Universiteit Utrecht

The effect of prominence on speaker-specificity in Dutch vowels

Bachelor's thesis Linguistics

Linda Albers
5526329

Supervisor: dr. W. F. L. Heeren
Second assessor: dr. ir. G. Bloothooft

Contents

Abstract

This study aims to contribute to forensic voice comparisons by discussing speaker-specificity (or speaker-dependency) in Dutch vowels as a function of prominence. Previous studies demonstrated that prominent and non-prominent Dutch vowels differ greatly, which leads to the question whether prominence therefore also influences speaker-specificity. This research was done by measuring pitch, duration, intensity, spectral tilt, spectral slope and formants of prominent and non-prominent vowels in spontaneous conversations in a corpus of spoken Dutch. An interaction effect between prominence and speaker was sought on all acoustic variables to determine whether prominence affected speaker-dependency. In addition to this, linear discriminant analyses were performed over all variables together and separately. This was done to determine in which prominence condition the classification of the speakers was the most successful, since this could point out that either prominent or non-prominent vowels are more suitable for performing forensic voice analyses. By doing separate linear discriminant analyses for all acoustic variables, it was determined which variable was the most successful in correctly classifying speakers and therefore the most useful for forensic voice comparisons. Most acoustic variables appeared to be speaker-specific to some extent in both the prominent and non-prominent condition. However, prominent vowels generally seemed to be more speaker-specific than non-prominent vowels.

Introduction

Knowledge of speaker-specificity is of great importance for forensic phonetics, as speech recordings can serve as evidence in court cases. By researching how speaker-specificity manifests itself acoustically, more accurate analyses of this phenomenon can be done in forensic voice comparisons. Knowledge of speaker-specificity may also contribute to a better understanding of the perception of voices: people are able to discriminate or even recognise voices (e.g., Latinus & Belin, 2011). New insights in speaker-dependency might clarify how this is possible.

Determining speaker-specificity has to be done with caution. Rose (2002, p. 26), for instance, states that speakers differ from each other in their voices (inter-speaker variation), but their voices also overlap considerably. At the same time, speakers show variation within their own voice (intra-speaker variation). Speakers never produce the same utterance in the exact same way twice: acoustic differences between utterances are always present (e.g., Latinus & Belin, 2011). To draw accurate conclusions on speaker-specificity, one must therefore determine the ratio between inter-speaker variation and intra-speaker variation. When inter-speaker variation is significantly greater than intra-speaker variation, the acoustic correlate that was studied can be considered suitable for forensic voice comparisons. When there is a substantial amount of variation within and between speakers, the overlap between speakers is too large. In this case, there is less speaker-specific information available (Rose, 2002, pp. 72-73). However, it has to be noted that measurements of speaker-specificity can never provide full certainty of two speech recordings belonging to the same or different individuals due to the large amount of variation between and within speakers.

A substantial number of studies on speaker-specificity have been done. These investigations often analysed vowels, since they contain much information: vowels generally have more intensity than consonants. Besides that, Dutch vowels are generally voiced (Rietveld & Van Heuven, 2013, p. 230). Previous research mostly studied vowel formants as speaker-dependent features, as formants provide relatively much speaker-dependent information. Loakes (2004) examined speaker-specificity by measuring the second (F2) and third formant (F3) in vowels in spontaneous Australian English speech. Both formants proved to be speaker-specific and are therefore suitable for analysing speaker-specificity in a forensic context. Loakes (2004) also concluded that front vowels are more speaker-specific than back vowels. McDougall and Nolan (2007) studied speaker-dependency in vowel formants as well. Their research shows that both the first formant (F1) and F2 are useful for measuring speaker-specificity. Furthermore, they have found that F2 provided more speaker-dependent information than F1. Since formants have proved to be of great use in measuring speaker-dependency, they will be studied in this research as well. Nevertheless, an analysis of speaker-specificity based only on formants is insufficient. A more reliable analysis would include several characteristics of vowels (Nolan, 2001). It is for this reason that this study will also analyse fundamental pitch (F0), duration, intensity, spectral tilt and spectral slope. These variables were selected because these are correlates of prominence as will be explained below.

The aim of the present research is to shine new light on the topic of speaker-specificity by considering the phenomenon prominence, by which stress and accent are meant, in relation to speaker-specificity. As will be explained in the next paragraph, stress and accent cause acoustic changes in vowels, which might influence speaker-dependency in those vowels. By studying the relation between speaker-dependency and the factors stress and accent, analyses of speaker-dependency might be improved. Forensic voice comparisons can then be performed using only the most useful category of vowels, which makes the analysis more reliable.

It is important to note that stress and accent are two prosodic phenomena that differ considerably. Stress is a structural property of a word that indicates which syllable in the word is emphasised, whereas accent is an emphasis on a specific word or syllable in a sentence which may vary depending on the intention of the speaker (Rietveld & Van Heuven, 2013, pp. 286 & 298). Despite being aware of the difference between these phenomena, the present study elected to merge them into one factor: prominence. This allows for an investigation of a general effect of emphasis on vowel acoustics (in relation to speaker-dependency). Previous studies show that stress and accent affect the acoustics of Dutch vowels (e.g., Van Bergem, 1993; Van Heuven & De Jonge, 2011). Stressed and accented vowels tend to have a higher F0 and a higher intensity than non-stressed and non-accented vowels. Moreover, Nooteboom and Slis (1972) and Koopmans-Van Beinum (1980) demonstrated that stressed vowels have a longer duration than non-stressed vowels. Stress and accent also cause differences in formant values, because prominent vowels are pronounced with more care and are therefore spectrally expanded (Rietveld & Van Heuven, 2013, p. 302). Although intensity is a reliable acoustic correlate for stress and accent, it is not a strong cue for perception of stress and accent (Sluijter & Van Heuven, 1996; Rietveld & Van Heuven, 2013, p. 302). The study by Sluijter and Van Heuven (1996) does show, however, that spectral balance is a reliable correlate for stress and accent. They demonstrated that in stressed and accented vowels the intensity of higher frequencies increases more than the intensity of lower frequencies. Rietveld and Van Heuven (2013, p. 204) also state that spectral balance is a reliable cue for perception of stress and accent. The present study will therefore analyse spectral tilt and spectral slope as well.

As is evident from these studies, prominent and non-prominent vowels in Dutch differ greatly. It is therefore possible that prominence influences speaker-specificity. It is however unknown to what extent this is the case. Researching this question is of importance, because it might provide evidence for which category of vowels is more useful for determining the degree of speaker-specificity in speech fragments. If prominent vowels prove to be the most speaker-specific, non-prominent vowels can be disregarded when evaluating speech evidence, and vice versa. Therefore, this study will attempt to contribute to the scientific field of forensic phonetics by posing the following research question: what is the influence of prominence on the availability of speaker-specific information in Dutch vowels?

Expectations

Previous research has shown that prominent and non-prominent Dutch vowels differ from each other in acoustic properties such as F0, duration, intensity, formants, spectral tilt and spectral slope. It is therefore expected that the results of this study will show this as well. The most likely results of this study are that prominent vowels turn out to be longer, have more intensity and a higher F0 than non-prominent vowels. It is also expected that the prominent vowels show a higher increase of intensity in high frequencies than in low frequencies. Furthermore, it is expected that both prominent and non-prominent vowels will prove to be speaker-specific to some degree. Since previous studies show that acoustic features of vowels are speaker-specific, it is expected that speaker-specificity will be found in the following acoustic variables that will be analysed in this study: F0, duration, intensity, spectral tilt, spectral slope, F1, F2 and F3.

The fact that prominent and non-prominent vowels are not pronounced identically may lead to differences in the degree of speaker-specificity between the two categories. This may yield new insights in speaker-dependency in relation to prominence, which previous studies have not investigated as extensively as the present study. On the one hand, it is possible that prominent vowels turn out to be more speaker-specific than non-prominent vowels. Because prominent vowels are articulated very precisely, they make optimal use of the vowel space, whereas non-prominent vowels do not. The difference in use of the vowel space may cause prominent vowels to contain more speaker-specific information than non-prominent vowels. On the other hand, it is possible that non-prominent vowels prove to be more speaker-specific. These vowels may be a better approximation of vowels in spontaneous speech in general. In spontaneous speech, many vowels are not articulated with optimal care and it is possible that there is more speaker-specific information to be found here. Because it is possible that either prominent vowels or non-prominent vowels come out as the most speaker-specific, an interaction effect between prominence and speaker-specificity was expected.

Method

This investigation made use of spontaneous conversations between native speakers of Dutch that are included in the *Corpus Gesproken Nederlands,* or CGN (Oostdijk, 2004). The collection was narrowed down to twenty-one conversations, since only conversations with an available prosodic annotation were used in this study. Prior to gathering data, it was determined which conversations contained enough utterances of the same vowel in the two different conditions per speaker. The criterion was that the vowel had to be uttered at least eight times in, respectively, a prominent and non-prominent position by each speaker. This ensured that enough data would be available for conducting this research and drawing valid conclusions. The pilot revealed that the vowel /i/ met the criterion. The use of this vowel is potentially advantageous to this study, because it is a front vowel. As was stated in the introduction, Loakes (2004) found that front vowels are more speaker-specific than back vowels. The use of /i/ might therefore yield better results than the use of a back vowel.

Twelve speakers who uttered the /i/ less frequently than the criterion of eight times per category of prominence were excluded from this study. The analysis was therefore done with a total of twenty-eight speakers (fifteen males, thirteen females).

The next step was to annotate the relevant intervals in the conversations using the program *Praat* (Version 5.4.08; Boersma & Weenink, 2015). In each conversation all instances of /i/ were annotated per speaker in both prominent and non-prominent positions (annotated as P and NP respectively). In this process, prosodic and phonetic annotations made available by the CGN were used as a reference for correctly identifying and categorising all /i/s. What must be noted is that the prosodic annotations of the CGN do not differentiate between stress and accent, but only mark prominent vowels in speech.

After annotating all 21 conversations, measurements of the acoustical features of all individual occurrences of /i/ were made. A script in *Praat* looped through all annotated files and took the following measurements. First of all, duration (in seconds) was measured by taking the ending point and starting point of the interval and subtracting these from each other. F0 (in Hertz) was measured four times at an equal distance throughout the interval (20%, 40%, 60% and 80%). Additionally, the maximum and mean intensity (in decibel) of the interval were determined. Spectral tilt (in dB/decade) was measured as a coefficient over a range from 500 to 8000 Hz. Spectral balance (in dB) was the difference between the intensities over the ranges of 500 to 2000 Hz and 2000 to 4000 Hz.

For F1, F2 and F3 (in Hertz), the mean values were determined. At first, these formants were measured repeatedly throughout the vowel over smaller intervals, but the majority of these automatic measurements proved to be unreliable or “undefined”. In case of the latter, *Praat* was not able to determine a value. Consequently, these measurements were discarded and instead, mean values and standard deviations for F1, F2 and F3 were determined for all intervals. These measurements mostly resulted in actual values instead of “undefined” cases. Nevertheless, not all formant values were accurate. Because of this, formant measurements with standard deviations above 700 Hz were excluded from further analyses.

All data was collected for each occurrence of /i/ per prominence category per speaker. As with the formants, the data was checked for any extreme values that had to be excluded from the analysis. These were, for instance, F0-values above 350 Hz for male speakers and F2 values above 3200 Hz. Next, normality tests were done using *IBM SPSS Statistics 22* to make sure that the data for each acoustic variable was distributed normally. Variables were transformed, when necessary, to achieve a normal distribution. This only applied to duration, for which a log(10) transformation was done.

Each of the acoustic variables of the vowel /i/ was subjected to an *ANOVA* to determine if the factor prominence interacted with the factor speaker. Prominence here served as a *fixed factor* and speaker as a *random factor*. In the case of an interaction, it was attempted to determine whether the degree of speaker-specificity was higher in either prominent vowels or non-prominent vowels. This was done with separate linear discriminant analyses (LDAs) for prominent and non-prominent vowels. The percentages that were returned by these tests demonstrated which of the two conditions is more suitable for forensic voice comparisons. These analyses were also done for each acoustic variable separately, as this could indicate which of these variables is more useful in discriminating speakers.

Results

In table 1 below, mean values and standard deviations for all variables are displayed per condition. The results will be discussed for each variable individually.

Table 1.
Mean values and standard deviations for each variable in the prominent and non-prominent condition.

	Prominent		**Non-prominent**	
Variable	**Mean**	**SD**	**Mean**	**SD**
F0 (Hz) - 1 (female)	220.6	41.6	205.0	60.1
F0 (Hz) - 2 (female)	223.0	45.0	202.2	59.8
F0 (Hz) - 3 (female)	225.6	46.7	202.0	61.4
F0 (Hz) - 4 (female)	224.1	48.1	199.7	60.5
F0 (Hz) - 1 (male)	151.1	49.5	135.0	36.6
F0 (Hz) - 2 (male)	150.5	50.8	132.6	36.1
F0 (Hz) - 3 (male)	150.6	53.8	132.0	36.6
F0 (Hz) - 4 (male)	153.5	53.6	132.7	37.7
Duration (ms)	102.3	38.0	88.3	36.4
Mean intensity (dB)	64.2	5.5	62.3	5.2
Maximum intensity (dB)	67.5	5.6	65.2	5.3
Spectral tilt (dB/decade)	-4.6	5.7	-5.4	5.5
Spectral slope (dB)	-26.0	10.0	-25.7	10.3
Mean F1 (Hz)	349.4	53.4	350.1	53.9
Mean F2 (Hz)	2389.2	244.4	2354.6	233.3
Mean F3 (Hz)	3236.4	225.6	3235.0	242.6

F0

Since male and female speakers differ greatly in their F0, these two groups were separated in this analysis. This resulted in a group of thirteen female speakers and a group of fifteen male speakers.

The data for the female speakers was distributed normally for measurement points 2, 3 and 4 ($W = .986$, $p = .110$; $W = .986$, $p = .104$ and $W = .986$, $p = .097$ respectively), but not for point

1 (W = .981, p = .030). Despite that, all data was analysed using *ANOVA*. This resulted in a significant effect of prominence for all four cases: [F(1, 15.8) = 6.4, p = .023] for point 1, [F(1, 11.5) = 5.6, p = .036] for point 2, [F(1, 11.2) = 6.0, p = .032] for point 3 and [F(1, 12.6) = 11.1, p = .006] for point 4. The mean values for all four points were higher in the prominent condition than in the non-prominent condition, as can be seen in table 1 and figure 1. The effect of speaker was significant for point 1, 3 and 4 ([F(12, 11.5) = 3.8, p = .016], [F(12, 12.0) = 2.7; p = .048] and [F(11, 11.2) = 3.6, p = .021] respectively), but not for point 2 [F(12, 11.2) = 2.5, p = .072]. Moreover, the interaction between speaker and prominence was not significant for three out of four F0-analyses. The only analysis that yielded a significant interaction effect was the analysis done at point 2 [F(10, 197) = 2.3, p = .016]. The others were not significant with [F(10, 197) = 1.6, p = .106] for point 1, [F(10, 197) = 1.5, p = .136] for point 3 and [F(10, 151) = 1.1, p = .337] for point 4.

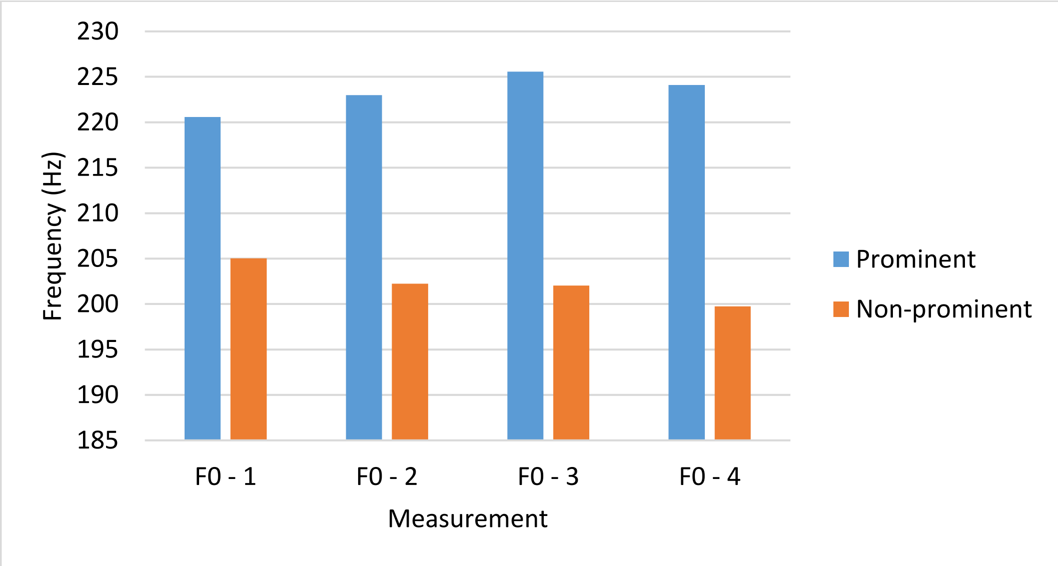

Figure 1.
Mean values in the prominent and non-prominent categories for the four different measurements of F0 of female speakers.

The data for the male speakers was not distributed normally in any of the analyses. The normality tests resulted in (W = .923, p < .001), (W = .913, p < .001), (W = 893, p < .001) and (W = .908, p < .001) for measurement points 1, 2, 3 and 4 respectively. Since this could not be solved with transformation, *ANOVA* was done regardless of this issue. This analysis showed that there was a significant effect of prominence on all four variables of F0: [F(1, 15.9) = 6.5, p = .022], [F(1, 16.2) = 10.1, p = .006], [F(1, 16.3) = 8.7, p = .009] and [F(1, 20.9) = 10.6, p = .004] for point 1, 2, 3 and 4 respectively. The mean values for F0 were higher in the prominent category than in the non-prominent category, as is shown in table 1 and figure 2. The effect of speaker was significant for all four variables as well: [F(14, 14) = 6.1, p = .001] for point 1, [F(14, 14) = 7.1, p < .001] for point 2, [F(14, 14) = 7.5, p < .001] for point 3 and [F(14, 14) = 9.6, p < .001] for point 4. The interaction effect for all measurements of F0 was also significant with [F(14, 289) = 3.2, p < .001], [F(14, 271) = 2.9, p < .001], [F(14, 255) = 2.8, p = .001] and [F(14, 212) = 2.0, p = .023] for points 1 until 4 respectively.

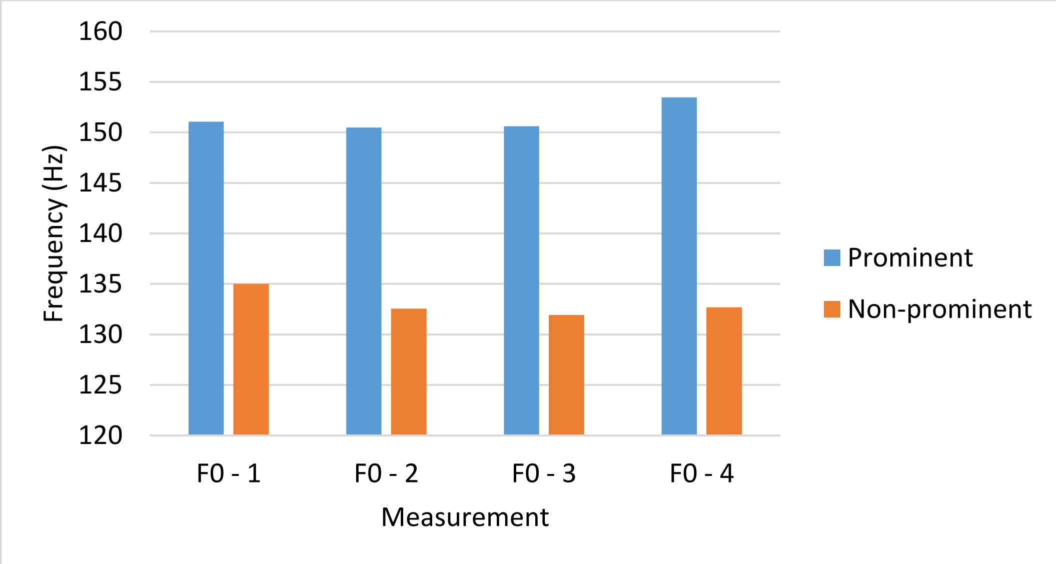

Figure 2.
Mean values in the prominent and non-prominent categories for the four different measurements of F0 of male speakers.

Duration

For this variable, a logarithmic transformation was necessary to achieve a normal distribution (W = .996, p = .052). *ANOVA* showed that the difference between the categories prominent and non-prominent was significant [F(1, 29.0) = 39.9, p < .001]. The mean duration for non-prominent /i/s was 88.3 ms (SD = 36.4 ms), while the mean duration for prominent /i/s was 102.3 ms (SD = 36.4 ms), as can be seen in table 1. The factor speaker also yielded a significant effect [F(27, 27) = 5.6, p < .001]. An interaction effect between prominence and speaker, however, was not found [F(27, 629) = 1.0, p = .471].

Intensity

The data for the variable mean intensity was distributed normally (W = .997, p = .200). The analysis showed that the difference in mean intensity between prominent and non-prominent vowels was significant [F(1, 27.9) = 11.9, p = .002]. The mean value for mean intensity was higher for prominent vowels than for non-prominent vowels, as is displayed in table 1. An effect of the factor speaker was found as well [F(27, 27) = 2.7, p = .005]. Additionally, a significant interaction between prominence and speaker was found [F(27, 638) = 2.3, p < .001]. Figure 3 illustrates this interaction: the values for the variable mean intensity are displayed as an error bar for each speaker in both prominence conditions.

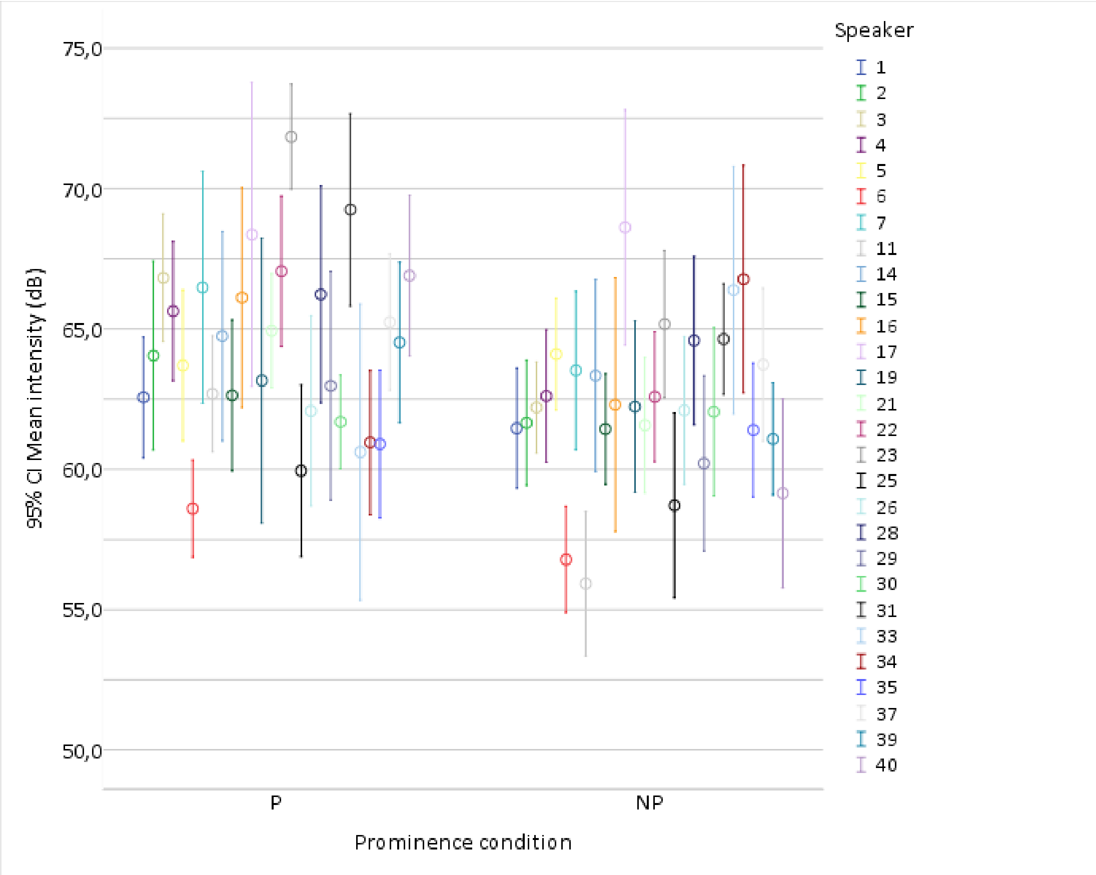

Figure 3.

Error bar plot for the variable mean intensity per prominence condition clustered by speaker.

The data for maximum intensity was distributed normally as well (W = .997, p = .380). The difference between prominent and non-prominent vowels was significant [F(1, 28.0) = 2.0, p < .001]. The mean value for the maximum intensity of prominent vowels was 67.5 dB (SD = 5.6 dB), compared to 65.2 dB (SD = 5.3 dB) for non-prominent vowels, as shown in table 1. A difference between speakers was also found [F(27, 27) = 3.3, p = .001]. As with mean intensity, there was a significant interaction between the factors prominence and speaker [F(27, 638) = 2.0, p = .002].

Spectral tilt

The data for this variable was distributed normally (W = .998, p = .499). Both prominence and speaker yielded significant effects: [F(1, 29.0) = 6.2, p = .019] and [F(27, 27) = 13.0, p < .001] respectively. The mean spectral tilt for prominent /i/s was -4.6 dB/decade (SD = 5.7 dB/decade) and the mean value for non-prominent /i/s was -5.4 dB/decade (SD = 5.5 dB/decade), as table 1 shows. An interaction effect between the two factors was not found [F(27, 638) = 1.0, p = .413].

Spectral slope

The results for this variable were not distributed normally (W = .950, p < .001) and this could not be improved by transforming the variable. It was therefore decided to perform the analysis without ensuring a normal distribution. Prominence did not cause significant differences in the data [$F(1, 28.4) = 1.7$, $p = .200$], while the factor speaker did [$F(27, 27) = 11.5$, $p < .001$]. An interaction effect was not found for spectral slope [$F(27, 638) = 1.5$, $p = .066$].

Formants

Mean F1 and F2 were distributed normally with ($W = .997$, $p = .362$) and ($W = .996$, $p = .129$) respectively). Mean F3 was not distributed normally ($W = .995$, $p = .045$), but the analysis was carried out nonetheless.

For the mean F1, there was no significant effect of prominence [$F(1, 30.0) = 0.7$, $p = .408$]. A significant effect of speaker, however, was found [$F(27, 27) = 5.4$, $p < .001$]. The interaction between prominence and speaker turned out to be significant as well [$F(27, 504) = 1.6$, $p = .033$].

The effect of prominence was significant for the mean F2 [$F(1, 28.2) = 5.5$, $p = .027$] with a mean value of 2389.2 Hz (SD = 244.4 Hz) for prominent /i/s and 2354.6 Hz (SD = 233.3 Hz) for non-prominent /i/s, as can be seen in table 1. Moreover, the effect of speaker and the interaction effect between prominence and speaker were significant for this variable: [$F(27, 27) = 12.2$, $p < .001$] and [$F(27, 634) = 1.7$, $p = .020$] respectively.

For the mean F3, the effect of prominence was not significant [$F(1, 28.8) = 0.4$, $p = .522$]. The factor speaker, on the other hand, did have a significant effect [$F(27, 27) = 7.5$, $p < .001$]. The interaction between prominence and speaker was significant as well [$F(27, 579) = 1.6$, $p = .034$].

Linear discriminant analyses

The LDAs that were done over all acoustic variables resulted in a percentage of correct speaker classification for each of the categories of prominence. However, as several variables in these analyses correlated highly ($r > .850$), they could not be included in the analyses together. The LDAs were therefore performed with the following variables: duration, F0-2, mean intensity, spectral tilt, spectral slope, F1, F2 and F3. F0-1, F0-3, F0-4 and maximum intensity were excluded from the analyses. In the prominent category, 42.3% was classified correctly, whereas in the non-prominent category, 32.9% was classified correctly (compared to a chance level of 3.6%, as 28 speakers were included in the analyses).

As mentioned above, the linear discriminant analyses were also done for all acoustic variables separately. The results of these analyses are displayed in table 2.

Table 2.
Percentages of correct classifications in the linear discriminant analyses for all acoustic variables in the prominent and non-prominent condition.

Variable	LDA	LDA prominent	LDA non-prominent
F0 - 1	9.1%	8.8%	12.2%
F0 - 2	10.6%	8.6%	10.6%
F0 - 3	10.1%	10.8%	10.9%
F0 - 4	11.2%	12.6%	11.2%
Duration	7.2%	4.4%	6.2%
Mean intensity	6.1%	7.5%	7.3%
Maximum intensity	6.9%	5.9%	4.9%
Spectral tilt	8.6%	5.0%	9.1%
Spectral slope	7.9%	8.1%	7.0%
Mean F1	7.3%	7.9%	8.8%
Mean F2	7.5%	10.0%	9.2%
Mean F3	7.4%	10.4%	9.8%

Discussion

The results from this study give clear insights in the acoustics of Dutch vowels. Most importantly, it was determined whether the factors prominence and speaker interacted with each other in the Dutch vowel /i/. This interaction effect was found for mean intensity, maximum intensity, spectral tilt, all three formants and all measurements of the F0 of male speakers. For female speakers, only the second measurement of F0 had a significant interaction effect, whereas the others were not significant. This could imply that prominence does not affect speaker-specificity in the female F0 of the Dutch vowel /i/, which would be interesting to investigate further. This also counts for the variables duration and spectral slope, as neither of those yielded a significant interaction effect. For the variables for which an interaction was found, the results suggest that the degree of speaker-dependency of the vowel is influenced by whether the vowel is in a prominent position or not. In general, one can say that speaker-specificity in the Dutch vowel /i/ depends on prominence to some extent, because nearly all variables showed the interaction.

In addition to *ANOVA*, linear discriminant analyses were done to further explore the relation between prominence and speaker-dependency. These returned a higher percentage of correct speaker classification for prominent /i/s than for non-prominent /i/s (42.3% versus 32.9% respectively). This suggests that in Dutch, prominent /i/s are more speaker-specific and therefore more suitable for forensic voice comparisons than non-prominent /i/s. The LDAs for the separate variables did not demonstrate very clearly which prominence condition is the most useful in determining speaker-specificity, since the percentages of correct classification were higher for prominent vowels in half of the variables and higher for non-prominent vowels in the other half of the variables. The results of these analyses imply that prominent /i/s were more speaker-dependent in F0-4, mean intensity, maximum intensity, spectral slope, mean F2 and mean F3, whereas non-prominent /i/s were more speaker-dependent in F0-1, F0-2, F0-3, duration, spectral tilt and mean F1. This might be because speaker-specificity manifests itself differently in different acoustic variables, which causes some variables to show more speaker-dependency in a prominent position, while others show more speaker-dependency in a non-prominent position. For all variables except F0-2, duration and maximum intensity, the percentages of correct speaker classifications were higher in either one of the prominence conditions than in the LDA done over all measurements regardless of prominence. This suggests that it might be useful to perform forensic voice comparisons using only one of the two categories for these acoustic variables instead of analysing prominent and non-prominent /i/s together.

The general effect of prominence on the different characteristics of the vowel /i/ was also determined. An effect of this condition was found on the variables duration, mean intensity, maximum intensity, spectral tilt, F2 and F0 (the latter for males and females separately). These variables all yielded higher values in the prominent condition than in the non-prominent condition. For spectral slope, F1 and F3, a significant effect of prominence was not found, which suggests that these acoustical features are not affected by stress and accent in this study. This contradicts the expectations of this study and the findings of earlier work (Sluijter & Van Heuven, 1996; Rietveld & Van Heuven, 2013, p. 302). Because several previous studies have shown that spectral slope and vowel formants are affected by prominence, it is possible that the contradicting data for these variables in the present study are due to the confounding factors that will be described below. Apart from the four variables that did not conform to the expectations, the results are (quite) in line with those of previous research that demonstrated a difference between prominent and non-prominent Dutch vowels, as was discussed in the introduction of the present study (Nooteboom & Slis, 1972; Van Bergem, 1993; Van Heuven & De Jonge, 2011). In Dutch, prominent /i/s are indeed longer, higher and have a higher intensity than non-prominent /i/s. F2 is also higher in prominent /i/s than in non-prominent /i/s. These variables thus behaved as expected.

Additionally, the effect of the factor speaker was analysed. This effect was found on all variables except the second measurement of the F0 of female speakers. Since all other F0 measurements yielded significant results, it is plausible that there is an effect of speaker on F0 in general for both male and female speakers (even though F0-2 was not significant for females. It is possible, for example, that the measurements taken for this variable were not accurate). The findings for the three vowel formants conform to previous literature that stated that formants are speaker-specific (Loakes, 2004; McDougall & Nolan, 2007). This is further confirmed by the LDAs of the three formants, as these resulted in 7.3%, 7.5% and 7.4% for F1, F2 and F3 respectively. Since these percentages are relatively high, formants seem to be important to use in forensic voice comparisons. Formants did not, however, yield the highest percentages of correctly classified cases, as the LDAs of the four measurements of F0 resulted in higher percentages. This might mean that F0 is a better predictor of speaker-specificity.

Nevertheless, the high percentages of correct speaker classification for F0 may also be explained by a factor concerning the group of speakers. Both male and female speakers were included in these analyses and since these two groups differ greatly with respect to their F0, this might have influenced the outcome of the LDAs. In this case, the higher percentages of correct classifications do not necessarily mean that the variable F0 is more suitable for forensic voice comparisons than other variables. The measurements of mean and maximum intensity yielded the lowest percentages of correct classification, which implies that these variables are less suitable for forensic voice comparisons. A possible explanation for the variation between the variables is that speaker-dependency manifests itself differently in the different acoustic variables, which causes them to yield differing percentages in the LDAs. Overall, the results imply that all variables in this study provide (at least some) useful information on speaker-specificity.

Not all results were as expected, which might be due to confounds in this research. An example of a factor that has to be taken into account is that all annotations were done by hand and just by one person. It is not unlikely that small mistakes were made in the process, since over eight hundred intervals had to be annotated. Furthermore, it is possible that some of the measurements done by *Praat* are slightly inaccurate (several formant measurements were excluded from the analysis for this reason). It would have been better to check the measurements by hand, but this was not feasible within the timespan of this study. In some cases, the measurements might also have been corrupted due to background noise during the interval. While annotating, intervals that contained severe noise were skipped, but it is possible that a few intervals with (less audible) noise have still been included.

Despite these confounds, this study found rather highly significant effects of (the interaction between) prominence and speaker. Even if the data contained some errors, they can hardly have been the cause of very faulty results. It is more likely that such errors caused only minimal differences in the data and that the results of this study generally represent reality relatively well.

There is, however, still room for improvement. It would be wise to do similar research, perhaps on a larger scale, with vowels other than /i/. Although the analysis of this vowel has provided much useful information, the /i/ might behave differently from other vowels, because the vowels differ acoustically with respect to, for instance, their formants and duration (Rietveld & Van Heuven, 2013, pp. 230-232).To achieve a more general insight in the relation between prominence and speaker-dependency in vowels, a diversity of vowels needs to be studied. One could, for example, choose to study a back vowel such as /o/ or /a/ instead of the front vowel /i/. It also makes sense to treat stress and accent as two separate phenomena instead of as prominence in general. Since stress and accent differ acoustically, one might find different results for these factors with respect to speaker-specificity.

Overall, the present study has provided clear conclusions on the relation between prominence and speaker-dependency. It is evident that both prominent and non-prominent vowels are to some degree speaker-specific. This manifests itself in most, but not all, acoustic features of the vowel /i/. Furthermore, prominent vowels seem to be more suitable for forensic voice comparisons than non-prominent vowels.

References

Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 24 March 2015 from http://www.praat.org.

Oostdijk, N. (2004). Corpus Gesproken Nederlands. Nederlandse Taalunie.

Koopmans-Van Beinum, F. J. (1980). *Vowel contrast reduction: an acoustic and perceptual study of Dutch vowels in various speech conditions.* Amsterdam: Academische Pers B. V.

Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*, *21*(4), 143-145.

Loakes, D. (2004). Front vowels as speaker-specific: some evidence from Australian English. In *Proceedings of the Australian International Conference on Speech Science* (pp. 289-294).

McDougall, K. & Nolan, F. (2007). Discrimination of speakers using the formant dynamics of /uː/ in British English. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1825-1828).

Nooteboom, S. G. & Slis, I. H. (1972). The phonetic feature of vowel length in Dutch. *Language and Speech*, *15*(4), 301-316.

Rietveld, A. C. M. & Van Heuven, V. J. (2013). *Algemene fonetiek*. Bussum: Coutinho.

Rose, P. (2002). *Forensic speaker identification*. London: Taylor & Francis.

Sluijter, A. M. & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical society of America*, *100* (4), 2471-2485.

Van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, *12*(1), 1-23.

Van Heuven, V. J. & De Jonge, M. (2011). Spectral and temporal reduction as stress cues in Dutch. *Phonetica*, *68*(3), 120-132.