Chapter Five

Acoustic analysis of vowels[1]

5.1. Introduction

In this chapter we will provide an acoustical analysis of the vowel tokens produced by the 20 Chinese, 20 Dutch and 20 American speakers of English, which were recorded in the course of the project. A description of the materials and the method of data collection were given in Chapter four.

As was explained in Chapter three, the vowel systems of (Mandarin) Chinese, Dutch and American English differ considerably, both in the number of vowels in the inventory and in the details of their position within the articulatory vowel space, and possibly also in terms of their durational characteristics. Although the phonetic differences are typically described in terms of articulatory properties, we have not tried to determine articulatory properties if the vowels through physiological measurements – as we had no recourse to the type of equipment needed, such as X-ray photography, Magnetic Resonance Imaging (MRI) or Electromagnetic Mid-sagittal Articulography (EMMA). Rather we used acoustic measurements that are known to have rather clear correspondences with articulatory properties of vowels. How this is done will be explained briefly in the next section.

5.1.1. Objective measurement of vowel quality

There is agreement among experimental phoneticians that vowel quality can be quantified with adequate precision and validity by measuring the center frequencies of the lower resonances in the acoustic signal. Specifically, the center frequency of the lowest resonance of the vocal tract, called first formant frequency or F1, corresponds closely to the articulatory and/or perceptual dimension of vowel height (high vs. low vowels, or close vs. open vowels). For an average male voice, the F1 values ranges between 200 hertz (Hz) for a high vowel /i/ to some 800 Hz for a low vowel /a/. The second formant frequency (or F2) reflects the place of maximal constriction during the production of the vowel, i.e., the front vs. back dimension, such that the F2 values range from roughly 2200 Hz for front /i/ down to some 600 Hz for back /u/. For female voices the formant frequencies are some 10 to 15% higher, on account of the fact that the resonance cavities in the female vocal tract are smaller (shorter) by 10 to 15% than those of a male speaker.

---

[1] This chapter is a slightly adapted version of H. Wang and V. J. Van Heuven (2006) Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. In J. M. Van de Weijer and B. Los (eds.) *Linguistics in the Netherlands 2006*. Amsterdam/ Philadelphia: John Benjamins, 237–248.

The relationship between the formant frequencies and the corresponding perceived vowel quality is not linear. For instance, a change in F1 from 200 to 300 Hz brings about a much larger change in perceived vowel quality (height) than a numerically equal change from 700 to 800 Hz. Over the past decades experimental phoneticians and psycho-physicists have developed an empirical formula that adequately maps the differences in hertz-values onto the perceptual vowel quality domain, using the so-called Bark transformation (for a summary of positions see Hayward, 2000). Using this transformation, the perceptual distance between any two vowel qualities can be computed from acoustic measurements.

We used the Bark formula advocated by Traunmüller (1990):

$$\text{Bark} = [(26.81 \times F) / (1960 + F)] - 0.53,$$

where F represents the measured formant frequency in hertz.

For many languages formant measurements have been published, so that an adequate determination can be made of the vowel systems of those languages. Probably the best known set of formant measurements was produced for American English, in the early fifties by Peterson and Barney (1952) for male and for female speakers separately (see also Chapter three). These authors used the same stimuli that we used, i.e. vowels embedded in a /hVd/ consonant frame. Similar vowel sets were recorded for 50 male and 25 female speakers of Dutch by Pols and co-workers in the seventies (Pols, Van de Kamp and Plomp, 1973 and Van Nierop, Pols and Plomp, 1973, respectively). Formant measurements for the vowels of Mandarin (Beijing dialect) became available only recently (Li, Yu, Chen and Wang, 2004).

Formant measurements for Chinese-accented English (aiming at the American pronunciation norm) were published by Chen, Robb, Gilbert and Lerman (2001). The authors recorded a subset of the American English vowels (eleven monophthongs) in the same /h_d/ monosyllables that we used ourselves. However, their speakers (20 male and 20 female adults) had been living in the USA for at least two years after having received intensive exposure to spoken English in China in order to qualify for the TOEFL test required to enter a university in the USA. This is clearly a different type of ESL speaker than we target in our study, so that it makes every sense that we should measure the formants in our speaker group separately. We would predict, of course, that certain vowels that are acoustically indistinct in our dataset will be more clearly differentiated in Chen et al.'s (2001) data but not so clearly as when spoken by American native speakers. Moreover, no data are available in Chen et al. (2001) on the perception of the ESL tokens; so that it is unclear to what extent the vowels produced by their advanced Chinese learners of English were correctly identified by either Chinese or American listeners.

No formant data have ever been published for Dutch-accented English vowels. However, several studies have been done on the perceptual mapping of English vowels by Dutch ESL speakers. In such studies, a large number of vowel tokens were generated by speech synthesis covering the acoustical vowel space according to a finely-meshed grid. Listeners, whether native or foreign, were then instructed to indicate for each artificial vowel sound which of the vowels in the target language would be most compatible with it (often with a goodness or typicality rating). The responses allow the researcher to reconstruct the perceptual vowel space of the listener in terms of the prototypical vowel exemplar for each perceptual category and some area of tolerance around the prototype, where more or less acceptable tokens of the category may occur. Unfortunately, the Dutch ESL listeners were all university students of English (Schouten, 1975) or Dutch-English bilinguals (Broerse, 1997), and therefore cannot provide a basis for comparison for our study. [2] There is no alternative, then, but to measure the acoustical properties of Dutch-accented English vowels ourselves, using the data we collected in the present study.

5.1.2. The problem of vowel normalization

Unfortunately, formant values measured for the same vowel differ when the vowels are produced by different individuals. The larger the differences between two speakers in shape and size of the cavities in their vocal tracts, the larger the differences in formant values of perceptually identical vowel tokens are. Given that the vocal tracts of women are some 15 percent smaller than those of men, comparison of formant values is especially hazardous across speakers of different sex. Numerous attempts have been made, therefore, to factor out the speaker-individual component from the raw formant values such that phonetically identical vowels spoken by different individuals would come out with the same values. None of these vowel normalization procedures have proven fully satisfactory (Adank, Van Heuven and Van Hout, 1999; Labov, 2001: 157-164; Rietveld and Van Heuven, 2001). Broadly, two approaches to the normalization problem have been taken in the literature (for a detailed discussion of the issue of vowel normalization, see also Nearey, 1989). The first approach, called intrinsic normalization, tries to solve the problem by considering only information that is contained in the single vowel token under consideration, typically by computing ratios between pairs of formant values such as F1/F0 and F2/F1.[3] The alternative, extrinsic normalization, looks at tokens of all the vowels in the phoneme inventory of a speaker and expresses the position of one vowel token relative to the other tokens within the individual speaker's vowel space.

In the present study we have opted for a straightforward extrinsic vowel normalization procedure, first used by Lobanov (1971), which is simply a z-normalization of the F1 and F2 frequencies over the vowel set produced by each individual speaker. In a z-normalization, the F1 and F2 values are transformed to z-scores by subtracting the individual speaker's mean F1 and mean F2 from the raw formant values, and then dividing the difference by the speaker's standard deviation. Z-transformed F1 values less then 0 then correspond to relatively close (high) vowels, values larger than 1 refer to rather open vowels. Similarly, negative z-scores for F2 refer to front vowels, whilst positive z-scores for F2 represent back vowels. In our case we applied the Lobanov normalization after first transforming the hertz values to Bark values.

---

[2] Also, in Broerse's study only the perceptual norms were determined for the checked (short, lax) vowels in the inventory.

[3] When formant values are rescaled to Bark, the numerical difference (F1–F0; F2–F1, etc.) is preferred over the ratio.

5.1.3 Vowel duration

The vowels of English and Dutch can be divided into two major groups on the basis of their phonological behavior, which largely correspond with phonetically short (and lax) versus long (and tense) vowels (for details, see Chapter three). Typically, the short/lax and long/tense vowels are in paired oppositions. In English, examples of such pairs are /iː ~ ɪ/ and /uː ~ ʊ/. Since vowel duration plays an important role in marking the contrast, next to vowel quality differences, we also measured the vowel duration in the tokens recorded in our dataset. Since some speakers speak faster than others, raw vowel duration cannot be used in the comparison. Rather, durations should be normalized within speakers. Here, too, we used a simple z-normalization procedure (see above) so that negative normalized durations refer to relatively short vowel tokens, and positive values represent relatively long vowel durations.

Chinese does not use exploit length as a vowel feature at the phonological level. We would predict (see Chapter three) that Chinese ESL speakers will make less difference between the short (lax) and long (tense) vowels of English – whether as subsets in the vowel inventory or in pairwise oppositions – than Dutch ESL speakers, and certainly less clearly so than native speakers of English.

5.1.4 Selecting vowels for analysis

Our recordings contain tokens of 19 vowel types, that is, if the speakers had indeed spoken British English. Given that our speakers, including the Dutch speakers, without having been instructed to do so used an American-style of pronunciation, without r-coloured vowels (so-called murmur diphthongs), there seems little point in measuring the vowels that were followed by /r/. Therefore we eliminated the tokens representing *here'd, haired, hard, hoored* and *heard*. Next, we decided not to include any full diphthongs as these would introduce the complication of having to trace the spectral change over the course of the vowels. This eliminated the types *hide, how'd* and *hoyed*. What remained is precisely the set that was also measured in Chen et al. (2001). We finally decided also to eliminate the /ɔː/ type. It appeared that our speakers did not systematically differentiate between this vowel and /ɔ/. Moreover, quite a few of our L2 speakers pronounced *hawed* as /haud/.

5.2 Formant plots

Using the Praat speech processing software (Boersma and Weenink, 1996) the beginnings and end points of the target vowels were located in oscillographic and/or spectrographic displays. Formant tracks for the lowest four formants, F1 through F4, were then computed using the Burg LPC algorithm implemented in Praat, and visually checked by superimposing the tracks on a wideband spectrogram. Whenever a mismatch between the tracks and the formant band in the spectrogram was detected, the model order of the LPC-analysis was changed *ad hoc* until a proper match was obtained between tracks and spectrogram. Once a satisfactory match was obtained, the values for F1 and F2 were extracted at 25, 50, and 75% of the duration of the target vowel, as well as the vowel duration as such, and stored for off-line statistical processing.

Formant values were then converted to Bark (see § 5.1.1) and averaged over the ten male and ten female speakers in each speaker group separately. These mean F1 and F2 values are plotted in acoustical vowel diagrams in figures 5.1a-f for male and female Chinese, Dutch and American speakers of English. Each plot contains the position of the ten monophthongs selected as explained in § 5.1.4.

The **Chinese** ESL speakers' vowels show tight clustering, and therefore little spectral distinction between intended /i:/ and /ɪ/ (see figure 5.1a-b for the Chinese speakers). This result was predicted from the contrastive analysis of the Chinese and American English vowel systems in Chapter three. The lack of differentiation between the two vowels is very clear for the male speakers; there is some measure of spectral distinction in the Chinese female tokens. Similarly, there is hardly any spectral difference between intended /ɛ/ and /æ/, nor between /u:/ and /ʊ/. The lack of distinction in these two vowel pairs was also predicted by the contrastive analysis.

In spite of the lack of distinctive vowel pairs, we may observe that the Chinese ESL speakers spread their vowels over a large portion of the acoustical vowel space. Although the number of (phonological) vowels in Chinese is relative small (between seven and ten, see Chapter three), this does not prevent Chinese ESL speakers from using a very large vowel space. Probably, this is a consequence of the much larger number of distinct vowel allophones in Chinese, which gives Chinese ESL speaker an advantage. The substitution of context-dependent allophones is not predicted, however, by Flege's Speech Learning Model (Chapter two).

We divided our American English inventory of ten monophthongal vowels into two subsets, corresponding to five tense vowels and five lax vowels. Here, the vowel /ɔ/ is classed as a tense vowel on the grounds that it is a merger of tense /ɔ:/ and lax /ɔ/. Its location in the vowel space (see figure 5.1e-f for the American speakers) motivates this choice quite clearly. Also, we classified the open front vowel /æ/ as tense, though not on phonological or distributional grounds (it would be phonologically lax since it cannot occur at the end of a word, see Chapter three). Phonetically, however, there is good reason to consider American /æ/ a tense vowel: it is clearly longer than all other lax vowels, and is in fact as long as any tense vowel in the system, and it is also peripheral, that is, on the outer edge of the vowel space. This must also have been the (implicit) reason prompting Strange, Bohn, Nishi and Trent (2004) to classify American /æ/ as tense.[4] In figure 5.1 the five tense vowels have been linked with a solid line; the lax vowels have been linked with a dotted line. We observe, in figure 5.1a-b that the tense and lax vowel polygons largely overlap, indicating that the Chinese ESL speakers basically fail to spectrally distinguish between the spectrally more peripheral tense set and the spectrally reduced (centralised) lax set.

The ESL tokens produced by the **Dutch** speakers are generally distributed over a much smaller portion of the vowel space than the Chinese ESL tokens. One reason for the apparently shrunken vowel space in Dutch ESL may be that Dutch speakers

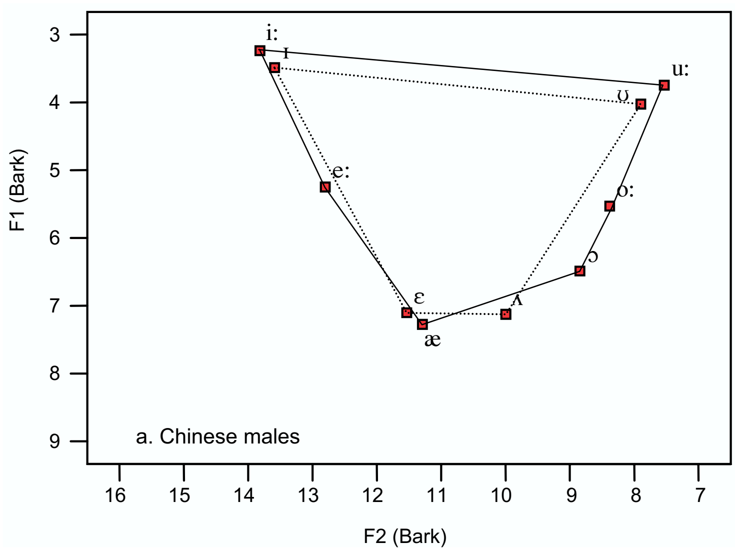
---

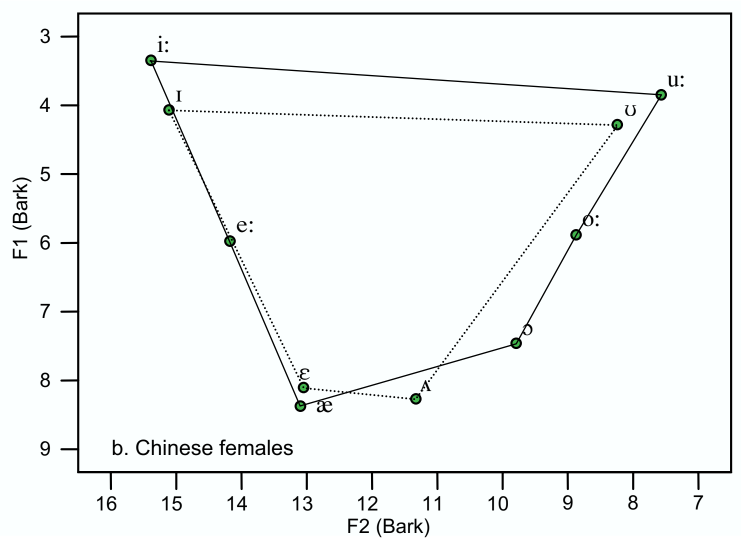
[4] Strange et al. (2004) plot the eleven monophthongal vowels of American English (same as in Chen et al., 2001) recorded by four male speakers in disyllabic /hVba/ frames.

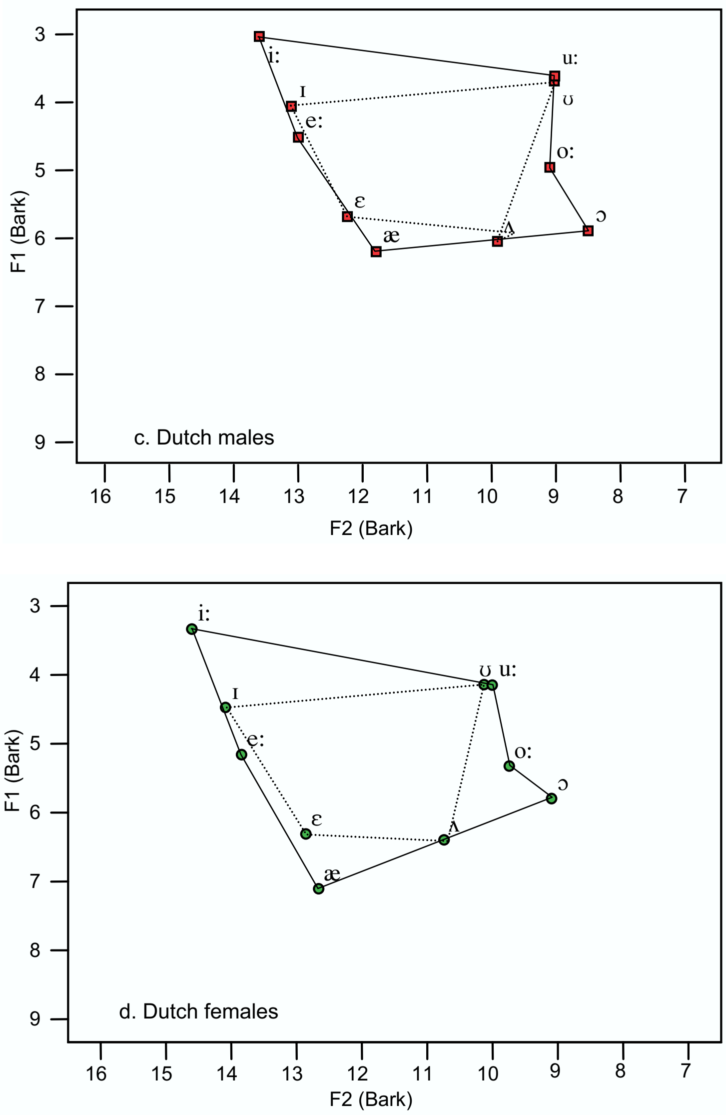
reserve the most open part of their vowel space for the vowel /a:/, as in Dutch *taak* /ta:k/ 'task', for which they have no use in English. In spite of the rather contracted vowel space, the vowels within the space seem spectrally more distinct than those of the Chinese speakers. There is a clear spectral difference between intended /i:/ and /ɪ/, which is predicted as positive transfer should occur from Dutch to English (see Chapter three). There is a fair degree of separation between intended /ɛ/ and /æ/. Although the separation is not as large as in the native American speech (see below), the success on the part of the Dutch speakers is unexpected, and in fact runs counter to the prediction from the contrastive analysis in Chapter three. The /ɛ/ ~ /æ/ contrast is typically listed as a cause for the formation of new sounds (Flege) and we are surprised to find that in our group of ESL speakers some notion of the difference has already been established. Interestingly, the other vowel pair that has traditionally been mentioned as a cause for the formation of new sounds, /u:/ ~ /ʊ/, remains completely undifferentiated in the Dutch ESL speakers – as predicted by the contrastive analysis in Chapter three.

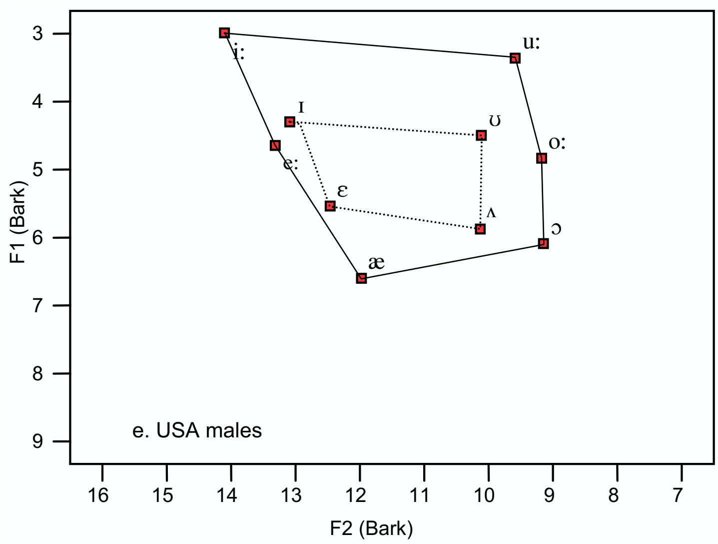
Dutch and English both have tense and lax vowel subsets. Inspection of figure 5.1c-d, however, shows that the tense and lax subsets are not very clearly separated in Dutch ESL. One reason for the relatively poor separation between the subsets is the lack of a /u: ~ ʊ/ contrast in Dutch. The Dutch speakers do not spectrally distinguish between the two, so that here the two subsystems merge. Also, at the lower edge of the vowel space there is little differentiation between more centralized (half) open lax vowels and peripheral open tense vowels as the Dutch ESL speakers do lower /ɔ/ as much as they should for American English, and at the same time observe insufficient contrast between /ɛ/ and /æ/.

If we now turn to the **American** native realisation of the vowels, in figure 5.1e-f, we notice that the vowel spaces are larger than those found for the Dutch ESL speakers, but much smaller than those of the Chinese ESL speakers. Nevertheless, the American native vowels are spectrally much more distinct than those produced by the Dutch speakers, and even more so than the Chinese ESL vowels. There are very large spectral differences between the members of the pairs /i: ~ ɪ/, /ɛ ~ æ/ and /u: ~ ʊ/. Moreover, the figure illustrates quite convincingly that the tense and lax vowel subsets are organised in terms of an outer (peripheral) and an inner (more centralised) area. In this respect, too, the L1 speakers clearly differ from both the Dutch and (even more) from the Chinese ESL speakers


i:
ɪ
u:
ʊ
e:
o:
ɔ
ɛ
ʌ
æ
a. Chinese males
F1 (Bark)
3
4
5
6
7
8
9
16
15
14
13
12
11
10
9
8
7
F2 (Bark)


i:
ɪ
u:
ʊ
e:
o:
ɔ
ɛ
ʌ
æ
b. Chinese females
F1 (Bark)
3
4
5
6
7
8
9
16
15
14
13
12
11
10
9
8
7
F2 (Bark)


c. Dutch males
F1 (Bark)
F2 (Bark)
i:
ɪ
e:
ɛ
æ
ʌ
u:
ʊ
o:
ɔ
d. Dutch females
F1 (Bark)
F2 (Bark)
i:
ɪ
e:
ɛ
æ
ʌ
ʊ u:
o:
ɔ



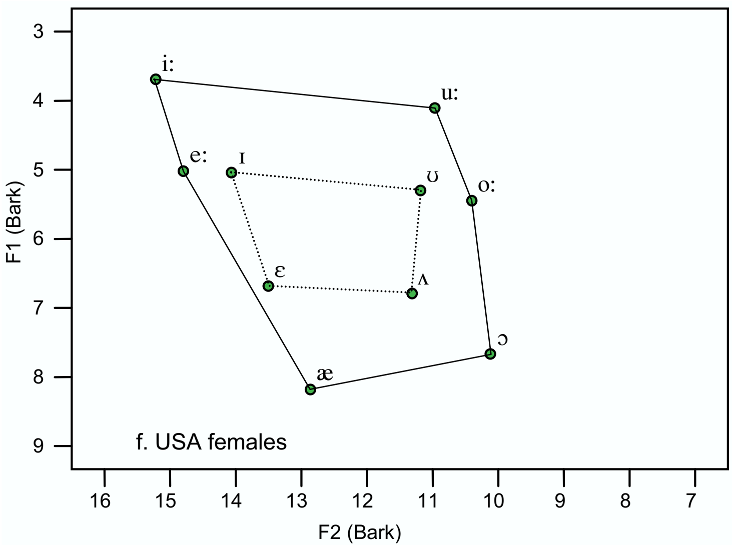


Figure 5.1. The mean values of F1 and F2 (in Bark) of the ten American English monophthongs plotted separately for tense (solid polygons) and lax (dotted polygons) vowels for six groups of speakers (indicated in each panel). Male speakers are represented by squares, female speakers by circles.

5.3 Vowel duration in Chinese, Dutch and American English

The vowels of American English are often separated into two length categories, short and long (see Chapter three). Phonetically, the four short vowels, /ɪ, ɛ, ʌ, ʊ/ do not only have short durations, they also take up more centralized positions in the vowel space. For this reason, the set of four may be called lax as well. The other vowels of American English are long and have positions along the outer perimeter of the vowel space. These are, in the present restricted dataset, the vowels /i:, e:, æ, ɔ, o:, u:/.

Since vowel duration may be expected to contribute to the perceptual identification of vowel tokens by English listeners, we measured vowel duration in each of the 600 tokens in our dataset, and plotted mean vowel duration for each of the ten types, separately for lax and tense categories in figure 5.2a for Chinese ESL speakers, in panel b for the Dutch speakers and in panel c for the American L1 speakers.

Taking the native speakers as our starting point, figure 5.2c clearly shows that the four lax/short vowels have much shorter duration (with means between 169 and 185 ms) than the six long/tense vowels (with means between 225 and 266 ms). As a result of this, vowels that are spectrally close to each other, such as /e:/ (266 ms) and /ɪ/ (184 ms), are yet acoustically distinct. Note also that when the vowels are ordered from short to long, as has been done in figure 5.2c, the increment between adjacent vowels in the figure is never more than 14 ms (which is the difference in mean vowel duration between /o:/ and /æ/). However, the discrepancy between the longest of the short vowels (/ɛ/, 185 ms) and the shortest of the long vowels (/u:/, 225 ms) is 40 ms. These results can be taken in evidence of the phonetic correctness of the subdivision of the American English vowels into the short and long categories made here.

If we now consider the vowel durations produced by the Chinese ESL speakers (figure 5.2a) we note that the short vowels are roughly within the duration range of the American L1 speakers. Also, the long vowels are generally within the native range for long vowels, with the exception of the vowels /æ/ and /ɔ/. Interestingly, these are precisely the vowels that distributionally pattern with the short vowels, as they cannot occur at the end of a word in English. When foreign learners are trained to pronounce English according to British (RP) norms, short vowel duration for /æ/ and /ɔ/ could reasonably be expected. However, given the fact that the Chinese ESL speakers were taught according to American pronunciation norms, this explanation is ruled out. We must assume, therefore, that the vowel duration of /æ/ and /ɔ/ has an incorrect perceptual representation in Chinese ESL speakers.

The Dutch ESL vowel durations are surprisingly similar to the Chinese realisations. Again, there are two gross duration categories, one for short vowels with durations less than 200 ms, and one for long vowels with durations in excess of 240 ms. As in the Chinese ESL tokens, the Dutch speakers make the long vowels /æ/ (208 ms) and /ɔ/ (172 ms) too short by American-English standards. Moreover, the Dutch speakers, who did not differentiate between /u:/ and /ʊ/ in spectral terms (see figure 5.1c-d), also have a tendency to make the short /ʊ/ too long (202 ms) – even though this is still some 40 ms shorter than their mean duration for long /u:/. Unexpectedly, then, it seems as if the Dutch ESL speakers are not more successful in keeping the American-English lax and tense vowels distinct than the Chinese speakers, even though Dutch is language with a tense ~ lax subdivision, which is not the case for Mandarin.

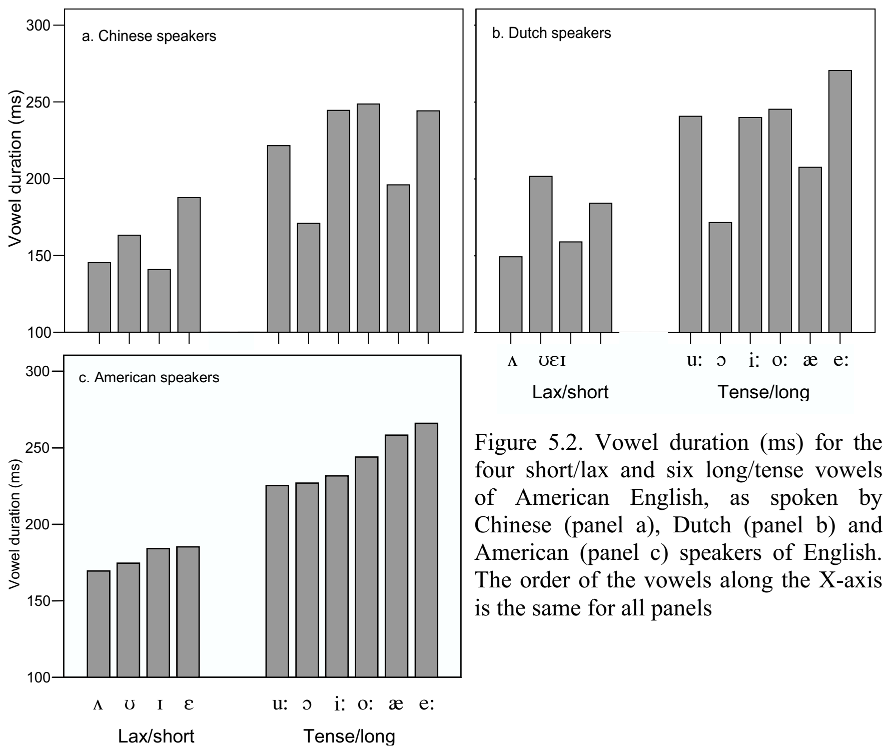


Figure 5.2. Vowel duration (ms) for the four short/lax and six long/tense vowels of American English, as spoken by Chinese (panel a), Dutch (panel b) and American (panel c) speakers of English. The order of the vowels along the X-axis is the same for all panels

5.4 Automatic vowel classification

So far we have only considered the means of the realisations of the vowels – in terms of vowel quality (F1 and F2) and of duration – averaged over groups of ten male and ten female speakers. The means do not tell us anything about how well the individual speakers keep the vowels distinct in their pronunciation of English. Figure 5.3a-c plot the individual realisation of the vowels in the F1 by F2 plane as scatter clouds, enclosed by spreading ellipses. These ellipses are drawn along the principal component axes, optimally capturing the directionality of the scatter of the vowel tokens within one vowel type. The ellipses have been plotted at + and – 1 SD away from the F1-F2 centroids. Before computing the scatter points and the ellipses based upon them, however, speaker normalization had to be carried out – as explained in § 5.1.2 – in order to make the vowel tokens produced by different individuals of different genders comparable.

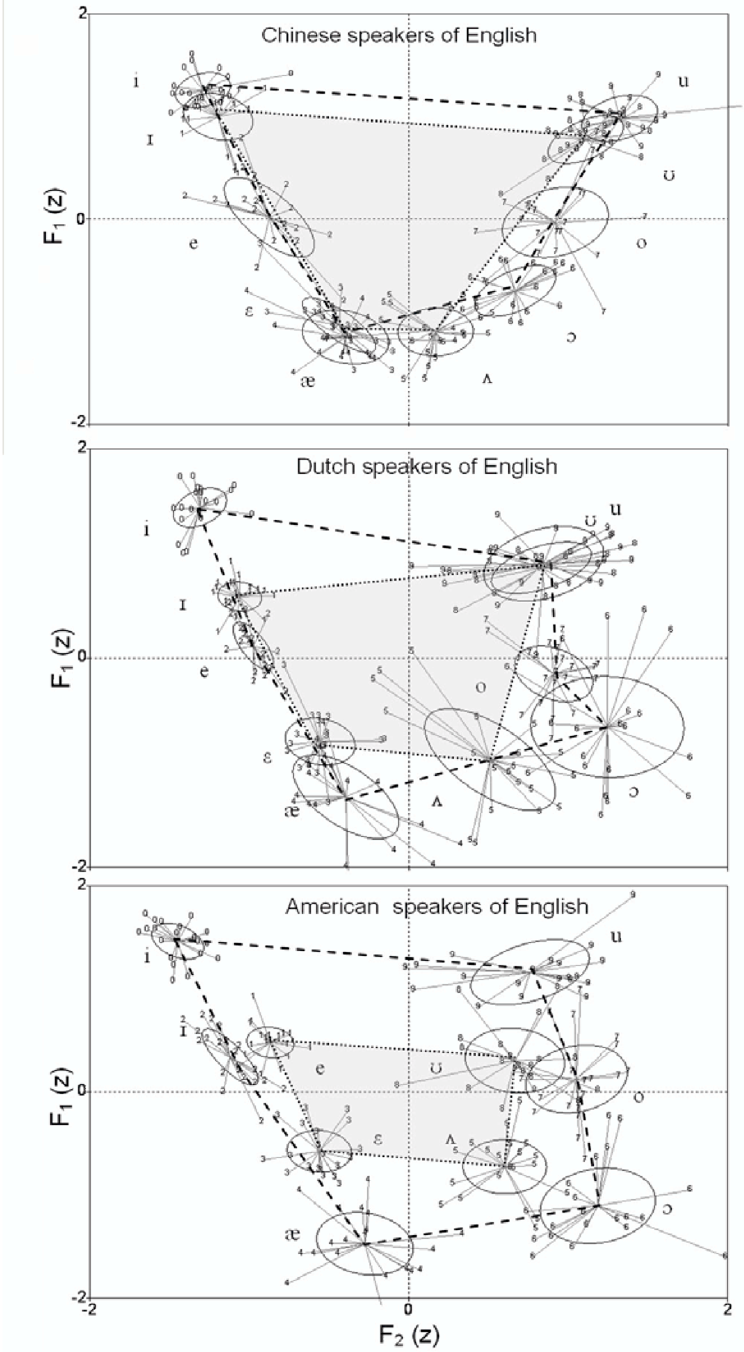


Figure 5.3a-c. Individual vowel points for Chinese speakers of English (after speaker-individual z normalization) plotted in the F1 by F2 plane, with spreading ellipses drawn at +/−1 SD away from the centroid along the first two principal component axes of the scatter clouds.

The figures show that, generally, the Chinese speakers (Figure 5.3a) have more overlap between the ellipses of neighbouring vowels than is the case in the Dutch ESL realizations (Figure 5.3b). The American native L1 speakers have the smallest (Figure 5.3c).

We will now attempt to quantify the difference between the three speaker groups in terms of the degree of success in keeping the ten vowels distinct. We have used Linear Discriminant Analysis (LDA) for this purpose. LDA is an algorithm that computes an optimal set of parameters (called discriminant functions) which automatically classifies objects in pre-established categories. For a comprehensive treatment of LDA in research on vowel identification, see Weenink (2006). The more distinct the categories are in the dataset, the fewer the number of classification errors yielded by the algorithm. In the case at hand, the discriminant functions are based on linear combinations of weighted acoustic parameters F1, F2 and duration. Again, before running the LDA, speaker normalization was carried out using the z-transformation on the formant frequencies in Barks. The results of the LDA are presented in terms of confusion matrices (see Table 5.1 on the next page), which show the intended vowels in the rows against the vowels as classified by the algorithm in the columns. Correctly classified vowel tokens are in the cells along the main diagonal. All off-diagonal cells contain confusions.

I will first present the overall percentage of correctly classified vowel tokens of Chinese, Dutch and American speakers of English. Moreover, we ran the LDA twice. The first time we just included the two spectral parameters as possible predictors of vowel identity, i.e. F1 and F2 (converted to Bark and z-normalized within individual speakers). The second time we extended the set of predictors by also including (z-normalised) vowel duration. Figure 5.4 presents these results.

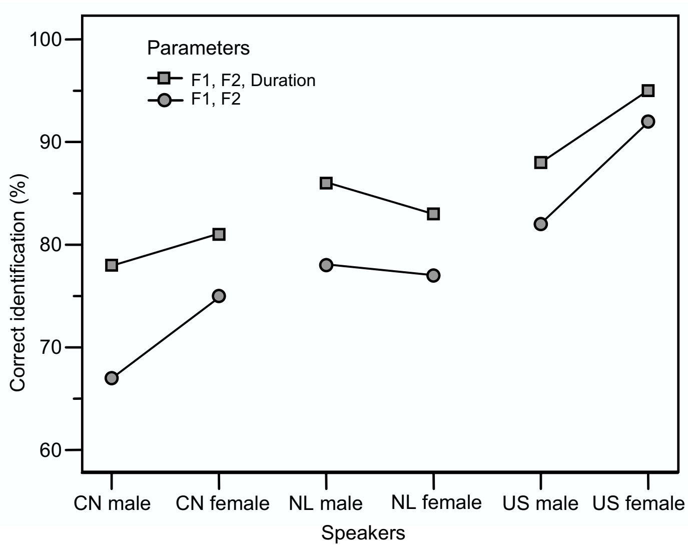


Figure 5.4. Percentage of correctly classified vowel tokens by Linear Discriminant Analysis with F1 and F2 as predictors, and with duration added as a third predictor for six groups of speakers (male and female Chinese, Dutch and American speakers of English).

Figure 5.4 shows at once that the vowels as spoken by the native speakers afford the best automatic identification, those spoken by the Dutch learners can be less successfully identified, and the Chinese ESL tokens are poorest. Adding duration to the set of predictors boosts the correct identification by some 10 percentage points (a little less for the American L1 vowel tokens, possibly due to a ceiling effect). Finally, the vowel tokens produced by the female speakers tend to be more distinct, and therefore better identified, than those spoken by the males. However, there is no such gender effect in the Dutch vowel set.

A more detailed view of the LDA results is presented in Table 5.1, where percent predicted vowel identity is crosstabulated against the actual vowel identity for Chinese, Dutch and American native speakers in the upper, middle and lower panels, respectively. The results presented in this table were based on the output of the LDAs which used F1, F2 and vowel duration as predictor variables.

The results obtained for the Chinese-accented vowel tokens reveal two major problems, viz. the more or less symmetrical confusion of /ɛ/ and /æ/ and an asymmetrical confusion of lax /ʊ/ with tense /u/ (but not vice versa). These pronunciation errors follow from a traditional contrastive analysis, and were also noted in a pedagogical textbook (Zhao, 1995).

In the results for the Dutch speakers of English we find two symmetrical error patterns, i.e. /ɛ/ ~ /æ/ and /ʊ/ ~ /u/ and their counterparts, all of which were predicted by contrastive analyses (Table 3.4) and were noted in the pedagogical literature (Tables 3.5 and 3.6 for Dutch and Chinese speakers, respectively). One incorrect classification type was never predicted, however. This is the incorrect classification of intended vowel /ʌ/ as a front vowel /ɛ/.

We will have occasion to review the LDA results from a different perspective in Chapter ten, where we will make an attempt to use the LDA to make predictions of cross-linguistic vowel perception, thus simulating for instance the perception of Dutch-accented vowel tokens by Chinese listeners of English. Before we discuss such attempts, we will first deal with the results of human perception of vowels, consonants, consonant clusters and words in meaningless and meaningful contexts in Chapters six through nine.

Table 5.1. Classification matrices of observed and predicted vowel identity of English vowel tokens produced by Chinese (upper panel), Dutch (middle panel) and American native speakers. Prediction of vowel identity made by Linear Discriminant Analysis using F1, F2 and vowel duration as predictors. Percent correct in parentheses.

Presented vowels	Vowel identity predicted from Chinese production data (80.0%)										
	æ	e:	ɛ	i:	ɪ	ɔ	o:	ʊ	ʌ	u:	Total
æ	**60**		30						10		100
e:		**85**	10	5							100
ɛ	20	5	**65**						10		100
i:				**95**	5						100
ɪ					**100**						100
ɔ						**80**	5		15		100
o:						20	**70**			10	100
ʊ							5	**70**		25	100
ʌ			5			10			**85**		100
u:								10		**90**	100

Presented vowels	Vowel identity predicted from Dutch production data (83.0%)										
	æ	e:	ɛ	i:	ɪ	ɔ	o:	ʊ	ʌ	u:	Total
æ	**75**		25								100
e:		**95**			5						100
ɛ	15		**80**		5						100
i:				**100**							100
ɪ					**100**						100
ɔ						**80**	5	10	5		100
o:						5	**90**	5			100
ʊ								**60**		40	100
ʌ			15			10			**75**		100
u:							5	20		**75**	100

Presented vowels	Vowel identity predicted from American production data (92.5%)										
	æ	e:	ɛ	i:	ɪ	ɔ	o:	ʊ	ʌ	u:	Total
æ	**100**										100
e:		**95**			5						100
ɛ			**100**								100
i:				**100**							100
ɪ					**100**						100
ɔ						**85**	10		5		100
o:						5	**80**	5		10	100
ʊ			5				5	**80**		10	100
ʌ						10		5	**85**		100
u:										**100**	100