fortis) member of the pair will be realized. We will not discuss this problem in the table below; rather we consider this a consequence of a rule difference between Dutch and English depending on a sound's position in the syllable; the matter will therefore be discussed in § 3.3.

Table 3.9 Survey of pronunciation problems with consonants by Dutch learners of English, derived from Collins and Mees (1981) and Gussenhoven and Broeders (1976: 142−143). D = Dutch, E = English.

English target	Dutch substitutions/typical errors
/f/	D /f/ is identical to E /f/.
/v/	E /v/ has less friction than its D counterpart. Most D speakers substitute D /v/ for E / v/; these are similar sounds. [10]
/θ, ð/	Both are absent in D. These two sounds pose major problems of recognition and articulation for the learner. /ð/ is far harder for D learners than /θ/. Replacement of /ð/ by /d/ is one of the most common and persistent D errors. /θ/ is easier for D learners; the traditional instruction of tongue between teeth obtains the slit tongue shape characteristic of /θ/, which distinguishes it from /s/.
/s, z/	The articulation of /s, z/ is different from that of E. D /s, z/ are typically articulated with a portion of the tongue between front and blade whilst the tip is kept down behind the front teeth. With some speakers there may also be some lip-rounding. D /s/ has less firmly held stricture than E /s/; the jaw is more open with a laxer articulation. As a result, the friction of D /s/ is *graver* than the *sharp* friction which characterises the English sound. Some D speakers produce a D /s/ which is acceptable if transferred into E, whilst others produce a sound which is between E /s/ and E /ʃ/. Some of the D accents lack a contrast /s ~ z/. Other accents have no contrast /s ~ sj/ and /z ~ zj/.
/ʃ, ʒ/	The D sequence /sj/ in *chef* has more obvious palatal off-glide than its E counterpart. The articulation is often unrounded; the effect of this is to make D /sj/ sharper in friction.
/h/	E /h/ tends to have somewhat stronger glottal friction than D /h/, and voiceless *pharyngeal* friction can be heard from some speakers. E /h/ is only voiced between some voiced sounds, whereas D /h/ tend to have breathy voice in all contexts. Breathy (voiced) /h/ does not compromise its identity in E.

[10] The Dutch labio-dental semivowel /ʋ/ would be a better substitute but Dutch speakers do not do this.

Table 3.9. Continued.

/m, n, ŋ/	These are identical in D and E. One difficulty likely to arise is excessive nasalization of preceding vowels (plus deletion of the conditioning nasal). This is especially noticeable in open vowels. Nasal release of /t, d/ may provide problems for D students, particularly into syllabic consonants. D learners tend to insert /ə/ between stop and nasal, e.g. *rotten* /ˈrɔtən/.
/l/	The distribution of clear [l] and dark [ɫ] is similar in D and in E, though for many D speakers intervocalic /l/ is dark. Many D accents (Rotterdam and Amsterdam) have dark [ɫ] in all contexts including initial position. Articulation of clear [l] is similar in D and in E. Dutch /l/ is not devoiced following fortis plosives, compare E *plan* [pl̥æn] and D *plan* [plɑn]. D dark [ɫ] is significantly different from E dark [ɫ]. No perceptual confusions will arise from the differences.
/j, w/	D /j/ is similar to the E sound, but is often realized with friction, thus giving a voiced palatal fricative [j̥]. Because of the similarities of E /j/ and D /j/ there are few significant problems for the learner.
/w/	For E bilabial /w/ the typical substitution is D labio-dental /ʋ/. /w/ presents a major problem for Dutch learners both in terms of articulation and in confusion of E /w ~ v/ contrast.
/r/	D onset /r/ is either an alveolar or uvular trill (or fricative in clusters). E /r/ is a retroflex approximant. D coda /r/ may also be an approximant. Although substitution of trill and fricative may sound foreign, no perceptual confusion will arise.
/p, b/ /t, d/ /k/	D /p, t, k/ have very short VOT and are not aspirated. These realisations are substituted for E /p^h, t^h, k^h/, and may be confused with E /b̥, d̥, g̊/. D lenis stops /b, d, (g)/ have negative VOT (prevoicing); no perceptual confusion should arise when these sounds are substituted for their E counterparts.
/g/	Absent in D; D /g/ occurs mostly in loanwords or as an allophone of /k/. It is not available as a substitute for E /g/. D /k/ may be substituted, even in the onset. Perceptual confusion with /k/ is expected.
/tʃ, dʒ/	Absent in D; these affricates are either replaced by/confused with the fricatives /ʃ, ʒ/ or by some sequence of /t(s)j/, /d(z)j

3.2.3.2 Chinese – English consonant transfer

The following Table 3.10 summarizes the typical errors and substitution patterns observed for English consonants spoken by Chinese learners of English. Again our main source of information is Zhao (1995).

Table 3.10. Survey of pronunciation problems with simplex consonants by Chinese learners of English, derived from Zhao (1995). M = Mandarin, E = English.

English target	Mandarin substitutions/typical errors
/b, d, g/	E voiced (lax) plosives have 0 VOT (and therefore have no voice lead nor voice lag). This is as in M; no problems are predicted.
/p, t, k/	Voiceless (tense) plosives are aspirated both in M and in E but more strongly in E. Confusion with E /b, d, g/ may result as a result of insufficient aspiration
/f, v/	/v/ is absent in M, /f/ exists. /f/ is not a problematic target sound but /w/ and /f/ are substituted for /v/ (the latter especially in the coda)
/θ, ð/	Both are absent in M. /t/, /s/ and /f/ are substituted for /θ/, and /d/, /dz/ and /v/ for /ð/
/s/	/s/ in M is articulated with the tongue blade against the back of the upper teeth, in E with blade against alveolar. Substitution is either unnoticed or confusion with /θ/ arises
/z/	/z/ does not exist in M; the unaspirated voiced affricate /dz/ is substituted, which may be confused with E /dʒ/ or even with /tʃ/
/ʃ, ʒ/	These fricatives do not exist in M. No substitutes are given; no confusions are predicted.
/tʃ, dʒ/	/tʃ/ is approximated by M [ts^h] and /dʒ/ by [ts]. No specific confusions are predicted.
/w/	M and E /w/ are quite similar. M /w/ is in free variation with /v/. As a result /w/ is often incorrectly replaced by /v/ (and vice versa). /w/ ~ /v/ confusion is predicted.
/j/	M /j/ is similar to E. No problems predicted
/h/	M no /h/; the uvular fricative [χ] is substituted. This will not lead to confusions but the substitution will be unacceptable.
/l/	The clear /l/ is exactly the same as the M lateral /l/ The dark /ɫ/ is a more difficult sound for M learners, because in M, the lateral consonant never occurs in the coda
/m, n, ŋ/	English /m/, /n/, /ŋ/ are quite similar to M /m/, /n/, /ŋ/. However, M /m/ and /n/ never appear in the coda; M learners tend to pronounce the last phoneme unclearly, or even omit it unintentionally.[11] Word-medial /ŋ/ is claimed to be difficult for M learners.
/r/	E onset /r/ is replaced by the M fricative /ʐ/ which is quite similar to the target but has slight friction; confusion with E /z/ is predicted.[12] No problems are predicted with coda-/r/; here the Chinese retroflex vowel is an adequate substitute.

[11] Some Chinese learners, especially people from Hunan, Sichuan, Fujian and Anhui provinces, may replace /n/ with /l/ or /l/ with /n/, as these sounds are free variants in the local dialects.

[12] One common error among Southern Chinese learners of English is the confusion of /r/ with /l/ and also with /n/. They produce *right* /rait/ as *light* /lait/ or *night* /nait/. Since this is not a problem for Northern Chinese (Mandarin) speakers, we have not included this confusion in the table.

3.3 Syllable structure

Human speech is basically spoken as a sequence of opening gestures of the mouth. Of course, once the mouth has been opened, it has to be closed before it can be opened a second time. One cycle of opening and closing the mouth produces a phonetic syllable. The alternation of opening and closing gestures takes place at a rate of some five cycles per second. When the mouth is maximally open, vowel sounds are produced; when the mouth is completely or partially closed, consonants are produced. The segments (consonants and vowels) within a syllable are subject to the sonority principle: louder and more sonorous sounds are produced in the middle of the syllable when the mouth is maximally open, and sounds of decreasing sonority are produced as they are closer to the edges of the syllable.

Languages differ widely in the complexity of syllable structures they allow. The simplest type of syllable structure is a regular alternation of a single consonant (C) and a single vowel (V). Many languages only allow regular CVCV alternation and in all languages CV is the most frequent syllable type. Mandarin comes rather close to such a CV language. English and Dutch have a richer variety of syllable types, and they allow up to three consonants in sequence in the beginning of a syllable and up to four in the final part of the syllable. Many consonants have rather different pronunciations depending on whether they precede the vowel or follow it within the syllable. Research has indicated that positive transfer of consonants is limited to source and target segments that have the same position in the syllable (Lado, 1957; Flege, 1995). Also, speakers of a language that has a simple CV structure find it difficult to produce sequences of consonants that are not interspersed with vowels. It is therefore important to review some of the differences in syllable structure among the three languages under consideration.

3.3.1 English

English is a language that allows complex syllable structures. Syllables are split up in an onset and a rhyme portion; the rhyme is further subdivided into the vocalic nucleus and the coda, which contains all postvocalic consonants. Onsets in English may vary in length from zero to three consonants. If the onset has its maximal length, i.e. three segments, the very first segment must always be /s/. Given this severe restriction the /s/ is considered to be outside the onset and given special appendix status. The vocalic nucleus either contains a long (or tense) vowel or a short (or lax) vowel. A word (or syllable) may not end in a lax vowel; lax vowels have to be followed by at least one coda consonant. Tense vowels may occur at the end of a word (or syllable). Given that diphthongs may occur at the end of a word, it follows that a diphthong functions as a tense vowel in English. The maximal number of consonants that can follow the vowel is three if the vowel is lax and two if the vowel is tense. In maximally long coda strings the last consonants are restricted to {t, d, s, z}, on the grounds of which this final constituent has been given appendix status. These can only occur as realisations of some suffix (past tense, past participle, plural, third person singular, as in *milked*, *ranged*, *milks*, *fields*). The velar nasal takes up the position of two coda consonants. Semivowels (glides) /j, w and h/ cannot occur in the coda; they are restricted to the onset. Voiced (lenis) and voiceless (fortis) obstruents may occur in the onset and in the coda.

3.3.2 Dutch

The syllable structure of Dutch, which is closely related to English, has much in common with the English system. The syllable is hierarchically subdivided in much the same way. Dutch has zero to three consonants in the onset, with special appendix status for initial /s/. The vocalic nucleus contains either a short/lax vowel or a long/tense vowel, which again is functionally equivalent to a diphthong. Lax vowels may not occur at the end of a word or syllable; they have to be followed within the rhyme by at least one coda consonant. The maximum number of coda consonants is four, which can only occur after a lax vowel and then contains appendix consonants {s, t, st, ts} as in *herfst* /hɛrfst/ 'autumn'. The velar nasal counts as two consonants; /h/ cannot occur in the coda. However, other than in English, semivowels /w, j/ may occur in the coda but only after a long/tense vowel, as in *haai* /ha:j/ 'shark', *geeuw* /ɣe:w/ 'yawn'. Voiced as well as voiceless obstruents occur in the onset; in coda position voiced obstruents are impossible; these are neutralized to their voiceless counterparts.

Coda clusters are often broken up in Dutch by the insertion of an epenthetic vowel schwa. The insertion typically takes place when two adjacent consonants in the code do not differ enough in sonority, as in *melk* > [mɛlək], herfst > [hɛrəfst]. No vowel epenthesis takes place before obstruents which may occur in the appendix (i.e. /s/ and /t/) (see e.g. Van der Hulst, 1984).

3.3.3 Chinese

Traditional Chinese phonology divides the syllable into an Initial and Final. The Initial is the way a syllable begins, usually with a consonant. The Final is the syllable minus the Initial. For example, in *ta*, *chi*, *jin*, *chuang*, the Finals are *a*, *i*, *in*, and *uang*, respectively. The longest form of a Final consists of three parts: a medial (or: semivowel), a main vowel (or: head vowel), and an ending (or, in the case of retroflex suffixes, sometimes two endings, as in the *er*-sound *ming'er* 'tomorrow').

A Final in Mandarin comprises one of four medials: ∅ (empty), /i/, /u/, or /iu/ (= [y], one of three vowels: /a/, /e/, or /o/, and one of six endings: ∅, -i, -u, -n, -ŋ, and [ɻ] (phonetically -r).[13] Actually, there are only 40 different Finals (if Finals involving retroflex suffixes are not counted). As a result of these very severe restrictions on possible syllables in Mandarin, no obstruent clusters are possible in the onset (Initial) nor in the coda (Final). Onset clusters can maximally have a length of two segments, in which case the consonant closest to the vocalic nucleus must be a semivowel. Coda clusters are disallowed; in fact, syllables are generally open, i.e. end with a vowel. The only possible coda consonants are the nasals /n/ and /ŋ/. In compound vowels with /a, e, o/ as the first segment and /i, u/ as the second element, the latter are phonetically realised as semivowels, creating a diphthong. Phonetically,

[13] This gives rise to 4 × 2 × 6 = 48 possible Finals, since *a* and *o* count as allophones of one phoneme.

the retroflex approximant [ɻ] could also be considered a coda but this sound functions as a vowel.

3.3.4 Dutch versus English syllable structures

Generally, the syllable structures of Dutch and English are highly similar; in fact, Dutch syllable structure seems even less constrained than English, given that Dutch allow onsets such as /kn, pn, ɣn/ (in written obstruent+/n/ clusters the obstruent is not pronounced in English). As a result, Dutch speakers of English are expected to have few problems in realizing the complex syllable structures of English.

Complex clusters are no problems as such. However, due to some language-specific restrictions and peculiarities of Dutch some interference phenomena may arise. A very serious difficulty for Dutch speakers of English is to maintain the fortis ~ lenis (voiceless ~ voiced) contrast in coda obstruents. Lax/voiced coda obstruents are consistently realised as their fortis/voiceless counterparts, which may lead to perceptual confusion in English in minimal pairs such as *bad* ~ *bat*, *lies* ~ *lice*, *ridge* ~ *rich*, *leave* ~ *leaf*, *mouth* (verb) ~ *mouth* (noun), and many more.

Dutch speakers have a predictable tendency to break up English coda clusters, using their epenthetic vowel rule. Although the pronunciation of *milk* as [mɪɫək] sounds foreign, intelligibility will not be compromised by the epenthetic vowel.

3.3.5 Chinese versus English syllable structures

Since Mandarin allows no onset clusters except C+glide, Chinese speakers of English are predicted to have problems with the pronunciation of all other CC and CCC clusters of English. They are expected to break up awkward clusters by inserting an epenthetic vowel. Examples given by Zhao (1995: 95) indicate that /ə/ is inserted in between the members of CC clusters (*spy* > /səpai/, *pray* > [pəre^{i}]. No examples are given of pronunciation problems involving CCC onset clusters.

Even more problems are expected in the realisation of English coda clusters. Given that Mandarin only allows /n/ and /ŋ/ in the coda, any other consonant in that position will be awkward. Problems will increase when the coda contains two or more consonants. Chinese learners of English employ two strategies to cope with coda consonants. One is to add an epenthetic vowel [ə] after the coda consonant, which is then resyllabified to the onset of a separate syllable; this is what often happens in single C codas. When the coda is a cluster, it is often simplified by deleting one of the members of the cluster (after which epenthesis and re-syllabification may take place). Given the absence of obstruents in Mandarin codas and the absence of coda clusters, it is an open question how Chinese learners of English will deal with the fortis ~ lenis opposition in English codas. English is one of a minority of languages that maintains this contrast in coda position; in the majority of the world's languages the contrast is neutralised and only the voiceless member surfaces. One would predict that the realisation of marked phenomena in the target language (English) are a learning problem when these phenomena are absent in the source language (Mandarin). This prediction follows from the Markedness Differential Hypothesis (MDH, Eckman, 1977).

Zhao mentions one special strategy whereby Chinese learners arguably substitute Mandarin onset affricates [ts^h] and [ts] for English coda clusters /ts/ and /dz/, respectively. Since the place of articulation of the Mandarin affricates (tip of the tongue against the back of the upper teeth) is not the same as that of the English targets (tongue blade and the teeth ridge), this strategy will only be partially successful.

3.4 Concluding remarks

In this chapter we have reviewed the extensive literature on differences in sound structures between Chinese, Dutch and English. Some of the literature, especially that relating to the acoustical properties of vowels, was experimental in nature. The vast majority of the sources consulted, however, is based on observations made by teachers of English as a foreign/second language or by linguistic phoneticians using observation unaided by instrumental analysis. We will not be able to test each individual observation against experimental data to be collected in the next chapter(s). However, Chapter three will provide a database of observations we may turn to when discussing our experimental results. Very often we will point out correspondences between observations made in Chapter three and later experimental results, and on a few occasions we will also discuss experimental findings that have gone unnoticed in the (pedagogical) literature.