The pronunciation of /εi/ by male and female speakers of avant-garde Dutch

Vincent J. van Heuven, Loulou Edelman and Renée van Bezooijen
Leiden University/ULCL (van Heuven) / University of Nijmegen/CLS (Edelman, Bezooijen)

1. Introduction

In de last decade a new variety of Modern Dutch has emerged, which was first noted and commented on by the dialectologist Reker in 1993 in a press release by Radio-Noord (for details see Stroop 1998: 13, 107). Taking a cue from Reker’s observation, the emerging variety has been closely monitored by Stroop, who christened it Polder Dutch. The variety was elegantly described and placed in a wider sociolinguistic perspective in a popular-scientific brochure by Stroop (1998). More recently the variety got an entire website devoted to it (www.hum.uva.nl/poldernederlands), with very thorough documentation — both in Dutch and in English — on the phenomenon.

1.1 The phonetics of the sound change

The new variety differs from the Standard language only in its phonetics. Stroop presents the change as a chain shift, whereby the low-mid diphthongs /εi, œy, au/ are lowered. As a result, the onset of the low-mid diphthongs assumes a position very close to open /a/, so that the three diphthongs are no longer clearly differentiated in their onsets. However, the end points of the diphthongs — which may possibly be lowered as well — still differentiate adequately between the front unrounded /εi/, the front rounded /œy/, and the back rounded /au/.1

It remains unclear from the descriptions provided whether the degree of diphthongization is affected by the sound change. If it is only the onset of the diphthongs that is more open, and the end point remains stable, then the strength of diphthongization (the size of the diphthong trajectory) should have increased.

1

However, if the onset and the endpoint have been lowered together, then the strength of diphthongization should have remained the same. If the size of the diphthong trajectory should have increased, then we would — ceteris paribus — expect the duration of the diphthongs to have increased as well.

Stroop claims that the lowering of /ei/ reflects a natural tendency. In fact, low-mid diphthongs are rare in the world’s languages. Cognates of Dutch /ei/ in the neighbouring languages English and German are fully open diphthongs, e.g., English *wine /waɪn/*, German *Wein /vaɪn/*. The Dutch mid-open diphthongs, especially /ei/ and /eij/, are notoriously difficult for foreign learners of the language. The fully open alternative would be easier to produce on the strength of the argument that the speaker just has to open his mouth as wide as he can to get the vowel right, while more intricate articulatory control would be required for the low-mid variety.

In this paper we concentrate on the phonetic details of the diphthong /ei/ in Polder Dutch. This sound has been advanced as the exponent of the new variety, as is evidenced by the following quotation from Stroop (1998:25, our translation) 2:

‘The most conspicuous feature of Polder Dutch is the pronunciation of the diphthong /ei/, which is spelled as either ei or ij.’

The first aim of this study, then, is to clarify the phonetics of the sound change in so far as it relates to the pronunciation of the diphthong /ei/. Moreover, rather than using the traditional impressionistic phonetic approach, i.e. listening and transcription, we wish to settle the issue using acoustic measures of greater mouth opening and/or stronger diphthongization and lengthening.

1.2 The sociolinguistics of the change

Stroop (1998) claims that the change is typical of (relatively) young, highly educated, progressive Dutch women, who wish to make a statement through speech that they are unconventional and emancipated. The variety is often found among women with high-prestige social positions such as authors, actors, film producers, artists, left-wing politicians (either local or national), high-ranking academics, and pop-singers. It is for this reason that we prefer the use of the term ‘avant-garde’ Dutch for the new variety, rather than Polder Dutch, which in hindsight seems a misnomer. It should be pointed out here that the avant-garde variety is found throughout the country. It is not based on any existing dialect of Dutch, and it has no documented geographic epicenter. So, avant-garde Dutch truly qualifies as a sociolect rather than a dialect or regiolect.

Although the change is led by women, Stroop predicts that men will follow suit.3 It is in fact common for sound changes to be initiated by women (see Labov 2001:366–382 for an elaborate treatment of the issue), and the avant-garde Dutch variety is no exception to the general rule. Van Bezooijen and van den Berg (2001)
have shown that young women, as opposed to older women, identify with the new variety, and clearly more so than their male counterparts. The speakers of the new variety are not aware of the fact that their pronunciation of the language differs in any linguistically relevant way from the standard variety. When accused of using the avant-garde pronunciation their first reaction is denial. The sound change has all the characteristics of what Labov (2001) calls a change from below (see also van Bezooijen, Kroezen and van den Berg, 2002).

The second aim of the present study is to test the claim that the avant-garde variety of Dutch is more widespread among women than among men of similar social status and age.

2. Experimental approach

Vowel quality, and change of vowel quality in diphthongs, can be quantified by measuring the lower resonances in the acoustic signal. Specifically, the center frequency of the lowest resonance of the vocal tract, called first formant frequency or $F_1$, corresponds closely to the articulatory/perceptual dimension of vowel height. For an average male voice, the $F_1$ values range between 200 Hertz (Hz) for a high vowel /i/ to some 800 Hz for a low vowel /a/. The second formant frequency (or $F_2$) reflects the place of maximal constriction during the production of the vowel, i.e., the front vs. back dimension, such that the $F_2$ values range from roughly 2200 Hz for front /i/ down to some 800 Hz for back /u/.

The relationship between the formant frequencies and the corresponding perceived vowel quality is not linear. For instance, a change in $F_1$ from 200 to 300 Hz brings about a much larger change in vowel quality (height) than a numerically equal change from 700 to 800 Hz. Using the Bark-transformation (Rietveld and van Heuven 2001: 371), the perceptually realistic distance between two vowel qualities can be computed from acoustic measurements.

The acoustic characterization of diphthongs compares the formant values computed at the onset and at the offset of the diphthong. The degree of diphthongization is then expressed as the distance between the onset and offset vowel quality. This procedure is equivalent to measuring the length of the arrow that represents the diphthong in a traditional impressionistic vowel diagram.

Unfortunately, formant values measured for the same vowel differ across individuals. The more two speakers differ in shape and size of their vocal tracts, the larger the differences in formant values of perceptually identical vowel tokens. Comparison of formant values is especially hazardous across speakers of the opposed sex. Numerous attempts have been made, therefore, to factor out the speaker-individual component from the formant values such that phonetically identical vowels spoken by different individuals would come out with the same values.
None of these vowel normalization procedures have proven fully satisfactory (Adank, van Heuven and van Hout 1999; Labov 2001:157–164; Rietveld and van Heuven 2001:327–330). In the present problem, however, we need not include the full vowel system of the speakers in the analysis. Since the study is limited to the sound change in /si/ — a front, unrounded vowel — we only require reference vowels that allow us to determine the individual implementation of the front region of the speaker’s vowel space. All that is required, therefore, is a reliable estimation of the speaker’s /i/ (maximally high front vowel) and /a/ (maximally open front vowel).

The second basic point to consider is the sampling of the speakers. Remember that we test the hypothesis that women lead the sound change. It seems imperative, therefore, that we compare groups of male and female speakers that are equivalent in all sociolinguistically relevant aspects, such as socio-economic status and age. The speakers should not be aware of the fact that their speech production is being recorded for linguistic analysis, and their speech should be non-scripted, i.e., unpremeditated and spontaneously produced rather than rehearsed or — even worse — read out from paper. To aggravate matters, the type of speaker we were targeting is not easily accessible. These are typically well-known public figures, celebrities who will not be persuaded to participate in a scientific study. As a feasible alternative we decided to record a televised series of talkshows featuring precisely the type of speakers that we were looking for. The particular talkshow, Het Blauwe Licht, was produced by the 'high-brow' VPRO television network in the Netherlands. In each show two guests discussed recent television programs, press photos and newspaper articles.

3. Method

From the winter season of 1998/99 onwards, 16 male and 16 female Dutch-speaking guests who appeared in the television talkshow Het Blauwe Licht, were recorded on tape. The mean ages of the men and women were the same (47, ranges 28–64 and 32–52, respectively). Per speaker some six minutes of spontaneous, non-rehearsed speech were recorded (see §2 above).

For each speaker ten tokens of the target diphthong /si/ were selected from the recordings, along with five tokens of /i/ and five tokens of /a/ (see §2). Tokens preferably occurred before obstruents in stressed syllables of content words; multiple tokens of the same word by the same speaker were avoided.

The audio recordings were digitally sampled (16 kHz, 16 bits) and transferred to computer disk. Using the Praat speech processing software (Boersma and Weenink 1996) the beginnings and end points of the target vowels were located in oscillographic displays. Formant tracks for the lowest four formants (F₁ through F₄) were then automatically computed using the Burg LPC algorithm implemented in
The pronunciation of /si/ by male and female speakers of avant-garde Dutch

4. Results

We will begin by presenting a few overviews of vowel data for individual speakers. Figure 1 represents the acoustic vowel space for one male and one female speaker, whose vowel tokens are dispersed in a perfectly regular fashion.

Observe that vowel height ranges between 2 and 8 Bark for the male speaker and between 3 and 9 for the female. Similarly, the front-back values range between 10 and 14 Bark for the male as opposed to 12 and 16 Bark for the female. This is a direct consequence of the cross-sex difference in the size of the oral and pharyngeal cavities. Still, by virtue of the Bark transformation, equal distances across the male and the female vowel spaces are perceptually the same. The data typically show that the reference point vowel tokens are tightly clustered in the left-hand top corner for /i/ and the open-central area for /a/. There is more variability for the target diphthong /xi/ for both speakers, possibly indicating within-speaker instability for this diphthong. Also, visual inspection reveals that the cloud of /xi/ onsets for the male speaker finds itself roughly halfway between the /i/ and /a/ clusters. For the female speaker, however, the cloud of /xi/ onsets seems to have dropped to a relatively lower position between the /i/ and /a/ reference clusters, leaving a wide gap between the /i/ and /xi/ onset clusters.

More problematic cases arise in Figure 2, which illustrates two problems. Female speaker A.M.’s /i/ tokens (left-hand panel) are far from tightly clustered; in fact, one token seems completely off target. This is due to extreme rounding and centralization of this /i/ token (which sounds more like /y/). Centralisation and/or vowel reduction is also manifest in the /a/ reference tokens of female speaker A.L. (right-hand panel). Although we would expect the female /a/ tokens to have F1 values around 9 Bark — as we did indeed find for the female speakers D.L. (Figure 1) and A.M. (Figure 2 left) — A.L.’s /a/ tokens never extend below 8 Bark.

In view of the susceptibility of the reference point vowels to reduction (centralization) we decided to select the single most extreme (i.e. front-most) token within the speaker’s /i/ cluster as the high-front endpoint of the dimension, and the most extreme (i.e. most open) /a/ token as the other endpoint. Consequently, the speaker’s /i/ token with the highest F2 value and the /a/ token with the highest F1 value were adopted as the extremes of the speaker-individual vowel height dimension.
This procedure allows us to express vowel height speaker-individually as a relative measure. The spectral distance between the extreme /i/ token and the extreme /a/ token is set at 100%, such that /i/ has 100% vowel height and /a/ 0%. When some /si/ onset finds itself exactly midway between the extreme /i/ and /a/ tokens, its relative height will come out as 50%. This measure is implemented by computing (a) the euclidian distance between the reference endpoints, (b) the euclidian distance between the /si/ onset and the /a/ reference value, and (c) the percentage of b relative to a. The smaller the percentage c, the lower the relative starting point of the diphthong.

By the same reasoning we define a relative spectral change measure so as to express the speaker-individual degree of diphthongization for the /si/ tokens. First
we compute the euclidian distance in the Bark-transformed $F_1$ by $F_2$ plane between onset (formant measurements at the 25% temporal point) and offset (measurements at the 75% point) and then take this distance as a percentage of the total distance between extreme /i/ and /a/ of the speaker. A relative glide measure of 25% would then indicate that the /ai/ glide extends over one quarter of the entire front edge of the speaker’s vowel diagram.

These speaker-normalized measures of (relative) vowel height of the /ei/ onset and of the magnitude of the diphthongization are shown in Figures 3 and 4, respectively. In these figures the values have been plotted separately for the male (light grey) and female (dark grey) speakers. In both figures speakers are ordered from left to right in ascending order of ‘conservatism’.

It is obvious from Figure 3 that the female speakers, on the whole, have lower /ei/ onsets than the males. There is one man and one woman with an extremely open /ei/ onset of 20% vowel height. It seems that the change from [ei] to [ai] has been completed for these two speakers. At the conservative end of the scale, there is one woman with a higher (i.e. more conservative) /ei/ onset than the most conservative of the male speakers. For the $2 \times 14$ remaining speakers the women consistently lead in the change from [ei] to [ai]. The effect of sex is significant by a paired t-test, $t(15) = 5.46$ ($p < .001$, one-tail).

Figure 4 reveals the same state of affairs with respect to the (normalized) magnitude of the spectral change in the diphthongs. Clearly, the women generally have a larger difference between onset and offset of the diphthongs than the men, $t(15) = 2.93$ ($p = .005$, one-tail).

Figures 3 and 4 together indicate that the phonetics of the sound change in progress are best characterized as a combined lowering and magnification of the
low-mid diphthong: the onset changes from low-mid to fully low but the offset remains more or less stationary, such that a larger spectral distance has to be covered between onset to offset, which would perceptually enhance the diphthongal nature of the vowel.

Figure 5 plots the relationship between onset lowering and magnitude of spectral change in the /ei/ diphthongs of the 16 men (in grey) and 16 women (in black). The figure, and subsequent statistical analysis, reveals that there is a moderate but significant correlation between onset lowering and strength of diphthongization for the female speakers, \( r = .481 \) (\( p = .030 \), one-tail) but not for the males, \( r = .168 \) (ins.). This finding strengthens the claim that the sound change in progress is predominantly found with female speakers.

One might still argue that the acoustic differences between the male and female diphthongs /ei/ presented so far do not reflect a difference in phonetic vowel quality but are merely due to non-uniform scaling differences between the dimensions of the vocal tracts of men versus women. In order to bear out that the acoustic measures adopted truly reflect differences in auditory vowel quality we have regressed the acoustically defined /ei/ onset values against a perceptual measure for vowel opening that was reported earlier in Edelman (1999). Edelman reported a perceptual index (based on narrow phonetic transcriptions of diphthongs) for the strength of the Polder Dutch impression that was made by (a random selection of) 13 out of the 32 speakers in the present study. The relationship between acoustic and perceptual strength of the avant-garde quality of these 13 speakers is plotted in Figure 6. The correlation between acoustic measure and perceptual impression is considerable (\( r = .742 \)), which clearly indicates that we have not just been measuring the acoustic consequences of differences in the shapes and sizes between male and female vocal organs.
The pronunciation of /ei/ by male and female speakers of avant-garde Dutch

Let us, finally, consider the duration issue of the [ei] to [ai] change. Figure 7 presents the duration of the target /ei/ as well as that of the reference point vowels /i/, which is a phonetically short vowel when immediately followed by an obstruent (Nooteboom 1972), and /a/, which is a long vowel.

Figure 7 shows that the /ei/ of both male and female speakers has the same duration as the long, tense reference vowel /a/, and that both /a/ and /ei/ are some 50 ms longer than the short reference vowel /i/. We computed a speaker-normalized duration measure for /ei/ by dividing the duration difference between /ei/ and /i/ into the difference between /a/ and /i/. A paired t-test on the normalized /ei/ durations revealed no effect of sex of speaker.

5. Conclusions and discussion

The results that were obtained from the acoustic analysis of the 320 targets diphthongs (10 tokens of /ei/ for each of 16 male and 16 female speakers) allow us to answer the phonetic issues raised in the introduction. The phonetic characterization of /ei/ in the emerging avant-garde variety of standard Dutch is that it has a lowered onset. The offset, or end-point of the diphthong, tends to keep its original vowel height, so that the quality change between the onset and offset of the diphthong has increased accordingly. The duration of the new variant of /ei/ has not changed; as a result the speed of the spectral change (rate of formant change as visible in a spectrogram) must have increased as well.

The analysis bears out that the onset of the new /ei/ variety [ai] has the phonetic quality of a low front vowel, close to or even identical to the Dutch tense monophthong /a/ that was used as a reference vowel in the present study. The phonetic quality was therefore judged correctly by Stroop (1998). The lowered /ei/ should not
be equated with the loan diphthong \( /ai/ \) that occurs in some Dutch words that were borrowed from French, such as \textit{detail} ‘id., \textit{email} ‘enamel’. The onset of the imported \( /ai/ \) would qualify as a back vowel, pronounced close to the lax Dutch vowel \( /a/ \).

The analysis also shows that the timing properties of the lowered \( /ei/ \) have remained the same before and after the change. The new variant \( [ai] \) cannot be equated to the long vowel plus glide combination \( /ai/ \) as in \textit{haai} ‘shark’, which combination should be some 50 to 90 ms longer than \( /ei/ \). Consequently, we predict that avant-garde speakers will continue to observe a phonetic contrast in pairs such as \textit{hei} ~ \textit{haai} [hai ~ hai] ‘heather ~ shark’ and \textit{mei} ~ \textit{maai} [mai ~ mai] ‘May ~ mow’.

Sociolinguistically, the data bear out that the avant-garde variant of \( /ei/ \) is more strongly present in the female speaker group than in the male counterparts. Although extremely progressive and conservative speakers are found among both sexes, the women lead the change quite noticeably, especially in the middle portion of the range. This conclusion supports Stroop’s (1998) observation that the avant-garde variety of standard Dutch was initiated by women in precisely the socio-economic group that we targeted in this study.

Methodologically, our study has the added advantage that the sound change in progress could be studied in more detail, and in a non-impressionistic fashion, through the use of acoustic measurement procedures. Moreover, although cross-speaker and cross-sex comparison of acoustic measures of vowel quality are hazardous in principle, the procedure that we applied in our study, i.e., recording reference vowels and performing partial extrinsic speaker normalization on Bark-transformed formant measurements, affords valid comparison across vowels produced by male and female speakers.

As far as we have been able to ascertain, we are the first researchers to have adopted this specific normalization procedure. It bears a resemblance to Gerstman’s (1968) so-called end-point normalization, but differs from it in two details: (i) our procedure specifically looks for the front-most \( /i/ \) and the open-most \( /a/ \) in the front vowel continuum only, while the Gerstman procedure indiscriminately adopts the lowest and highest \( F_1 \) and \( F_2 \) values in an entire vowel set as the end-points, and (ii) our procedure is applied after Bark-transformation.

Notes

1. As a second part of the drag chain, the tense high-mid vowels \( /ce, ce, oe/ \), which have slight diphthongization in the standard language, are somewhat lowered and more noticeably diphthongized.
2. Indeed, whenever Stroop presents illustrations of the new variety on the website or during lectures, it is the diphthong \( /ei/ \) that is used. In fact, the website even invites the public at large to submit a recording of their own production of \( /ei/ \) — rather than some other vowel or diphthong in the language — in order to test whether the particular speaker has already fallen victim to the Polder Dutch sound change.
3. In fact, Stroop produces some examples of male pop-singers who feature the open diphthongs, specifically /i/u/, even today.

4. Given that the first and last portions of any vowel, monophthongs and diphthongs alike, are strongly influenced by (the articulation place of) the neighboring consonants, it is customary to sample the formant values for the starting point of the diphthong at one-quarter of the time-course of the diphthong, and to measure the formants for the endpoint of the diphthong at 75% of its duration.

5. We make the explicit assumption that the point vowels /i/ and /a/ do not participate in the sound change in progress that affects the Dutch mid vowels. We are probably correct in making this assumption as Stroop's vowel diagram does not indicate any involvement of the point vowels /i,a,u/ (1998:28).

6. Ms. Lies Kulsdom, producer of Het Blauwe Licht, characterized the guests as “people with considerable cultural payload such as film producers, authors and intellectuals, who are able to uncover the deeper layers of meaning in film footage and who are not afraid to voice their opinions”.

7. No full-scale studies are available on these durations. The differences given here are visible in spectrograms provided by Cohen and Nooteboom (1976:62) and Rietveld and van Heuven (2001:153–154). Also, the vowel in haai and maai remains stationary for a relatively long time, and then abruptly glides off toward /i/, whilst the vowel quality in hei and mei changes from the beginning onwards (‘t Hart and Collier 1983; Nooteboom and Cohen 1976, Rietveld and van Heuven 2001).

References


