Pre-print version of chapter 5. To appear as:

Vincent J.van Heuven (2017). Perception of English and Dutch checked vowels by early and late bilinguals. Towards a new measure of language dominance. In S.E. Pfenninger & J. Navracsics (eds) *Future research directions for Applied Linguistics* (Second Language Acquisition nr. 109). Bristol, Buffalo, Toronto: Multilingual Matters, 73-98.

Chapter 5

Perception of English and Dutch checked vowels by early and late bilingual. Towards a new measure of language dominance

VINCENT J. VAN HEUVEN

Introduction[1]

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 (e.g. Piske et al., 2001; Flege et al., 2006).The adult second or foreign language (henceforth L_2*) learners acquire the new language after that their 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 (e.g. Kuhl & Iverson, 1995). According to the motor theory (MT) of speech perception (e.g. Liberman et al., 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 & Iverson, 1995). Together, the mechanisms explained by MT and NLM conspire to make it difficult for adults to fully acquire a new sound system. The difference in success of the acquisition of the L2 phonology between young and older learners is not fully understood and subject to ongoing debate. Loss of neural plasticity with age has been advanced as one cause that prevents learners from modifying the perceptual categories that were developed in the early stages of life (e.g. Lenneberg, 1967; see also Navracsics & Sáry, and Pfenninger & Singleton, this volume). However, L2 acquisition at a young age (in a natural, non-supervised learning environment) was also found to be strongly correlated with intensity of exposure, which by itself could explain the greater success of young learners (e.g. Flege et al., 2006; DeKeyser, 2013).*

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, despite the fact that the learning process did not involve early L2 exposure. For instance, Bongaerts et al. (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).*

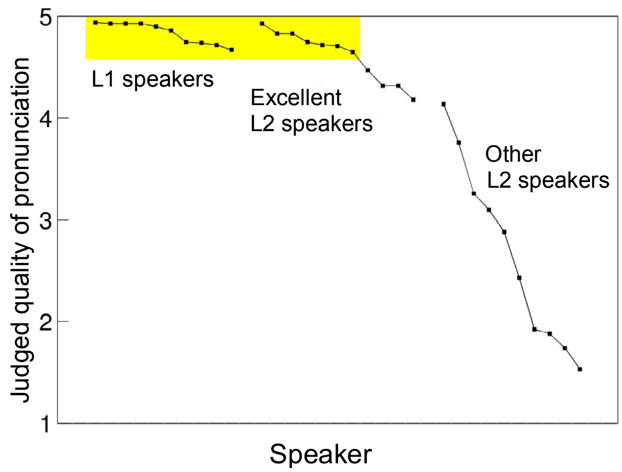


Figure 1 *Quality of British-English pronunciation rated by native English listeners (5: 'No accent at all, definitely native' – 1: 'Very strong foreign accent, definitely nonnative') of three samples of speakers: native speakers of British English (L_1), excellent L_2 speakers with Dutch as the mother 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).*

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 (see also Navracsics & Sáry, and Pfenninger & Singleton, this volume). 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 a (virtually) 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 (typically between the age of four to six) 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 'natural' or 'early' bilinguals. 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 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.*

Within the category of early bilinguals a further subdivision has been proposed depending on language dominance. It is hardly ever the case that even an early bilingual has a perfectly balanced command of both languages. Rather it will be the case that one language is more often heard or produced than the other, so that tasks in this language can be carried out faster and with fewer errors. For instance, Sebastián-Gallés and Soto-Faraco (1999) had Spanish dominant and Catalan dominant early bilingual Spanish-Catalan listeners identify members of binary vowel ad consonant contrasts in Catalan (/e-ɛ/, /o-ɔ/, /ʃ-ʒ/, /s-z/). The former three contrasts are unique to Catalan and are not part of the Spanish sound system. The fourth contrast, /s-z/, however, occurs in both languages. The results showed that the Catalan-dominant early bilinguals differentiated the Catalan contrasts more accurately and earlier in the time course of the stimulus than the Spanish-dominant counterparts. Crucially, no difference between the two groups was found for the shared /s-z/ contrast. The Spanish-dominant bilinguals had been monolingual in Spanish during the first 24 months of their lives. Likewise, the Catalan-dominant bilinguals did not hear any Spanish before the age of three. These results indicate that even in near-perfect bilinguals early exposure to one language over the other may have a lasting influence on the sharpness with which the sound categories of the two languages are defined perceptually.

It is well-known 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 & 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 & 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 may 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 is 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.

The first question I address in the present paper is whether such excellent L_2 speakers as exemplified in Bongaerts et al. (1997, see also Figure 1) have a mental representation of the English vowel sound system that is equal to that of native L_1 speakers of British English. I would hypothesize 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.

Secondly, we would like to know to what extent the late bilingual listeners differ in theor mental representation of the English and Dutch vowel systems from monolingual listeners, not only in terms of the location of the prototypes in the vowel space but also in terms of the well0definedness of the vowel categories.

Thirdly, we want to know if language dominance is reflected in the relative strengths of the vowel categories in the early bilingual listeners. Answering this question requires the availability of instruments ot determine (the degree of) language dominance in bilinguals. This chapters aims to contribute to the discipline by proposing – and provisionally testing – a new method for estimating language dominance in the phonology of bilinguals. In the next section we will introduce a short-cut method to establish the well-definedness (in lieu of *establishing boundary widths between categories) of the sound categories in the phonological systems of bilinguals and show how the measure can be transformed into an estimate of the relative dominance of one language over the other in the mental representation of the vowel systems in a bilingual listener.*

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.[2] *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 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.[3] *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, œ, 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 & 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, I decided to establish the boundary widths by a shortcut method 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. Language dominance can then be established by determining the ratio between the mean consistency observed in sound categories of the the two competing languages.*

Methods

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*,* œ*,* 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 a so-called robust analysis of formant frequencies and bandwidths, using the split-Levinson algorithm for Linear Predictive Coding (Willems, 1987) with a 25-ms window and a 10-ms frame shift. Fundamental frequency (vocal pitch) was computed by the method of subharmonic summation described in Hermes (1988). The same algorithm was used to decide* *for each analysis frame whether it was voiced or voiceless. 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]
F_1	300	980	790	360
F_2	2640	1620	1210	1157
B_1	30	98	79	36
B_1	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).[4] 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 Table 2), 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.

Table 2 *The vowel stimulus space, defined by 10 F2 (front-back) by 9 F1 (close-open) values. Step number, Hertz and Bark values are indicated. Three combinations were omitted from the stimulus set.*

				Second formant F2 (front – back)									
	Step #			1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
	Hertz			2732	2464	2221	2001	1800	1617	1451	1299	1160	1033
	Bark			15.0	14.3	13.6	12.9	12.2	11.5	10.8	10.1	9.4	8.7
First formant F1 height)	1.	308	3.2	1, 1	2, 1	3, 1	4, 1	5, 1	6, 1	7, 1	8, 1	9, 1	10, 1
	2.	381	3.9	1, 2	2, 2	3, 2	4, 2	5, 2	6, 2	7, 2	8, 2	9, 2	10, 2
	3.	459	4.6	1, 3	2, 3	3, 3	4, 3	5, 3	6, 3	7, 3	8, 3	9, 3	10, 3
	4.	541	5.3	1, 4	2, 4	3, 4	4, 4	5, 4	6, 4	7, 4	8, 4	9, 4	10, 4
	5.	628	6.0	1, 5	2, 5	3, 5	4, 5	5, 5	6, 5	7, 5	8, 5	9, 5	10, 5
	6.	722	6.7	1, 6	2, 6	3, 6	4, 6	5, 6	6, 6	7, 6	8, 6	9, 6	10, 6
	7.	822	7.4	1, 7	2, 7	3, 7	4, 7	5, 7	6, 7	7, 7	8, 7	9, 7	10, 7
	8.	932	8.1	1, 8	2, 8	3, 8	4, 8	5, 8	6, 8	7, 8	8, 8	9, 8	
	9.	1050	8.8	1, 9	2, 9	3, 9	4, 9	5, 9	6, 9	7, 9	8, 9		
	#	Hz	Bk										

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.

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 was 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 native speakers of each language, (ii) they made regular use of both languages, and (iii) they subjectively perceived themselves as bilingual. 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.

An attempt was made to post hoc categorize early bilinguals in terms of English versus Dutch language dominance. This was done on the basis of a written questionnaire, in which the participants provided details with respect to their language acquisition, use and preference.

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 on 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').[5]

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).

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 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 and Van Houten (1989).

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).

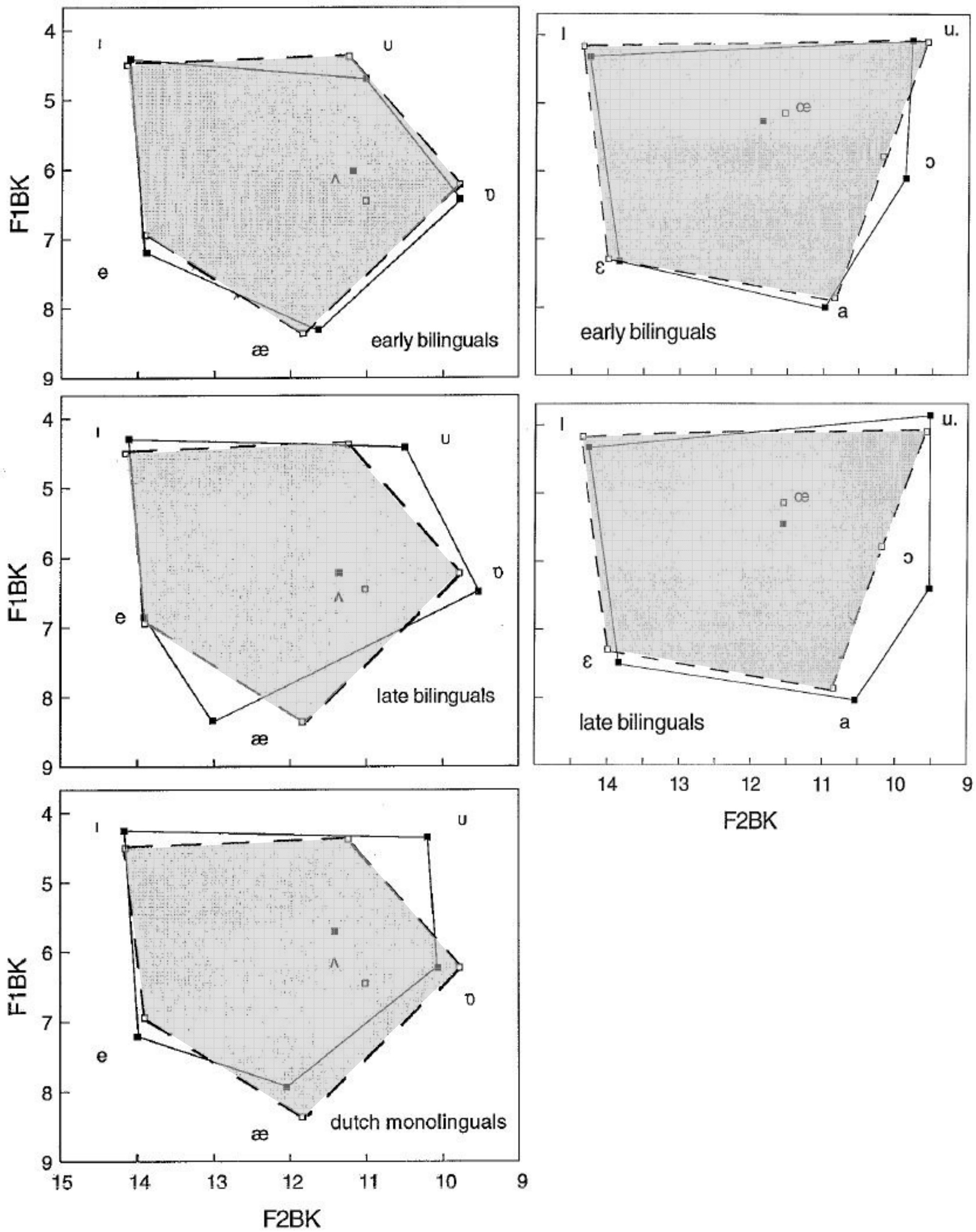


Figure 2 *Vowel centroids in the F_1-by-F_2-space (in Bark) by early bilinguals (top row), late bilinguals (middle row) and Dutch monolinguals, as perceived in the English (left column) and Dutch (right column) response modes. Monolingual English listeners (left column) and Dutch (right column) listeners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

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/.

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 terms 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. The origin of this might lie in the fact 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 than 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.

Labeling consistency

A first exploration

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 illustrated in Table 3.

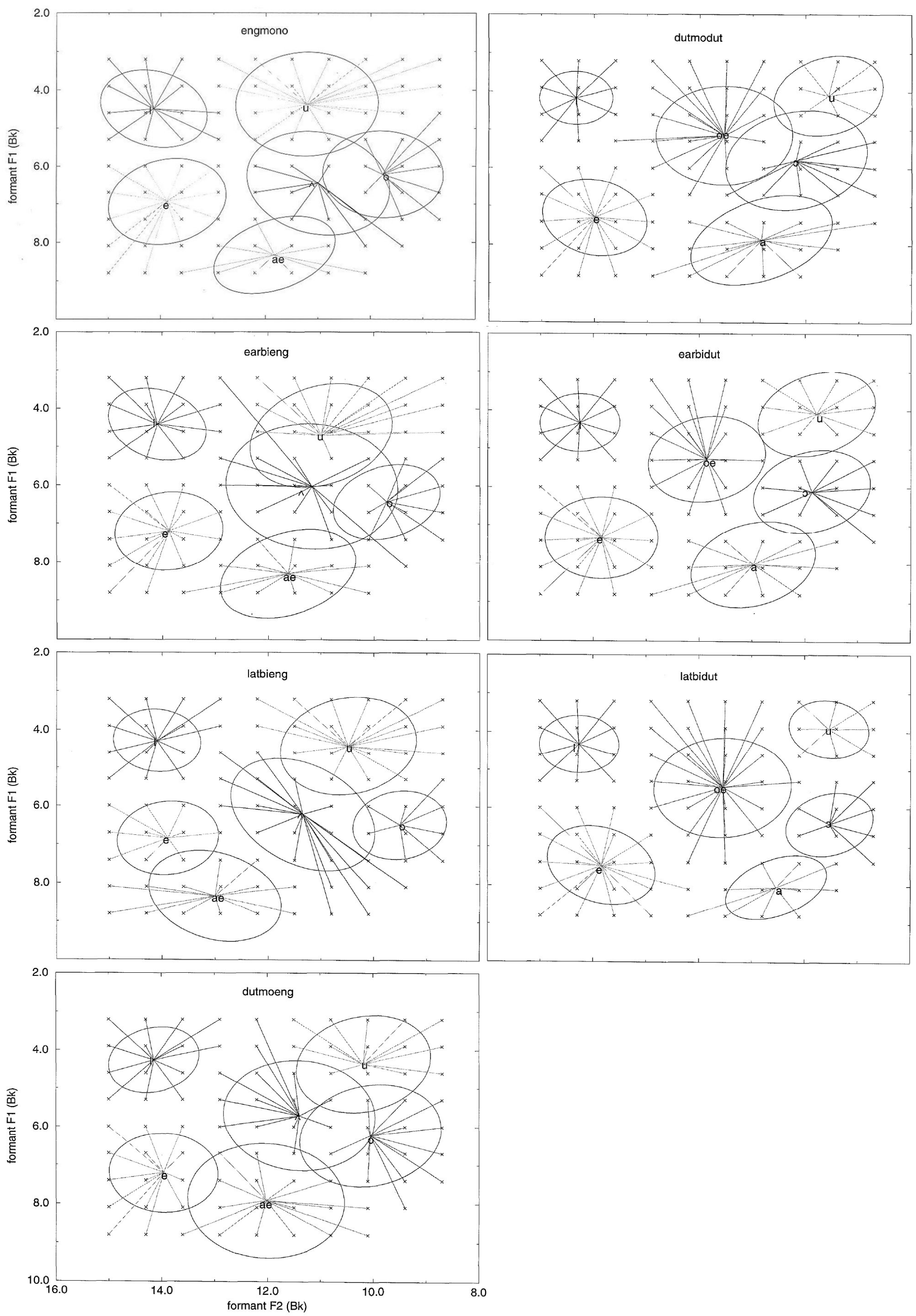


Figure 3 *Centroids and dispersion ellipses (± 1 standard deviation) in the F_1-by-F_2 space (Bark) for six perceived short vowels in the English listening mode (left column) and in the Dutch listening mode (right column) by monolingual native listeners (top row), early*

bilinguals (second row), late bilinguals (third row) and Dutch L_2 listeners of English (bottom).

Table 3 *Consistency index per language (listening mode) and subject group.*

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

First, Table 3 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/ in English 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 (.70). 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. The early bilinguals tend to be less consistent in either language than the monolinguals. 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 3.*

The consistency data were submitted to a Repeated Measures Analysis of Variance (RM-ANOVA) with stimulus language (Dutch, English) as a within-subject factor and with type of bilingualism (early, late, poor) as a between-subjects factor. Since the RM-ANOVA requires correlated data, the English monolingual group could not be included (since no Dutch stimuli were presented to these listeners). The RM-ANOVA indicates that the effect of stimulus language is significant, $F(1, 27) = 14.2$ *(*$p = .001$*,* $p\eta^2 = .344$*). However, the effect of listener group fails to reach significance,* $F(2, 27) = 1.1$ *(*$p = .363$*,* $p\eta^2 = .072$*) as does the language × group interaction,* $F(2, 27) = 1.1$ *(*$p = .362$*,* $p\eta^2 = .073$*). A one-way ANOVA on the consistency indexes obtained in the English presentation mode by all four listener groups, i.e. including the English monolinguals, again fails to show significance for listener group,* $F(3, 41) < 1$*.*

Consistency as a correlate of language dominance

As a last exercise we may see if there is any value in the consistency index as a tool to differentiate bilinguals in terms of language dominance. Given that the consistency index is significantly higher for Dutch than for English we defined for each listener a difference value Δ *as the index for Dutch minus the index for English. Table 4 lists, among many other variables, the index scores for English, Dutch and the* Δ *for each of the eight early bilinguals.*

Table 4 *Consistency scores (raw and z-transformed, see text) and variables relating to language acquisition for eight early (natural) bilingual listeners.*

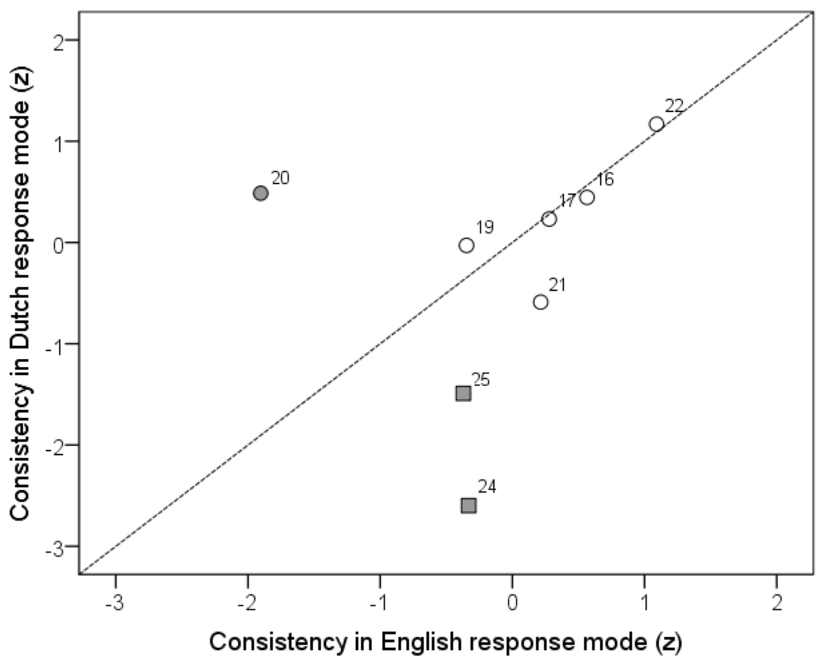
	ID number	16	17	19	20	21	22	24	25
1.	Gender	F	F	F	F	F	F	M	M
2.	Age	22	22	23	24	23	26	21	24
3.	Father	E	E	E	E	D	D	D	E
4.	Mother	E	E	D	D	D	D	D	E
5.	Home child	D/E	D/E	D/E	E	D/E	D/E	D/E	D/E
6.	Prim school	D/E	D	D	D	D	D/E	D/E	D/E
7.	Sec school	D	E	D	E	D	E	E	E
8.	Home now	E	D/E	D/E	E	D/E	D/E	D/E	D/E
9.	Work now	E	D	E	D	D	E	E	E
10.	Preference	E	D/E	E	E	D/E	E	E	E
11.	Age D	5-10	1.5-12	1-12	5-12	4-8	0-2, 8-12	9-12	6-12
12.	Age E	0-5, 10-10	0-1.5	≥ 12	0-5	0-4, 8-12	2-8	0-9	0-6
13.	Years D	8	20	23	19	23	26	21	24
14.	Years E	22	22	23	24	19	25	21	24
15.	Consist D	82.56	80.46	77.91	82.7600	72.41	89.66	52.8700	63.2200
16.	Consist E	77.01	73.26	65.06	44.8300	72.41	83.91	65.4300	64.7100
17.	ΔConsist	5.55	7.2	12.85	37.9300	0	5.75	−12.5600	−1.4900
18.	Consist D(z)	0.4453	0.2313	−0.0290	0.4657	−0.5891	1.1689	−2.5815	−1.5267
19.	Consist E(z)	0.5649	0.2784	−0.3470	−1.8908	0.2141	1.0911	−0.3186	−0.3740

20.	*ΔConsist(z)*	*−0.1196*	*−0.047*	*0.318*	*2.3565*	*−0.8032*	*0.0778*	*−2.2629*	*−1.1527*

3: L1 of father, 4: L1 of mother, 5: Language used at home when child, 6: Language of instruction in primary school, 7: Language of instruction in secondary school, 8: Language used at home at the time of the experiment, 9: Language used in public domain at the time of the experiment, 10: Preferred language in present daily life, 11: age when exposed mainly to Dutch, 12: Age when exposed mainly to English, 13: number of years using mainly Dutch, 14: number of years using mainly English, 15: Consistency index in Dutch, 16: Consistency index in English, 17: Difference between consistency in Dutch and English. 18-19-20 are the z-transormed versions of 15-16-17.

We have seen that the consistency scores were better overall when the stimuli were presented and responded to in the Dutch mode. For a fair comparison of the consistency scores in the two modes by the early bilinguals, some form of normalisation would therefore be needed. Moreover, the proper comparison should be between the early bilinguals in each response mode in separate in groups combined with monolingual listeners responding in the same language mode. Accordingly, z-transformation was applied to the English consistency scores for the early bilinguals combined with the group of monolingual English respondents. Similarly, z-scores were computed for the Dutch consistency indexes obtained by the members of the combined early bilingual plus monolingual Dutch listeners. In both arrays of z-scores the early bilinguals have lower scores on average than their monolingual comparison groups, but the difference is not significant (or else the effect of listener group should have been significant in the above analyses of the raw scores).

Language dominance should correlate with the difference between the z-scores obtained by the early bilinguals in the Dutch versus English listening mode. Figure 4 *displays the z-scores of the consistency z-scores obtained in the English response mode plotted against the z-scores in the Dutch mode for each of the eight early bilinguals, who are identified in the scatterplot by arbitrary two-digit numbers. Five out of eight early bilinguals have roughly the same consistency in their vowel labelling in both language modes. These would then be balanced bilinguals. Two individuals have clearly better consistency in English than in Dutch. They appear below the 45-degree reference line and are marked by filled squares. These would be clear cases of English-dominant bilinguals. It should be observed that these two individuals (nrs. 24 and 25) do not have exceptionally good*



consistency within their peer group. In fact, they belong to the poorer performers. The point is that within themselves they do better in English than in Dutch. Listener 20 (marked by a filled circle) represents the other extreme of the dominance continuum. This person has by far the poorest consistency in English of all eight early bilinguals but is in the upper half of the peer group in the Dutch response mode.

Figure 4 *Consistency index (z-scores) in English (horizontal axis) and Dutch (vertical axis) response mode for eight early bilinguals. The 45-degree reference line represents perfect balance of English and Dutch consistency (for further details see text).*

Now let us see if there is anything in the language acquisition history of the eight early bilinguals that might explain the differences in language dominance observed in the consistency indexes. The top part of Table 4 summarizes the information gathered on the early bilinguals by means of a questionnaire. The bottom part of the table presents the individual consistency scores in English and Dutch, the difference between them as well as the z-transformed values.

The most striking difference that distinguishes listeners 24 and 25 from the other early bilinguals is the fact that they spent the first nine and six years respectively of their lives in England, and only then crossed the North Sea to live in The Netherlands. These two listeners also turn out to be the only males in the sample, but it is hard to imagine that this would be causally related to their English language dominance. It is probably not accidental that the English dominance of the male who spent the first nine years in England is considerably stronger than that of the person who left England at the age of six. It suggests that spending time with English-speaking peers in an English school environment is of decisive importance here.

Listener 20 turned out to be a bit of a problem case. She spent the first five years of her life in England, and yet she obtained by far the lowest consistency score for English, which makes her the most Dutch-dominant listener in the sample. Reconstructing her language past from the table, it would appear that her mother spoke Dutch to her (or English with a Dutch accent), while the language the parents used to communicate with each other was English – which is understandable since the family lived in England at the time. Right at *the time when the child was about to go to school, the family moved to the Netherlands. This child was probably never exposed to a rich English-speaking school environment such as the one listeners 24 and 25 experienced. During primary school (until the age of 12) she grew up with Dutch-speaking peers. The family continued to live in the Netherlands (given that the young woman speaks Dutch outside the home). Most likely, the language she used at home continued to be English (with a Dutch-accented mother), while she attended an international school in the Netherlands, where the language of instruction was English but spoken by a great many non-natives. This, to my mind, leads to poorly defined perceptual norms for the English vowels – a hypothesis which is supported by the low consistency score in English.*

It would seem, in sum, that the (z-transformed) difference in labelling consistency for vowel categories would be a useful tool in determining language dominance in a bilingual speaker. It may well be, however, that a more sensitive measure of language dominance can be defined if we were to take specific vowel categories into account. Some vowels may be more prone to perceptual confusion than others. By zooming in on the poorly defined vowel categories we may enhance the sensitivity of the instrument.

Consistency in the perception of poorly defined vowel categories

Table 5 provides a full survey of the vowels provided by the listeners at the first and second presentation of the same stimulus. For ease of presentation, each of the four listener groups are listed separately, broken down further by the language mode of the stimulus presentation (Dutch versus English). Note that the English monolingual listeners were never asked to classify the vowels in the Dutch mode. Dutch monolinguals are not really monolinguals but always have a basic knowledge of English as a result of six to eight years of English lessons at school; they could therefore be asked to respond in the English as well in the Dutch mode.

The numbers in Table 5 are row percentages, and add up to 100. The actual number of responses is specified for each row under N.

The table shows, again, that the high tense vowels were impopular response categories in Dutch. Front /i/ was the least frequent response vowel in this group, followed by /y/. The response frequency of /u/ is better, although it still falls short of the non-tense vowel types. When comparing the performance of the groups, it would seem better to disregard the results for the high tense front vowels of Dutch. In that case there are no striking differences in the performance of the three Dutch groups (i.e. early bilingual, late bilingual and monolingual).

The most striking difference is seen in the English response mode, viz. between the consistency of the early bilinguals and that of the monolinguals. Although the mean difference is insignificant, as shown above, we now see that the English monolinguals have more consistent responses than the early bilinguals for the vowels /æ, ʌ, ʊ/, while the reverse is seen for the vowels /ɪ, ɒ/. Especially the inconsistency concerning the pair /ʌ, ʊ/ seems characteristic for the early bilinguals, with confusions (a change of response category between first and second presentation from one to the other) of 37 and 18 per cent. This suggests that the perceptual categories for these two vowels are less well defined in the mind the early bilinguals than for the monolinguals. Similarly, the vowel /æ/ is often confused with its neighbours in the vowel space, i.e. /e/ and /ʌ/. The triplet /æ, ʌ, ʊ/ is known to present pronunciation problems for Dutch learners of English (see e.g. Wang & Van Heuven, 2006). The vowel with the weakest definition is the central vowel /ʌ/. This is also the vowel with the largest dispersion ellipse in the corresponding panel in Figure 3. The competition between neighboring vowel categories would appear to be limited to the mid-range of F2 (front-back) values. As a result, there is vacillation or uncertainty between the lower vowel /æ/ and the *higher vowel /ʊ/. This perceptual uncertainly does not arise in the case of the more peripheral vowels /ɪ, e, ɒ/, which have more extreme F2 values, i.e. higher in the case of /ɪ, e,/ and lower in the case of /ɒ/. Moreover, the vowels English vowels /ɪ, e, ɒ/ would be rather good and readily identifiable examplars of Dutch /ɪ, ɛ and ɔ/, respectively. It would appear, then, that early bilingualism comes at a (small) price in that the perceptual representation of some vowels is poorer than is seen for monolingual speakers.*

Table 5 *Relative frequency (% of row total) of vowel response on first versus second presentation of the same stimulus, broken down by Language group by Response mode. Consistent responses are on the main diagonal (bolded in shaded cells). Total number of valid responses in listed under N.*

Language group			*Vowel category responded (first presentation)*																
			Mode: English							*Mode: Dutch*									
			ɪ	e	æ	ʌ	ɒ	ʊ	*N*		ɪ	ɛ	ɑ	ɔ	œ	i	y	u	*N*
Monolingual English (N = 15)	*Second presentation*	ɪ	**69**	*18*		*2*	*4*	*7*	*141*	ɪ									
		e	*3*	**80**	*7*	*5*	*2*	*4*	*271*	ɛ									
		æ		*7*	**77**	*11*	*3*	*2*	*131*	ɑ									
		ʌ	*1*	*9*	*10*	**51**	*8*	*20*	*178*	ɔ									
		ɒ	*1*	*1*	*7*	*15*	**62**	*14*	*163*	œ									
		ʊ	*4*	*3*	*1*	*12*	*7*	**74**	*382*	i									
										y									
										u									
Bilingual	*Sec on*	ɪ	**80**	*10*	*2*	*1*	*1*	*7*	*104*	ɪ	**72**	*10*	*1*	*1*	*9*	*2*	*4*	*1*	*93*

e	*4*	**82**	*3*	*2*	*3*	*6*	*140*	ɛ	*1*	**85**	*6*	*2*	*4*		*1*	*1*	*142*
æ	*2*	*18*	**45**	*20*	*12*	*3*	*60*	ɑ		*4*	**84**	*6*	*2*	*1*	*1*	*1*	*96*
ʌ	*3*	*4*	*11*	**40**	*5*	*37*	*114*	ɔ	*1*	*5*	*10*	**71**	*6*		*2*	*5*	*108*
ɒ		*2*		*7*	**84**	*7*	*90*	œ		*5*	*1*	*3*	**81**		*5*	*5*	*149*
ʊ	*4*	*3*	*3*	*18*	*4*	**69**	*175*	i	*15*		*15*	*8*	*8*	**15**	*31*	*8*	*13*
								y		*3*		*8*	*29*		**45**	*16*	*38*
								u		*2*	*2*	*7*	*9*	*4*	*9*	**67**	*55*

Table 5 Continued

Bilingual Late (N = 7)	ɪ	**80**	*8*		*7*		*5*	*92*	ɪ	**74**	*15*			*7*	*2*		*1*	*85*
	e	*3*	**70**	*18*	*8*	*1*		*96*	ɛ	*2*	**92**	*1*	*1*	*5*				*129*
	æ		*16*	**74**	*5*	*2*	*3*	*61*	ɑ		*15*	**59**	*7*	*18*				*71*
	ʌ	*4*	*1*	*10*	**75**	*4*	*6*	*159*	ɔ			*6*	**78**	*16*				*63*
Second presentation	ɒ				*14*	**84**	*2*	*58*	œ	*1*	*2*	*2*	*1*	**91**		*2*	*1*	*179*
	ʊ		*3*	*1*	*20*	*5*	**72**	*143*	i	*50*					**50**			*4*
									y	*5*				*30*	*5*	**55**	*5*	*20*
									u					*9*	*2*	*3*	**86**	*58*
Monolingual Dutch (N = 15)	ɪ	**76**	*9*	*2*	*8*	*1*	*4*	*193*	ɪ	**77**	*14*	*1*		*8*	*1*			*155*
	e	*5*	**68**	*18*	*8*	*2*		*194*	ɛ	*1*	**92**	*3*		*4*				*256*
	æ	*1*	*21*	**54**	*17*	*6*	*2*	*179*	ɑ		*3*	**85**	*5*	*6*		*1*		*209*
	ʌ	*3*	*2*	*5*	**70**	*6*	*14*	*424*	ɔ		*1*	*15*	**62**	*18*		*1*	*2*	*136*
Second presentation	ɒ		*3*	*7*	*22*	**56**	*12*	*164*	œ	*2*	*2*	*3*	*5*	**85**		*1*	*3*	*415*
	ʊ	*2*			*28*	*2*	**68**	*144*	i					*38*	**50**		*13*	*8*
									y	*5*	*5*	*10*		*67*		**10**	*5*	*21*
									u		*2*	*1*	*6*	*20*	*2*	*7*	**63**	*105*

In conclusion to this section it should be observed that the inconsistency patterns seen in Table 5, i.e. the vacillation in the responses to the repeated presentation of the same stimulus vowel sound, are a way to qualify the degree of overlap between the dispersion ellipses in Figure 3. Obviously, the more two ellipses overlap in Figure 3, the more strongly the vacillation between the two vowel categories involved.

Conclusion

In this chapter 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 only small differences 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.*

*Early bilinguals do show subtle differences in their mental representation of the vowels of English (more than of Dutch) in that the boundaries between /*æ, ʌ, ʊ*/ are less sharply defined than those of /*ɪ, ɛ, ɔ*/. We may note that this observation runs counter to the hypothesis we formulated in the introduction that the well-definedness of the vowels within the same language system would be uniform. This aspect of vowel representation in the mind of early bilinguals deserves more future research. The consistency analysis that we advocated in this study may turn out to be useful tool to analyse the stability and well-definedness of sound categories in monolingual and bilingual language users. We found evidence that the consistency index may serve as a diagnostic of language dominance in (early) bilinguals. The relative difference in labelling consistency between the two languages of the listener seemed to correlate well with the individual's language acquisition history. The sample of early bilinguals, however, was small. More and more systematic research is needed in this area in order to develop a better understanding of the potential of the labelling consistency index as a correlate of language dominance.*

Notes

1. This chapter appeared in abridged form as Van Heuven, Broerse & Pacilly (2011). It is based on an MA thesis written by Nicole Broerse author under the supervision of the present author. I thank Ing. Jos Pacilly for his invaluable help in generating the stimulus materials used in the experiment. I am also grateful to professor Antonella Sorace for her part in the supervision of the work done at Edinburgh University.

2. 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.

3. 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.

4. 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 × F) / (1960 + F)] – 0.53, where F represents the measured formant frequency in Hertz.

5. The response vowels were spelled as is usual in English and Dutch orthography in closed syllables. Capital versus lower case *u* were used to differentiate the English vowels in *putt* and *put*, respectively. Although this is somewhat contrived, the use of examples turned out to be adequate to resolve any remaining ambiguity.

References

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

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

Flege, J. (1995) Second language speech learning: theory, findings, and problems. In: W. Strange and J. J. Jenkins (eds) Speech Perception and Linguistic Experience: Issues in Cross-language Research (pp. 233-277). Timonium, MD: York Press.

Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sunge, H. and Tsukada, K. (2006) Degree of foreign accent in English sentences produced by Korean children and adults. Journal of Phonetics 34: 153-175.

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

Heuven, V. van, Broerse, N. and Pacilly, J. (2011), Perception of checked vowels by early and late Dutch/English bilinguals. In: W. Zonneveld, H. Quené and W. Heeren (eds) Sound and Sounds. Studies Presented to M.E.H. (Bert) Schouten on the Occasion of his 65th Birthday (pp. 103-116). Utrecht: Utrecht Institute of Linguistics OTS.

Heuven, V. van and Houten, J. van (1989) Vowel labeling consistency as a measure of familiarity with the phonetic code of a language or dialect. In M. Schouten and P. van Reenen (eds) New Methods in Dialectology (pp. 121-130). Dordrecht: Foris.

Jongman, A. and Wade, T. (2007) Perception of native and foreign-accented English by native and non-native listeners. In O.-S. Bohn and M. Munro (eds) Language Experience in Second Language Speech Learning (pp. 135-150). Amsterdam/Philadelphia: John Benjamins.

Kooij, J. and 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. and Iverson, P. (1995) Linguistic experience and the 'perceptual magnet effect'. In: W. Strange and J. Jenkins (eds) Speech Perception and Linguistic Experience: Issues in Cross-language Research (pp. 121-154). Timonium, MD: York Press.

Lenneberg, E. (1967) Biological Foundations of Language. New York: Wiley.

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

Piske, T., MacKay, I. and Flege, J. (2001) Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics* 29: 191-215

Schouten, M. (1975) Native-language Interference in the Perception of Second-language Vowels. Doctoral dissertation, Utrecht University.

Sebastián-Gallés, N. and Soto-Faraco, S. (1999) Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition* 72: 111-123.

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

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

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

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]
F_1	300	980	790	360
F_2	2640	1620	1210	1157
B_1	30	98	79	36
B_1	264	162	121	116

Table 2 *The vowel stimulus space, defined by 10 F2 (front-back) by 9 F1 (close-open) values. Step number, Hertz and Bark values are indicated. Three combinations were omitted from the stimulus set.*

				Second formant F2 (front – back)									
		Step #		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
		Hertz		2732	2464	2221	2001	1800	1617	1451	1299	1160	1033
		Bark		15.0	14.3	13.6	12.9	12.2	11.5	10.8	10.1	9.4	8.7
First formant F1 height	1.	308	3.2	1, 1	2, 1	3, 1	4, 1	5, 1	6, 1	7, 1	8, 1	9, 1	10, 1
	2.	381	3.9	1, 2	2, 2	3, 2	4, 2	5, 2	6, 2	7, 2	8, 2	9, 2	10, 2
	3.	459	4.6	1, 3	2, 3	3, 3	4, 3	5, 3	6, 3	7, 3	8, 3	9, 3	10, 3
	4.	541	5.3	1, 4	2, 4	3, 4	4, 4	5, 4	6, 4	7, 4	8, 4	9, 4	10, 4
	5.	628	6.0	1, 5	2, 5	3, 5	4, 5	5, 5	6, 5	7, 5	8, 5	9, 5	10, 5
	6.	722	6.7	1, 6	2, 6	3, 6	4, 6	5, 6	6, 6	7, 6	8, 6	9, 6	10, 6
	7.	822	7.4	1, 7	2, 7	3, 7	4, 7	5, 7	6, 7	7, 7	8, 7	9, 7	10, 7
	8.	932	8.1	1, 8	2, 8	3, 8	4, 8	5, 8	6, 8	7, 8	8, 8	9, 8	
	9.	1050	8.8	1, 9	2, 9	3, 9	4, 9	5, 9	6, 9	7, 9	8, 9		
	#	Hz	Bk										

Table 3 *Consistency index per language (listening mode) and subject group.*

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

Table 4 *Consistency scores (raw and z-transformed, see text) and variables relating to language acquisition for eight early (natural) bilingual listeners.*

	ID number	16	17	19	20	21	22	24	25
1.	Gender	F	F	F	F	F	F	M	M
2.	Age	22	22	23	24	23	26	21	24
3.	Father	E	E	E	E	D	D	D	E
4.	Mother	E	E	D	D	D	D	D	E
5.	Home child	D/E	D/E	D/E	E	D/E	D/E	D/E	D/E
6.	Prim school	D/E	D	D	D	D	D/E	D/E	D/E
7.	Sec school	D	E	D	E	D	E	E	E
8.	Home now	E	D/E	D/E	E	D/E	D/E	D/E	D/E
9.	Work now	E	D	E	D	D	E	E	E
10.	Preference	E	D/E	E	E	D/E	E	E	E
11.	Age D	5-10	1.5-12	1-12	5-12	4-8	0-2, 8-12	9-12	6-12
12.	Age E	0-5, 10-10	0-1.5	≥ 12	0-5	0-4, 8-12	2-8	0-9	0-6
13.	Years D	8	20	23	19	23	26	21	24
14.	Years E	22	22	23	24	19	25	21	24
15.	Consist D	82.56	80.46	77.91	82.7600	72.41	89.66	52.8700	63.2200
16.	Consist E	77.01	73.26	65.06	44.8300	72.41	83.91	65.4300	64.7100
17.	ΔConsist	5.55	7.2	12.85	37.9300	0	5.75	−12.5600	−1.4900
18.	Consist D(z)	0.4453	0.2313	−0.0290	0.4657	−0.5891	1.1689	−2.5815	−1.5267
19.	Consist E(z)	0.5649	0.2784	−0.3470	−1.8908	0.2141	1.0911	−0.3186	−0.3740
20.	ΔConsist(z)	−0.1196	−0.047	0.318	2.3565	−0.8032	0.0778	−2.2629	−1.1527

3: L1 of father, 4: L1 of mother, 5: Language used at home when child, 6: Language of instruction in primary school, 7: Language of instruction in secondary school, 8: Language used at home at the time of the experiment, 9: Language used in public domain at the time of the experiment, 10: Preferred language in present daily life, 11: age when exposed mainly to Dutch, 12: Age when exposed mainly to English, 13: number of years using mainly Dutch, 14: number of years using mainly English, 15: Consistency index in Dutch, 16: Consistency index in English, 17: Difference between consistency in Dutch and English. 18-19-20 are the z-transormed versions of 15-16-17.

Table 5 *Relative frequency (% of row total) of vowel response on first versus second presentation of the same stimulus, broken down by Language group by Response mode. Consistent responses are on the main diagonal (bolded in shaded cells). Total number of valid responses in listed under N.*

Language group			*Vowel category responded (first presentation)*																
			Mode: English							*Mode: Dutch*									
			ɪ	e	æ	ʌ	ɒ	ʊ	*N*		ɪ	ɛ	ɑ	ɔ	œ	i	y	u	*N*
Monolingual English (N = 15)	*Second presentation*	ɪ	**69**	*18*		*2*	*4*	*7*	*141*	ɪ									
		e	*3*	**80**	*7*	*5*	*2*	*4*	*271*	ɛ									
		æ		*7*	**77**	*11*	*3*	*2*	*131*	ɑ									
		ʌ	*1*	*9*	*10*	**51**	*8*	*20*	*178*	ɔ									
		ɒ	*1*	*1*	*7*	*15*	**62**	*14*	*163*	œ									
		ʊ	*4*	*3*	*1*	*12*	*7*	**74**	*382*	i									
										y									
										u									
Bilingual Early (N = 8)	*Second presentation*	ɪ	**80**	*10*	*2*	*1*	*1*	*7*	*104*	ɪ	**72**	*10*	*1*	*1*	*9*	*2*	*4*	*1*	*93*
		e	*4*	**82**	*3*	*2*	*3*	*6*	*140*	ɛ	*1*	**85**	*6*	*2*	*4*		*1*	*1*	*142*
		æ	*2*	*18*	**45**	*20*	*12*	*3*	*60*	ɑ		*4*	**84**	*6*	*2*	*1*	*1*	*1*	*96*
		ʌ	*3*	*4*	*11*	**40**	*5*	*37*	*114*	ɔ	*1*	*5*	*10*	**71**	*6*		*2*	*5*	*108*
		ɒ		*2*		*7*	**84**	*7*	*90*	œ		*5*	*1*	*3*	**81**		*5*	*5*	*149*
		ʊ	*4*	*3*	*3*	*18*	*4*	**69**	*175*	i	*15*		*15*	*8*	*8*	**15**	*31*	*8*	*13*
										y		*3*		*8*	*29*		**45**	*16*	*38*
										u		*2*	*2*	*7*	*9*	*4*	*9*	**67**	*55*
Bilingual Late (N = 7)	*Second presentation*	ɪ	**80**	*8*		*7*		*5*	*92*	ɪ	**74**	*15*			*7*	*2*		*1*	*85*
		e	*3*	**70**	*18*	*8*	*1*		*96*	ɛ	*2*	**92**	*1*	*1*	*5*				*129*
		æ		*16*	**74**	*5*	*2*	*3*	*61*	ɑ		*15*	**59**	*7*	*18*				*71*
		ʌ	*4*	*1*	*10*	**75**	*4*	*6*	*159*	ɔ			*6*	**78**	*16*				*63*
		ɒ				*14*	**84**	*2*	*58*	œ	*1*	*2*	*2*	*1*	**91**		*2*	*1*	*179*
		ʊ		*3*	*1*	*20*	*5*	**72**	*143*	i	*50*					**50**			*4*
										y	*5*				*30*	*5*	**55**	*5*	*20*
										u					*9*	*2*	*3*	**86**	*58*
Monolingual Dutch (N = 15)	*Second presentation*	ɪ	**76**	*9*	*2*	*8*	*1*	*4*	*193*	ɪ	**77**	*14*	*1*		*8*	*1*			*155*
		e	*5*	**68**	*18*	*8*	*2*		*194*	ɛ	*1*	**92**	*3*		*4*				*256*
		æ	*1*	*21*	**54**	*17*	*6*	*2*	*179*	ɑ		*3*	**85**	*5*	*6*		*1*		*209*
		ʌ	*3*	*2*	*5*	**70**	*6*	*14*	*424*	ɔ		*1*	*15*	**62**	*18*		*1*	*2*	*136*
		ɒ		*3*	*7*	*22*	**56**	*12*	*164*	œ	*2*	*2*	*3*	*5*	**85**		*1*	*3*	*415*
		ʊ	*2*			*28*	*2*	**68**	*144*	i					*38*	**50**		*13*	*8*
										y	*5*	*5*	*10*		*67*		**10**	*5*	*21*
										u		*2*	*1*	*6*	*20*	*2*	*7*	**63**	*105*