Perception of checked vowels by early and late Dutch/English bilinguals

Vincent J. van Heuven, Nicole N. Broerse and Jos J. A. Pacilly

1. Introduction

There has been considerable debate on the question to what extent someone can acquire a native pronunciation in a second or foreign language. It is generally observed that adults who immigrate to a foreign country, do not normally acquire the language of the new country without an accent that is reminiscent of their mother tongue. In the situation described above, the L_2 (second or foreign language) learners acquire the new language after that the first language (L_1) was well established. However, when young children immigrate, they usually learn to speak the new language without a trace of a foreign accent.

When one has learnt a first language, speech sounds are typically perceived in terms of the (phoneme) categories of the native language. These categories were shaped during the first twelve months after birth. According to the motor theory (MT) of speech perception (e.g. Liberman, Cooper, Shankweiler and Studdert-Kennedy 1967), sounds that belong to the same phoneme category, are not readily discriminated by native listeners; however, even a small phonetic difference between two sounds is easily perceived if these sounds belong to different phonemes, i.e. are separated by a phoneme boundary. A more recent theoretical development holds that a physically equal difference between two sounds is perceived as smaller the closer the sounds lie to a category prototype (Native Language Magnet theory, NLM, of speech perception, Kuhl and Iverson 1995). Together, the mechanisms explained by MT and NLM conspire to make it extremely difficult to fully acquire a new sound system after the age of puberty. Lack of neural plasticity has been claimed to prevent learners from modifying the perceptual categories that were developed in the early stages of life (e.g. Lenneberg 1967).

Nevertheless, there are indications that at least some adults have been able to learn to pronounce a foreign language in a way that cannot be distinguished from that of born-and-bred native speakers, even though the learning process did not start until after puberty. Bongaerts, Van Summeren, Planken and Schils (1997) recorded a number of intermediate and highly advanced Dutch L_2 speakers of English. Each L_2 speaker produced five English test sentences, which were subsequently judged by native English listeners and rated on a scale for nativeness. The recordings randomly alternated with control utterances spoken by native English speakers. Seven of the eleven highly advanced L_2 speakers (typically professors or lecturers in English at the level of university or teacher training college) obtained a pronunciation score that was equal to, or even better than, that of some of the native English speakers (see figure 1).

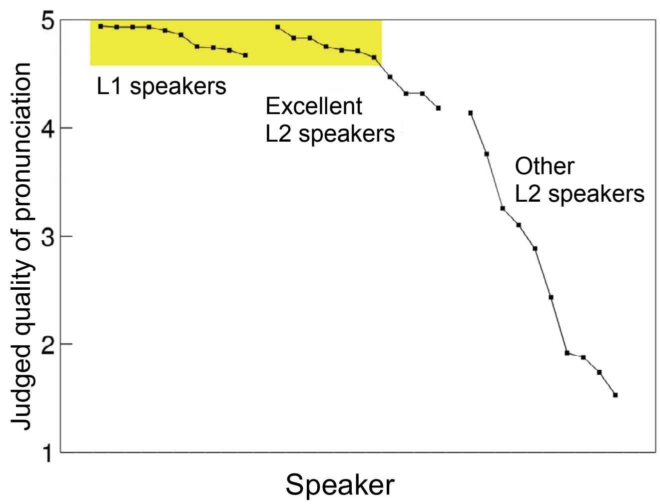

Figure 1. Quality of British-English pronunciation rated by native English listeners (5: 'No accent at all, definitely native' – 1: 'Very strong foreign accent, definitely nonnative') of three samples of speakers: native speakers of British English (L_1), excellent L_2 speakers with Dutch as the mother tongue, and an intermediate group of L_2 speakers (including some students of English). The grey area marks the overlap between native L_1 and excellent L_2 speakers (after Bongaerts et al. 1997).

The first question we address in the present paper is whether such excellent L_2 speakers have a mental representation of the English vowel sound system that is equal to that of native L_1 speakers of British English. It would be our hypothesis that if the excellent L_2 speakers cannot be differentiated from native speakers by their speech production, their perceptual representation of the sound categories should also be similar.

It is generally recognized that young children who grow up in a bilingual setting, acquire a native pronunciation in both languages they are exposed to. Bilingual settings may occur, for instance, when one parent is a native speaker of language A (e.g. English) and the other of language B (e.g. Dutch). When each parent communicates with the child in his/her native language, the child will end up with an equal command of, and native pronunciation in, both languages. Alternatively, the parents may have immigrated to a new country but rear the child in the native language. When the child goes to school (starting at the age of four) it is regularly and massively exposed to the language of the host country, which it will acquire at the native level – without losing its proficiency in the parents' language. Such bilingually raised children are referred to as early bilinguals – indicating that they acquired the bilingualism well before the age of puberty, i.e. when the critical period for acquiring a native pronunciation in any language had not elapsed. The term 'early bilingualism' is used in contrast with 'late bilingualism', which is the situation referred to above, when someone acquires the L_2 after the age of puberty, i.e.

after the complete acquisition of the L_1. It is the second aim of this paper to establish if there is any difference in the perceptual representation of the sound systems of two languages that are spoken by early and late bilinguals.

Not much is known about the research topics introduced above. It is well-known, however, that late bilingualism comes at a price. When one learns the sounds of an unfamiliar language, several possibilities may present themselves. Flege's (1995) Speech Learning Model (SLM) makes a distinction between three scenarios.

(i) A sound category (phoneme) in L_2 may be indistinguishable from its counterpart in the learner's L_1. In a phonetic transcription the L_1 and L_2 sounds would be written with exactly the same base symbol and diacritics. In this case of '*identical sounds*' the learner will use the L_1 category in the L_2 without introducing even a hint of foreignness. The nasal consonants /m, n, ŋ/ in English and Dutch would be examples of identical sounds

(ii) In certain other cases the L_2 has a sound category that has no match in the learner's L_1. The L_2 category may be in between two L_1 categories or it may find itself somewhere in the phonological space where the L_1 has no categories at all. Typically, the international phonetic alphabet (IPA) has distinct base symbols to represent the L_1 and L_2 sounds. Here the learner will – sooner or later – become aware of the difference between the '*new sound*' in the L_2 and his native L_1 categories. SLM predicts that the learner will ultimately set up new categories for these sounds, which will be faithful approximations of the L_1 sound categories. When these new sounds are in place, they will no longer contribute to foreign accent on the part of the learner. Moreover, since the new sounds are added to the set of categories that exist in the learner's L_1, the latter will not be affected in any way by the inclusion of the additional categories, so that the pronunciation of the L_1 will not change. English /æ/ as in *bad* would be an example of a new sound for Dutch learners of English; conversely, the Dutch /œy/ as in *huis* 'house' would be a new sound for an English learner of Dutch (Collins and Mees 1984).

(iii) The third scenario applies when the L_2 has sound categories that are similar but not identical to their nearest equivalent in the learner's L_1. The learner is not aware of a subtle difference between the L_2 sound and its L_1 equivalent but native listeners of the target language perceive the deviant sound as foreign. Similar sounds would be designated by the same IPA base symbol but differ in their diacritics. English /s/, for instance, has its spectral centre of gravity at a higher frequency (and therefore has a sharper timbre) than its Dutch counterpart. Dutch learners of English consistently fail to notice this difference and substitute their own, duller, Dutch /s/ when they speak English, and as a result sound foreign to the native English listener (Jongman and Wade 2007). Since the learner is not aware of the difference between the L_1 and the L_2 sound, a more widely defined sound category is formed which includes the observed variability of both the L_1 and the L_2 sound. The result of this fusion is a compromise category which is noticeably incorrect both in the L_2 and in the L_1 of the learner: the learner will not only sound foreign in the L_2 but also in the L_1.

We predict that the perceptual representation of the sound categories in L_1 and L_2 of late bilinguals will be different than that of early bilinguals. In the case of early bilinguals we expect the sound categories to be similar in nature within one language. That is to say, all the phoneme categories in L_1 will have the same sharpness of definition. The same applies to the categories in the other language. This hypothesis does not rule out the possibility that the sound categories of one language may be more sharply defined than those of the other language. Very likely, one of the two languages ~~which~~ are spoken by an early bilingual will be more prominent (or dominant) than the other; this difference in dominance may come to light in the well-definedness of the sound categories. We expect that the categories in the dominant language will be more sharply defined, i.e. have smaller boundary widths than those in the other language – but within each of the two languages the categories have similar boundary widths.

In the case of late bilinguals, we hypothesize that the categories in the L_1, which were formed during childhood, will be more sharply delineated than those in the L_2. Also, in the case of similar sounds, in terms of the Speech Learning Model, the ideal (prototypical) location of the sound in the phonetic space will be in between the preferred locations found for the L_1 sound and its equivalent in the L_2.

It is our ultimate prediction that the perceptual representation of the sounds in the L_1 and the L_2 will be different – in terms of the location of the prototypes and in the sharpness of the category boundaries – for monolingual speakers of a language as compared to bilingual speakers, whether early or late, and that there will also be differences between the early and late acquirers. More generally, we predict that even excellent L_2 learners, whose pronunciation can no longer be distinguished from native L_1 speakers, can still be shown to have perceptual representations of the L_2 sound system that deviates from that of monolingual speakers of the target language.

2. Approach

It follows from the introduction that we needed four different groups of subjects in order to find an answer to the research questions formulated. These are (i) a monolingual group of native English speakers, and (ii) a similar group of monolingual Dutch native speakers.[1] These two groups served as control conditions that establish the baseline against which the performance of the bilingual speakers was to be gauged. Two bilingual groups of listeners took part in the study, i.e. (iii) an early bilingual group, whose members all had learnt English and Dutch from childhood onwards in a typically

[1] Since Dutch children take compulsory lessons in English from the age of ten onwards, a purely monolingual adult speaker of Dutch is hard to find. It is generally accepted, however, that Dutch speakers of English have no clear idea of the sound categories of English – with the exception of students of English, either at the university or at teacher training colleges, who are explicitly trained to pronounce English without an accent. The latter type of subjects was not included in the monolingual speaker group.

bilingual setting (details see below), and (iv) a group a late bilinguals, which was comprised of the same type of learners that was targeted by Bongaerts et al. (1997).

In order to keep the experiment within reasonable bounds, we limited the study to the perceptual representation of only the short vowel phonemes in two related languages, i.e. (British) English and (Netherlandic) Dutch. Apart from six long vowels, three true diphthongs and – depending on the depth of the phonological analysis – a number of centering diphthongs, English has six short vowels (also termed lax vowels), which – phonotactically – cannot occur at the end of a word.[2] These are /ɪ, e, æ, ʌ, ɒ and ʊ/, i.e., the vowels that occur the words *hid*, *head*, *had*, *hud*, *hod*, and *hood*. In addition to four long vowels and three diphthongs, Dutch has eight short vowels, five of which are phonologically lax (and cannot occur at the end of a word), i.e. /ɪ, ɛ, ɑ, ɵ, ɔ/ as in *bid* 'pray', *bed* 'id.', *bad* 'bath', *put* 'well', and *bod* 'offer', and three more which are tense but remain phonetically short (unless followed by /r/ in the first syllable of a trochaic foot, cf. Kooij and Van Oostendorp 2003), i.e. /i, y and u/ as found in *bied* 'offer', *buut* 'terminal point' and *boet* 'make repairs', respectively.

We decided to generate an artificial vowel space that included all the short vowels in Dutch and English, and to sample vowels at regular intervals from this space. This approach is reminiscent of Schouten (1975) with the exception that we used (perceptually) uniform sampling whereas Schouten sampled his vowels by interpolating in small steps between the prototypes of all the vowels of English (whether long or short). We reasoned that such non-uniform sampling is inefficient and compromises the comparability between the vowels in the system. By limiting our study to only the short vowels, we could synthesize the vowels with realistic durations, while Schouten (1975) synthesized his vowels at a duration that was halfway between that of short and long vowels. As a result Schouten's vowels were always non-ideal, or non-prototypical, exemplars of their category. Another point of difference between Schouten's stimulus materials and ours is that, where Schouten used isolated vowels produced out of any spoken context, we synthesized the vowels in the context of a carrier phrase, embedded between an initial and a final consonant.

The location of the prototypes of the various vowel categories can be found by asking listeners to indicate, first of all, which of the six English or eight Dutch vowels they associate most readily with a given stimulus vowel, and second, how good they think the token is as an exemplar of the category chosen. In this way, we may map out the perceptual representation of the vowel space of the English or Dutch listener in terms of the location of the prototypes, i.e. the most preferred vowel tokens in the set of artificial vowel sounds. Although it is also possible to define the boundary widths from these judgments, we decided to establish the boundary widths by a shortcut method

[2] Schwa was not included in the sets of short vowels, neither in English nor in Dutch. We used vowels that could occur in stressed monosyllabic words, which requirement rules out the inclusion of schwa. Also in terms if its phonotactics, schwa is not on a par with the regular short vowels, since it can occur at the end of words.

proposed by Van Heuven and Van Houten (1989). The shortcut is to present each vowel token for identification twice (in different random orders) and to compute the consistency with which the listeners label the two tokens of each vowel type. The poorer the definition of the category boundary, the poorer the labeling consistency.

3. Methods

3.1 Stimulus materials

The English vowels /ɪ, e, æ, ʌ, ɒ and ʊ/ were placed in the context *I say m_f* and recorded in a soundproofed recording studio. The set was repeated three times at conversational speed by a late bilingual speaker of English and Dutch, who was an experienced phonetician with Dutch as his native language. This procedure was repeated with /ɪ, ɛ, ɑ, ɵ, ɔ, i, y and u/ in the Dutch context *Ik zei m_f*. The mean duration of all the Dutch vowel tokens (127 ms) and of all the English vowel tokens (133 ms) was measured. On the basis of this result, it was decided to adopt a single vowel duration of 130 ms for all the vowels in the stimulus set to be generated.

The same speaker then produced extreme versions of the Dutch point vowels /i, u, a, ɑ/ (as approximations of cardinal vowels 1, 4, 5 and 8) in the context of the Dutch carrier *Ik zei m_f* 'I said …'. The recordings were analog-to-digital converted (12 bit, 10 KHz) and submitted to an analysis of formant frequencies and bandwidths, using the split-Levinson algorithm (Willems 1987) with a 25-ms window and a 10-ms frame shift. Fundamental frequency was extracted by the subharmonic summation method described in Hermes (1988), which algorithm was also used to make a voiced/voiceless determination of each analysis frame. Formants F_1 to F_5 and the associated bandwidths B_1 to B_5 were computed. Table 1 below lists the values for the lowest two formants measured at the temporal midpoint in each of the four cardinal vowels.

Table 1. Formant centre frequencies and bandwidths measured for the four cardinal vowels which formed the basic framework for the stimulus space in the experiment.

	[i]	[a]	[ɑ]	[u]
F1	300	980	790	360
F2	2640	1620	1210	1157
B1	30	98	79	36
B1	264	162	121	116

A two-dimensional grid was defined as a vowel space which was spanned between the extreme point vowels in table 1. The vowel height dimension was sampled with nine steps at 0.7 Bark apart along the first formant interpolating between 308 and 1050 Hz (from 3.2 to 8.8 Bark).[3] The combined vowel backness and rounding dimension was sampled with ten steps of 0.7 bark along the second formant between 1033 and 2732 Hz (from 8.7 to 15.0 Bark). This defines a 9 (F_1) × 10 (F_2) grid (see figure 3), from which cardinal 5 and the two types closest to it were later eliminated on the grounds that these tokens sounded highly unnatural. The F_1 and F_2 values chosen were the targets to be attained at the temporal midpoints of the vowels. The target values of the higher formants were kept constant at the values found for the token containing cardinal vowel 4.

The resulting 87 vowels were embedded in either the Dutch or the English carrier phrase using LPC resynthesis in order to ensure smooth and realistic formant transitions between the initial /m/ and the final /f/ sound of the target word. The target vowel consisted of a steady state segment of six frames with a six-frame transition (i.e. lasting 60 ms) on either side, the entire phrase being 105 frames for both language conditions The stimulus sentence was resynthesized with constant segment durations and a fixed F_0 contour with an accent-lending rise-fall pitch movement on the target syllable.

For each language, the 87 stimulus utterances were converted back and recorded on analog audio tape in one random order, preceded by three practice items, and repeated in reversed order, preceded by the same three practice items. The interstimulus interval was 3 s (offset to onset); a short beep was recorded after every tenth item.

3.2 Subjects

Fifteen monolingual English speakers (seven males) were recruited at Edinburgh University with a mean age of 21;2 years and a range from 19 to 25. All but one were raised in the south of England and spoke RP English. The fifteenth subject was born in Portugal but raised by RP speaking parents and later attended public school in the UK. None of these subjects had any working knowledge of Dutch.

Fifteen Dutch-speaking subjects (four males) were recruited at Leiden University, The Netherlands. All were selected on the basis of their ABN (Algemeen Beschaafd Nederlands) variety, the Dutch equivalent of RP. The mean age for the monolingual Dutch subjects is 26;4 years with a range between 19 and 37. None of these subjects had a special affiliation with either the English language or English-speaking countries.

Fifteen bilingual speakers were found, eight of whom were early bilinguals. The remaining seven were native speakers of Dutch who had acquired an exceptional command of English after the age of puberty (late bilinguals). Inclusion proceeded on the basis of three main criteria: (i) they were assessed as having native competence by

---

[3] The Bark transformation is an empirical formula that adequately maps the differences in Hertz-values onto the perceptual vowel quality (or timbre) domain. A difference of 1 Bark, in whatever direction, is a perceptually equal difference in vowel quality, irrespective of its location in the F_1-by-F_2 space. We used the Bark formula proposed by Traunmüller (1990): Bark = $[(26.81 \times F) / (1960 + F)] - 0.53$, where F represents the measured formant frequency in Hertz.

native speakers of each language, (ii) they made regular use of both languages, and (iii) they subjectively considered themselves as bilingual.[4] In distinguishing between the natural (early) and artificial (late) bilinguals, a further requirement was that the former had acquired their languages, preferably simultaneously, before the age of ten.

3.3 Procedure

Stimuli were presented to listeners over good quality headphones (Sennheiser HD414) in individual sessions. Subjects indicated for each of the utterances they heard which vowel they recognized, with forced choice from either the six English short vowels, or from the eight Dutch short vowels – depending in the presentation mode. Listeners were issued response sheets that listed the six or eight vowels in a quasi-phonetic spelling, which was explained and exemplified with unambiguous sample words in the instruction text. The English response categories were *i* (as in *pit*), *e* (as in *pet*), *a* (as in *pat*), *o* (as in *pot*), *U* (as in *putt*) and *u* (as in *put*). The Dutch categories were *i* (as in *rit* /rɪt/ 'ride'), *e* (as in *red* /rɛt/ 'save'), *a* (as in *rat* /rɑt/ 'id."), *o* (as in *rot* /rɔt/ 'rotten'), *u* (as in *Ruth* /rɵt/ 'id.'), *ie* (as in *riet* /rit/ 'reed'), *uu* (as in *Ruud* /ryt/ 'Rudy') and *oe* (as in *roet* /rut/ 'soot').

The monolingual English listeners heard only the 2 × 87 items in their own language and responded with forced choice from the eight English short vowel categories. The bilingual listeners and the Dutch monolinguals took the test twice, the first time listening to the Dutch materials (and responding in the Dutch mode), the second time to the English version (and responding in the English mode).

4. Results

Preliminary screening of the responses revealed that the categories *ie* and *uu* (the high front unrounded and rounded vowel type, respectively) were severely underrepresented, with no more than 120 responses to either of these categories (against up to 864 responses for the other categories). Possibly, the frequency of the F_1 was not low enough to elicit a sufficient number of responses to these categories. Consequently, we excluded these two vowels as possible response categories from further data analysis – so that both the Dutch and the English response sets are limited to six possibilities.

We will now present the results in three successive stages. First, we will consider the location of the prototypical realizations of each of the six short vowel types in Dutch and in English, as determined from the responses by the four listener groups. Second, we will look at the size of the categories as can be derived, rather loosely, from the magnitude of the dispersion ellipses that can be drawn around the prototypes (or

---

[4] Bert Schouten was one of the seven late bilinguals who participated in the experiment. His performance will not be separately shown in the results below.

centroids), and third, we will quantify the sharpness of the perceptual representation of the various vowel types in terms of the response consistency measure defined by Van Heuven & Van Houten (1989).

4.1 Location of prototypes

Figure 2 shows the location of the vowel prototypes in the F_1-by-F_2 space, which was determined the mean F_1 and the mean F_2 value of all the vowel types identified by a group of listeners as an instance of that category weighted by the number of responses. The three graphs in the left column of figure 2 present the centroids of the short vowels of English as perceived by early bilinguals (top row), late bilinguals (middle row) and by Dutch learners of English as a foreign language (bottom panel). The vowels perceived by monolingual English listeners are repeatedly shown in the three graphs as the corners of the shaded polygons.

The early bilinguals have virtually the same locations of the six short English vowels as the monolingual English reference listeners (top panel). Also, both early and late bilinguals, as well as the Dutch learners of English, show only minor discrepancies in the locations for the vowels /ɪ, e, ɒ/. However, discrepancies can be observed for the remaining three short vowels /æ, ʌ, ʊ/. These are vowels that would be classified as new sounds in terms of Flege's (1987) Speech Learning Model. The Dutch learners of English ('monolingual Dutch') deviate considerably from the RP-English targets. The discrepancy is moderate for /æ/ (about 0.5 Bark) in the Euclidean vowel space, and larger for /ʌ, ʊ/ (roughly 1.0 Bark). Predictably, the centroids for /æ, ʌ/ are raised (closer vowel qualities) since the nearest vowel in Dutch would be /ɛ/ and /ɵ/, respectively. The location of /ʊ/ in the perceptual representation of the monolingual Dutch listeners is further back than for the RP-listeners. Again, this follows directly from the circumstance that the nearest Dutch vowel is a very back /u/. Interestingly, these findings correspond closely to the production data of Dutch learners of English and those of native English speakers as reported by Wang and Van Heuven (2006).

Crucially, the late bilinguals have their perceived ideal location of /ʌ, ʊ/ in positions that are intermediate between those of the native and early bilingual listeners on the one hand and those of the Dutch monolinguals on the other. It would seem, therefore, that the late bilinguals' perceptual representation of the English vowels is affected by the presence of an interfering vowel in the competing language, i.e. Dutch. A rather unexpected phenomenon is seen in the ideal location of /æ/ as perceived by the late bilinguals, which is as low as that of the native RP listeners but considerably more fronted. This might be seen as a tendency on the part of the excellent learners of English as a foreign language to overcompensate (or exaggerate) the difference between the ash-vowel and its competitors. Finally, the late bilinguals' results show a negative effect of their excellent command of English on the perceptual representation of the L_1. This is clearly seen in the lowered and backed location of the vowel /ɔ/, which is identical to these listeners' perceptual representation of English /ɒ/.

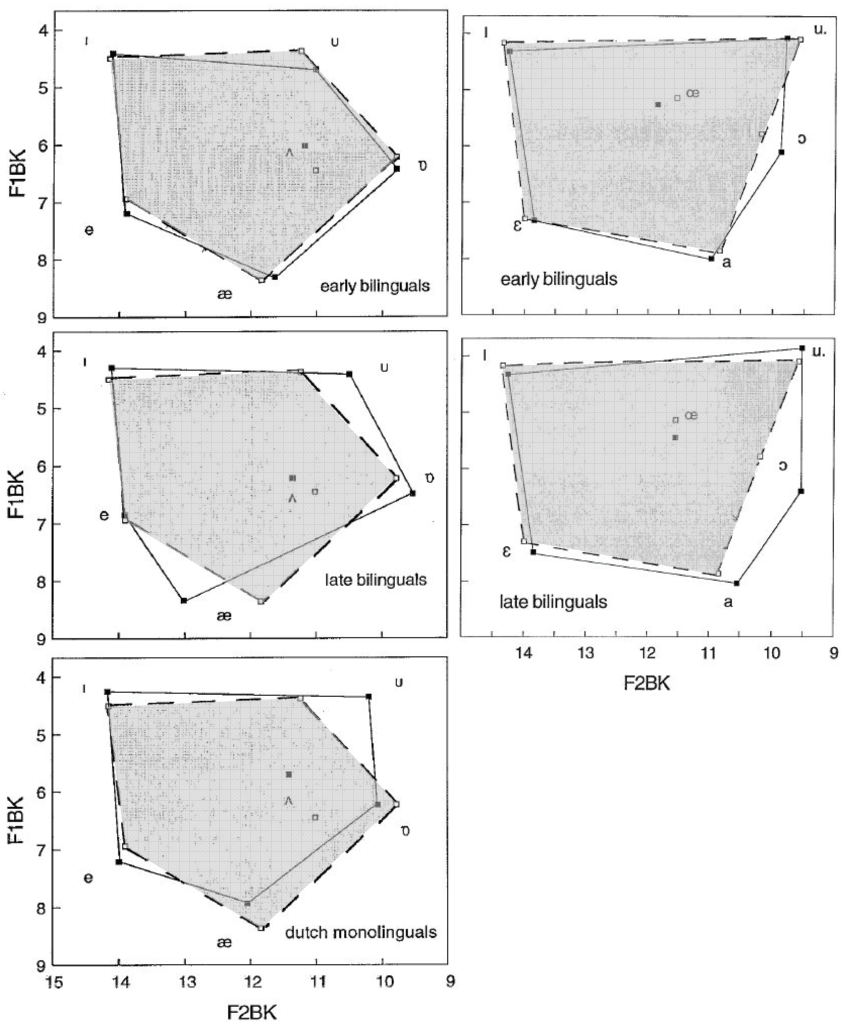

Figure 2. Vowel centroids in the F_1-by-F_2-space (in Bark) by early bilinguals (top row), late bilinguals (middle row) and Dutch monolinguals, as perceived in the English (left column) and Dutch (right column) response modes. Monolingual English listeners (left column) and Dutch (right column) listeners are repeatedly indicated by the shaded polygons for reference purposes.

4.2. Dispersion of perceived vowels

Figure 3 displays not only the centroids of the perceived vowels but also the dispersion of the vowels associated with a particular category. The dispersion ellipses were drawn at ± 1 standard deviation around the vowels centroids along the two principal components of the scatter cloud of each vowel type in the F_1-by-F_2 plane. The classification of each vowel type by the majority of the listeners is indicated by linking the vowel type to the centroid of the preferred category. The left and right top panels show the dispersion ellipses for the English (left) and Dutch (right) short vowels as perceived by monolingual English and Dutch listeners, respectively.

There is not much difference in term of dispersion between the early and late bilinguals' perceptual representation of the English vowels. The degree of overlap between adjacent categories is roughly comparable, with one exception. For both the late bilinguals and the Dutch learners of English the dispersion ellipses for the vowels /e/ and /æ/ overlap considerably, which can be understood from the circumstance that Dutch has just one low-mid front vowel where English has two. This type of native-language interference can be expected from learners of English as a foreign language but it is somewhat surprising that the interference should still be found with the late bilinguals.
A second observation would be that the bilingual listeners seem to have narrower vowel categories that the monolingual listeners. Possibly, then, the vowel space of bilingual speakers, whether early or late, is divided into a larger number of categories (i.e. the union of the vowel inventories of both languages), which would leave less space for each vowel in the combined inventory.

4.3 Labeling consistency

Each listener responded to each of the 87 stimulus vowel tokens twice. A consistency index was computed by dividing the number of vowel repetitions responded to in like manner by the total number of repetitions (= 87). The means are as in table 2.

Table 2: Consistency index per language (listening mode) and subject group.

Subject Group	Listening mode	
	English	Dutch
Monolingual	.70	.80
Bilingual	.71	.79
Early	.68	.75
Late	.75	.82
Dutch	.66	---

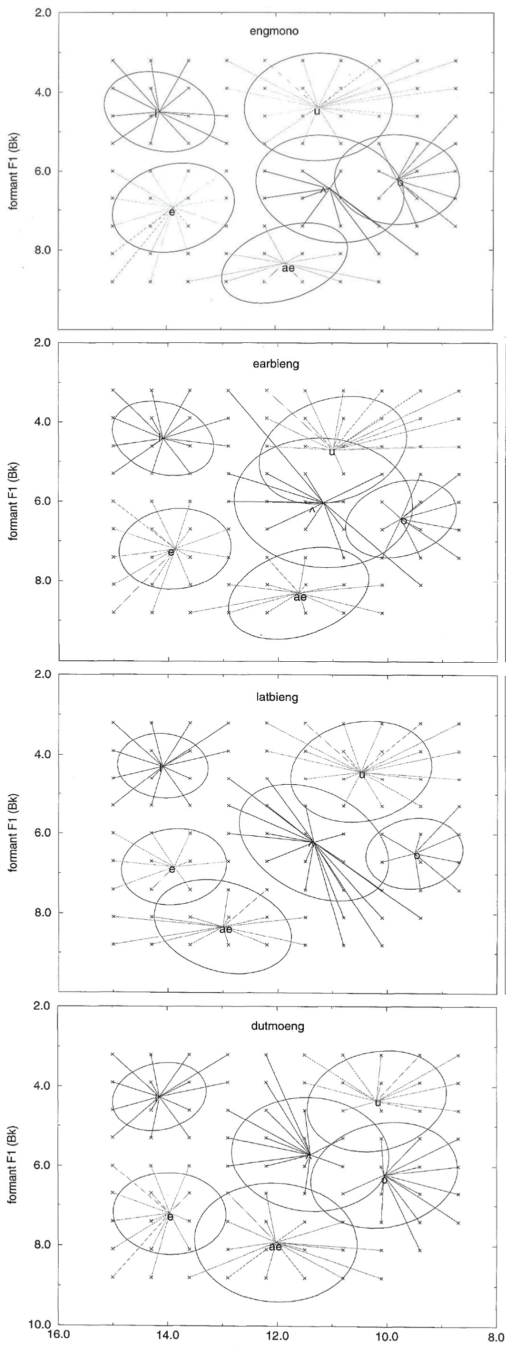

Figure 3. Centroids and dispersion ellipses (± 1 standard deviation) in the F_1-by-F_2 space (Bark) for six perceived short vowels in the English listening mode (left column) and in the Dutch listening mode (right column) by monolingual native listeners (top row), early bilinguals (second row), late bilinguals (third row) and Dutch L_2 listeners of English (bottom).

First, table 2 shows that the consistency in the Dutch listening mode is substantially better than in the English mode. We have no explanation for this difference, although we observe that the large spread of the centralized back vowel /u/ does lead to an unusual overlap (and therefore poorly defined category boundaries), which is avoided in Dutch, with its extremely back-articulated /u/ that has to be kept distinct from its front rounded neighbor /y/. The /y/, of course, is not a phonemic category in English, which is the reason why /u/ may have such a wide dispersion in RP. The results further confirm that non-native listeners categorize the L_2 vowel system less consistently (.66) than native listeners do. Also, although bilingual listeners, on average, have the same consistency in vowel identification as monolingual L_1 listeners, the data show a difference when we split up the group of bilingual listeners. Clearly, the early bilinguals are less consistent in either language than monolinguals are. Moreover, it would seem that the late bilingual listeners are the most consistent vowel identifiers, irrespective of the stimulus language. One reason why the late bilinguals are such consistent labelers, might be that these subjects – in contrast with the other listener groups – are professional linguists/phoneticians and pronunciation instructors. They are conscious of the differences between the Dutch and the English vowel in terms of their location in the vowel space, and may therefore be unusually intent on keeping the competing vowel categories separate – hence the smaller dispersion ellipses in figure 3 and the greater consistency in table 2.

5. Conclusion

We asked whether excellent speakers of a second of foreign language (L_2), whose pronunciation sounds as native as that of native (L_1) speakers, might still be different when it comes to their perceptual representation of the sound categories in the L_2 (and possibly even of their L_1). Our results show that the excellent adult Dutch learners of English have a perceptual conception of at least some of the vowels of RP English that differs from that of monolingual English listeners. Specifically the location of the 'new sound' /æ/, although more open than that in the pronunciation of less advanced Dutch learners of English, is more fronted than is the case in the perceptual representation of this vowel by monolingual English listeners (and of early bilinguals). We also noticed that the perceptual representation of the mid back vowel for the late bilinguals is the same in English and in Dutch – which it should not be since English /ɒ/ differs from Dutch /ɔ/; the late bilinguals' representation of this vowel is correct for English but wrong for Dutch. This shows that excellent learners of a foreign language whose pronunciation sounds perfectly native as judged by phonetically trained native listeners (Bongaerts et al. 1997), may still have an imperfect perceptual representation of the target-language sound system.

The second question we asked is whether the perceptual representation of the vowel systems of early bilinguals would differ from those of late bilinguals. The results indicate that this is indeed the case. Whereas we found little, if any, difference between the perceptual representation of the vowel systems of monolingual listeners and those of early bilinguals, the late bilinguals differed in several respects. Not only did the location of at least one vowel differ between early and late bilinguals (and monolinguals) in both English (L_2) and Dutch (L_1), see above, we also found that the late bilinguals' vowel categories, in both L_1 and L_2) tend to be more narrowly defined, i.e. with less allowed deviation from the category prototype (smaller spreading ellipses) and with sharper boundaries (better labeling consistency). We conclude, then, that early (or natural) bilinguals and late (artificial) bilinguals have different mental representations of the vowel systems of both languages they command.

References

Bongaerts, Th., Summeren, C. van, Planken, B., Schils, E. (1997) Age and ultimate attainment in the pronunciation of a foreign language. *Studies in second language acquisition*, 19, 447-465.

Collins, B.S., Mees, I. (1984) *The sounds of English and Dutch*. Leiden: E.J.Brill/Leiden University Press.

Flege, J.E. (1995) Second language speech learning: theory, findings, and problems. In: *Speech perception and linguistic experience: issues in cross-language research* (W. Strange, J.J. Jenkins, eds.), Timonium, MD: York Press, 233-277.

Hermes, D.J. (1988) Measurement of pitch by subharmonic summation. *Journal of the Acoustical Society of America*, 83, 257-264.

Heuven, V.J. van, Houten, J.E. van (1989) Vowel labeling consistency as a measure of familiarity with the phonetic code of a language or dialect. In *New methods in dialectology* (M.E.H. Schouten, P.Th. van Reenen, eds.), Dordrecht: Foris, 121-130.

Jongman, A., Wade, T. (2007) Perception of native and foreign-accented English by native and non-native listeners. In: *Language experience in second language speech learning* (O.-S. Bohn, M.J. Munro, eds.), Amsterdam/Philadelphia: John Benjamins, 135-150.

Kooij, J.G., Oostendorp, M. van (2003). *Fonologie. Uitnodiging tot de klankleer van het Nederlands [Phonology. Invitation to the sound structure of Dutch]*, Amsterdam: Amsterdam University Press.

Kuhl, P.K., Iverson, P. (1995) Linguistic experience and the 'perceptual magnet effect'. In: *Speech perception and linguistic experience: issues in cross-language research* (W. Strange, J.J. Jenkins, eds.), Timonium, MD: York Press, 121-154.

Lenneberg, E.H. (1967) *Biological foundations of language*. New York: Wiley.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M. (1967) Perception of the speech code. *Psychological Review*, 74, 431-461.

Schouten, M.E.H. (1975) *Native-language interference in the perception of second-language Vowels,* Doctoral dissertation, Utrecht University.

Traunmüller, H. (1990) Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88, 97-100.

Wang, H. Heuven, V.J. van (2006) Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. In *Linguistics in the Netherlands 2006* (J.M. van de Weijer, B. Los, eds.), Amsterdam/Philadelphia: John Benjamins, 237-248.

Willems, L.F. (1987) Robust formant analysis for speech synthesis applications. In: *European Conference on Speech Technology* (J. Laver, M.A Jacks, eds.), Edinburgh: CEP Consultants, 250-253.