The role of prosodic and verbal aspects of speech in the perceived divergence of Dutch and English language varieties

Charlotte Gooskens and Renée van Bezooijen

1. Introduction

Varieties of a language may differ at the same linguistic levels as languages. In the present study we make a distinction between the prosodic level and the verbal level. Prosody we define as all features which are not linked to specific segments, i.e., intonation, tempo variation, and loudness variation; verbal aspects of speech include syntax, lexicon, morphology, and segmental phonetics/phonology. We investigated the role of these two levels in the perception of language varieties, and in particular their relative contribution to the perceived divergence of six Dutch and six English language varieties from the respective standard languages.

Different methods have been developed to assess (directly or indirectly) linguistic distances between language varieties. In traditional dialectology isoglosses play an important role. Generally speaking, the larger the number of isoglosses separating varieties, the greater the linguistic distance is assumed to be. Not all isoglosses are assigned equal weights. However, there is as yet no consensus as to which criteria (e.g., linguistic level, markedness, popular ideas) should be applied to determine their relative importance.

The computation of linguistic distances has also been approached from a purely quantitative point of view. In this method lists of items (e.g., designations of a particular object or pronoun forms) are systematically compared between varieties, each difference contributing to a measure of dissimilarity. In principle dissimilarities can be computed at all linguistic levels, i.e., phonetic, phonological, morphological, lexical, syntactic, or prosodic. The dialectometric approach has originally been developed by Séguy (1973) and has been propagated by Goebl (1984). Dialectometry in its purest form has been applied mainly to Romance and English language varieties. However, related quantitative methods have been applied to Dutch language varieties (e.g., Hoppenbrouwers and Hoppenbrouwers 1988; Nerbonne et al. 1996;
Van Bezooijen (1996). The phonetic/phonological and lexical levels have been examined most often.

A completely different method entails the perceptual scaling of linguistic distance. This method differs in essential ways from the isogloss and dialectometric methods. Firstly, the linguistic distances are determined not on the basis of written records of speech but directly on the basis of auditorily presented, mostly continuous speech fragments. Secondly, the people establishing the distances are not scientists but lay people. Thirdly, distances are assessed not between any two varieties but between non-standard and standard varieties. Examples of this approach are Van Hout and Münstermann (1981) and Van Bezooijen (1994, 1995).

The present study fits in with the last approach, i.e., divergence between non-standard and standard varieties was assessed by means of perceptual judgments given by lay listeners. However, in contrast to previous research the judgments were not merely based on integral speech fragments containing information on all linguistic levels, but also on speech fragments from which either verbal or prosodic aspects of speech have been removed. Moreover, whereas previous perceptual studies were restricted to comparing language varieties within the same language, in the present study both Dutch and English varieties were included, thus enhancing the generalizability of the results.

Our study can be seen as a test of the validity of the common experience that both prosodic and verbal aspects of speech play a part when a short stretch of speech from a speaker not known to the listener is perceived as diverging from the standard language. In order to separate the two linguistic levels, modern signal analysis techniques were used. Both the manipulated and original versions were judged in listening experiments for their degree of divergence from the standard language, so that the perceived divergence based on prosodic information, verbal information, and prosodic and verbal information combined could be compared. In this way, not only could it be assessed whether the two aspects of speech in fact both play a role in the perceived divergence of language varieties, but also to what extent their contribution differs in size.

In what follows we will first describe the speech material (Section 2) and the method used in the experiments (Section 3). Next, the results of the two experiments will be presented and discussed, both separately and combined (Sections 4, 5, and 6). Finally, some general conclusions will be drawn (Section 7).
2. Speech material

Two perception experiment were carried out, one in the Dutch language area with six Dutch language varieties and one in the British English language area with six British English (henceforth English) language varieties. The selection of speech material used in the experiments followed from a number of choices.

First, we had to decide on the specific language varieties to be included. These varieties had to meet two criteria. First, they had to be geographically and linguistically distant from each other. Second, they had to differ in different ways from the standard language: only (mainly) at the verbal level, only (mainly) at the prosodic level, at both levels, or at neither level. On the basis of the dialectological and sociolinguistic literature, the following Dutch varieties were selected:

- Standard Dutch of the Netherlands (sn)
- Standard Dutch of Belgium (ss)
- Bedum (a town in the northeast of the Netherlands, 11,000 inhabitants)
- The Hague (a large city in the west of the Netherlands, 445,000 inhabitants)
- Maastricht (a city in the southeast of the Netherlands, 118,000 inhabitants)
- Uitbergen (a village in the west of the Dutch-speaking part of Belgium, 1,200 inhabitants)

The English varieties were the following:

- Standard English of Britain (sp)
- Liphook (a small town in the southeast of England, 7,000 inhabitants)
- Glasgow (a large city in the southwest of Scotland, 620,000 inhabitants)
- Newcastle (a city in the northeast of England, 260,000 inhabitants)
- Rhondda (a community in the south of Wales; the Rhondda valley has 78,000 inhabitants)
- Belfast (a city in Northern Ireland, 280,000 inhabitants)

The places are indicated on a map of the Netherlands and the Dutch-speaking part of Belgium and a map of Great Britain in Figure 1.
For each language variety three speakers were included. The speakers of the non-standard varieties had all been born and raised in the respective places. They were judged by linguistically trained locals to be good representatives of the language variety they spoke. The speakers of the standard varieties were judged by two phoneticians from the respective countries. All speakers were males between 27 and 45 years of age. Also, since research has revealed some relationships between voice quality and dialect (Trudgill 1974; Esling 1978; Van Bezooijen 1985), speakers were selected who were judged by trained ear phoneticians to have rather “neutral” voice qualities.

To obtain authentic speech fragments the speakers were interviewed by local interviewers speaking the same variety as the interviewees. The interview questions were identical for all speakers. They pertained, for example, to food preferences and favorite television programs. From the interviews for each speaker a speech fragment of between 15 and 20 seconds was selected which was judged by two judges to be geographically and socially neutral in content. So, fragments containing semantic information which would allow the listeners to identify the language variety spoken by the speaker were discarded. Furthermore, only fragments were included which were free of discontinuities. In the selection no attention was paid to the prosodic or phonetic features.
verbal properties of the speech fragments. The fragments were thus not selected for their having dialect-specific features. In this respect they were random samples.

Both in the English and Dutch listening experiments the 18 fragments (6 language varieties by 3 speakers) were presented in 3 different versions:

1. "PROSODY ONLY". The speech is electronically delexicalized. The speech is completely unintelligible and sounds like a low mumble as when someone is talking behind a thick wall. The only thing that can be heard is the prosody, i.e., pitch movements, tempo variation, and loudness contour.

2. "VERBAL INFORMATION ONLY". The speech is electronically monotonized by replacing all pitch values by a constant pitch value. All verbal information is retained, which means that the speech is completely intelligible. The speech sounds like a robot speaking on a flat tone, like in an old movie.

3. "ORIGINAL". This version has the original prosodic and verbal information.

The fragments which were to be manipulated were stored on computer disk with a sampling frequency 10 kHz, 12 bits. Delexicalization was carried out by lowpass filtering the speech signal by computer at 350 Hz. Our filter caused all values above approximately 450 Hz to be attenuated with 50 dB. Monotonization was carried out with the program PSOLA (Moulines and Verhelst 1995). The fundamental frequency was measured by subharmonic summation (Hermes 1988). The pitch contour was changed into a flat line at the mean pitch over all speakers of each language (109 Hz for Dutch and 110 Hz for English).

3. Method

In the Dutch listening experiment, 12 male and 12 female speakers of Standard Dutch participated. They came from different regions of the Dutch-speaking Netherlands and Belgium, but not from Bedum, Maastricht, The Hague, and Uitbergen, i.e., the places where the non-standard Dutch speakers came from. The English listener group consisted of 13 men and 11 women. They came from all over England, but not from Liphook, Glasgow, Newcastle, Rhondda, and Belfast, i.e., the places where the non-rp speakers came from. Most listeners were students from universities and other institutions of higher education. Students of linguistics or Dutch/English language
were excluded, since their knowledge of language varieties might differ from that of other Dutchmen or Englishmen.

The procedure was the same for both groups of listeners. The 54 fragments (6 language varieties by 3 speakers by 3 versions) were presented in separate blocks of “original”, “prosody only”, and “verbal only” fragments. The blocks were presented in two sequences, A and B. Within each block there were two orders of fragments, a random order and a mirrored image of this order. Within each group, half of the listeners heard sequence A and the random orders, the others heard sequence B and the random orders mirrored. Each block was preceded by six practice fragments, one from each of the six language varieties, so that the listeners could get used to the task. These practice fragments consisted of fragments by speakers who had been excluded during the selection procedure. The fragments were separated by 4 seconds response time.

The listeners were instructed to judge each fragment on a scale from 1 (dialect) to 10 (Standard Dutch as spoken in the Dutch/English newscast). The experiment lasted 20 minutes. All stimuli were presented to the listeners, individually or in small groups, through headphones. The listeners were paid for their participation.

4. Dutch results

Below, the results of post-hoc tests are reported which have been carried out to see (1) whether the same varieties were rated differently in the three versions and (2) whether different varieties were rated differently within the same version. In all cases use was made of the Tukey’s Honestly Significant Difference (HSD) procedure, with a significance level of 5 percent.
4.1. The perceived distance between the Dutch language varieties and Standard Dutch

Figure 2. Mean judgments (averaged over speakers and listeners) for six Dutch language varieties broken down for version. Enclosed within square brackets are language varieties that were not judged to be significantly different (p < .05).

In Figure 2 the mean scores per language variety (averaged over three speakers and 24 listeners) are presented, broken down for the three versions: "original", "verbal information only", and "prosody only". The language varieties which were not judged to be significantly different have been grouped within square brackets.

It is clear from Figure 2 that the judgments of the "original" version and the "verbal information only" version are highly similar. And indeed, none of the language varieties differed significantly in the two versions. So, apparently, not hearing any intonation does not affect the perceived distance.

In the "original" and "verbal information only" versions, four groups of varieties can be distinguished. The dialects of Uitbergen, Bedum, and Maastricht, spoken in the geographical periphery of the Dutch area, form one
group. These three varieties are judged to be most different from "perfect" SN (scale position 10). The judged distance is approximately eight scale positions. The dialect of The Hague, spoken in the urban agglomeration of Western Holland (the Randstad), and the two standard languages constitute three separate groups. SN and the dialect of The Hague are placed in the middle of the scale (four and five and a half scale positions from "perfect" SN, respectively). The SN fragments are judged to be closest to "perfect" SN, with a distance of approximately one scale position.

This outcome corresponds very well with linguistic reality in the sense that in fact the varieties spoken in the periphery differ from SN to a much greater extent than, in this specific case, the dialect of the Hague. The Hague dialect is really an accent, since it has (virtually) no lexical, morphological, or syntactic characteristics of its own. The same holds for SN: it (virtually) only differs from SN at the phonetic level.

The SN fragments had been included to function as anchor points. Ideally they would have received ratings close to 10. In fact the mean ratings were a bit below 9. We can think of several reasons why the SN fragments have not been judged to be perfect SN. Firstly, the listeners may have had an even more typical representative of SN in mind as a point of reference when judging the fragments. This explanation is quite plausible considering the fact that the SN speakers when originally judged for their degree of SN by trained listeners, were also not judged to be perfect speakers of this variety. On a scale from 1 (not SN) to 10 (SN) they were rated 8.4, 9.1, and 8.5. Furthermore, the listeners might not have used the whole scale because they wanted to reserve some space in case there was still a fragment to come that sounded even more SN-like. Finally, the listeners may have avoided scale position 10 because they were not absolutely sure about their judgments. In fact, the most extreme values of judgment scales are known to be avoided by judges.

It can be deduced from Figure 2 that the mean judgments for the "prosody only" version are limited to a small part in the middle of the scale. It was apparently very difficult for the listeners to form an opinion about the degree of standardness of the language varieties presented on the basis of the limited amount of information they received. Due to the small range of the ratings, the groups of varieties are not as distinct in the "prosody only" version as in the other two versions; they show some overlap.

Only SN is clearly distinguished from the other language varieties. This seems a remarkable outcome. It suggests that in the ears of the subjects there is a dichotomy in prosodic characteristics between Standard Dutch on the one hand and non-standard Dutch, of whatever kind, on the other. Of course,
it should be kept in mind that the listeners were speakers of Standard Dutch. We do not know to what extent their ratings for the non-standard fragments were a positive choice, in that they were indeed able to identify the prosodic characteristics presented, or a negative choice, in that they only perceived those characteristics not to be part of the standard variety.

The distance between the judgments of the SN fragments and “perfect” SN is about four scale positions. Unfortunately, it is not possible to tell how much of the distance is due to the fact that the listeners actually considered the prosody of the fragments to be “imperfect” SN and how much to a lack of confidence. In other words, it is unclear to what extent the central scores are a positive or a negative choice.

The prosody of the dialect of Maastricht is not judged to be more divergent from the standard language than the prosody of the other language varieties. This is contrary to our expectations, since the prosodic markedness of the Limburg dialects is often referred to in the literature. It is unlikely that lack of perceived divergence has to do with the Maastricht speakers not having been properly selected. The three speakers included in our study were judged to be the broadest speakers among eight speakers who were all considered to be broad speakers of the Maastricht dialect, by the Maastricht interviewer, by the speakers themselves, and by three linguistically trained natives. So, there must be another explanation. We can think of three.

Firstly, it should be kept in mind that only syllable-like entities are recognizable in low-pass filtered speech. Individual words and the position in the sentence cannot be recognized. It is possible - though not very likely - that the same kind of durational and intonational differences are present in SN as in the dialect of Maastricht, but that they are distributed differently or have a different function. If this is the case, little difference will be heard between the Maastricht dialect and SN in the “prosody only” version. This is a consequence of the method chosen.

Secondly, an explanation might be found in the nature of typical intonation contours in language varieties. The prosody of language varieties can differ in two respects. On the one hand, varieties may differ in that some have extremely divergent contours whereas others have contours which are similar to SN. This would be a qualitative difference. On the other hand, language varieties may differ in that some have few whereas others have many divergent contours. This would be a quantitative difference. If the Maastricht dialect is of the qualitatively different type, with markedly divergent intonation contours occurring infrequently, this might explain why the prosody was not perceived as particularly divergent by the listeners. The
fragments were only 15 to 20 seconds long, and since they were selected randomly without special attention to prosody there is a chance that these divergent contours were not included in the material. The outcome of the study would then not be valid in a general sense, since it suggests Maastricht prosody is not divergent whereas in fact it is. However, the study does resemble the real-life situation in which a short stretch of speech is overheard by chance in a noisy environment. With respect to this situation, the outcome would be a valid predictor of perceived divergence.

Thirdly, the Maastricht dialect might be less marked prosodically than (some) other Limburg dialects. If this is the case, another choice might have yielded results that are more in accordance with our expectations. The literature on the Limburg dialects, which is much more extensive with respect to lexicon and phonology than with respect to prosody, gave no clear indications in this regard. So much is sure, however, that the Maastricht dialect does have the lexical tones which are so typical of the Limburg dialects. Definitely, more descriptive and perceptual research is needed on Limburg prosody before it can be decided whether our results are typical of Limburg dialects or not.

4.2. The role of prosody compared with the role of verbal information in the perceived divergence of the Dutch language varieties

Above we have discussed the distances between the different language varieties for each of the versions ("original", "verbal information only", and "prosody only") separately. In order to assess the relative importance of prosodic and verbal information, we will now directly compare the distances to SN for the three versions. Results will be presented in two ways. First we will present the distances of the different versions relative to scale position 10, i.e., "perfect" SN. We will try to make clear that these absolute scores give a distorted picture of the role of prosody compared with that of verbal information. We will then propose an alternative way of presenting the distances, using perceived SN as a point of reference rather than "perfect" SN. In this way relative rather than absolute scores are obtained.

In Figure 3 the differences between "perfect" SN and the mean scores of each variety are presented for each of the three versions. As can be seen, the differences between the "original" version and the "verbal information only" version are small for all language varieties. This is what we expected on the basis of the results presented in Figure 2. For the two standard languages,
The “prosody only” version shows a greater distance to “perfect” SN than the “original” versions. This is an unexpected result, since prosodic features are included in the “original” version and form a subset, in addition to other possible divergent features. We attribute this seemingly strange result to the concentration of the ratings in the “prosody only” version in the middle of the scale.

Figure 3. Distances (averaged over speakers and listeners) between the judgments for six Dutch language varieties and “perfect” SN (scale position 10), broken down for version.

The distorting effect of “prosody only” on the score distributions is demonstrated in Figure 4. This figure makes clear that when relating SN to “perfect” SN (scale position 10), the distance (d) is larger in the “prosody only” version than in the “original” version. However, taking perceived SN as a point of departure, the distance (d) of SN to SN is smaller in the “prosody only” version than in the “original” version. This reversal is apparently due to the changed position of perceived SN relative to “perfect” SN. Below we will present distances in terms of relative rather than absolute scores.
Figure 4. Distances (averaged over speakers and listeners) between the judgments for SB and perceived SN (a) and perceived SN and “perfect” SN (b, scale position 10) for “prosody only” version and the “original” version.

For all three versions the relative distances to SN of each language variety (other than SN) were calculated as follows. For each listener separately, we subtracted the score for each non-standard speaker from the mean score given by that listener to the three SN speakers. The relative scores thus obtained, averaged over listeners and speakers, are presented in Figure 5.

For none of the language varieties the difference between the “original” version and the “verbal information only” version was found to be significant. It must therefore be concluded that it does not make a difference for the listeners whether prosodic information is present or not; the “verbal information only” version and the “original” version sound equally divergent from the SN fragments.
Figure 5. Relative distances to perceived SN of non-SN varieties. The percentages have been obtained by dividing the distances to perceived SN in the "prosody only" version by the distances to perceived SN in the "original" version, times 100.

The perceived divergence of the fragments is much smaller in the "prosody only" version than in the "original" and "verbal information only" versions. The difference is significant for all language varieties. Again it is shown that verbal information contributes much more to the perceived divergence of the language varieties than prosodic information.

On the other hand it must be noted that even in the "prosody only" version there is still some distance between the non-standard and the SN fragments. If the listeners had not been able to hear any difference among the language varieties in this version, the distances would have been zero. This is clearly not the case, and it can thus be concluded that prosody still plays a - be it limited - role for the perceived divergence of the language varieties in question.

In order to be able to express the relative importance of prosodic and verbal aspects of speech by means of a single value, we divided the relative difference scores for the "prosody only" version by those for the "original version". The proportions were subsequently multiplied by 100 to obtain
percentages (see Figure 5). Roughly speaking, prosody appears to be responsible for between 1/5 and 2/5 of the perceived distance to SN. For the dialect of Maastricht the percentage is smallest (18%). This means that the Maastricht prosody is of relatively little importance for the perceived divergence of this dialect. Prosody plays the most important role for the divergence of SN (39%). The latter result confirms the opinions expressed by respondents to written questionnaires (Knops 1984; Hagen 1980). They assumed the difference between Standard Dutch in the Netherlands and in Belgium to be largest at the level of pronunciation, but intonation was deemed important as well. Willeyns (1987) even thought intonation to be the most important differentiator, but this is not confirmed by our results. The percentages for the other language varieties lie between those of Maastricht and SN.

5. English results

5.1. The perceived distance between the six English language varieties and RP

In Figure 6 the judgments of the English listeners for the English varieties are presented, separately for the three versions. Here also, language varieties which were not judged to be significantly different from each other have been placed together within square brackets.

The varieties have been judged in the same way in the “verbal information only” and the “original” version except for the dialects of Rhondda and Belfast, where the judgments are significantly lower in the “original” version than in the “verbal information only” version. This means that for the perceived divergence of these two dialects intonation is so important that its removal has an effect on the ratings.

Not surprisingly, the speakers who had been included as representatives of RP were judged to be significantly more RP than any of the other language varieties. However, the RP fragments are not considered to be “perfect” representatives of the standard language. The same reasons may apply here as forwarded to explain the not completely “perfect” ratings for the SN fragments in the Dutch experiment (see Section 4.2).
In addition to RP, two other groups can be distinguished in the "original" and "verbal information only" conditions. The Liphook dialect forms one group by itself and the four dialects spoken in the periphery of the English language area form another group. The dialect of Liphook, representing a southern language variety, is judged to be much less divergent than the other four dialects in the investigation, which are given extremely low scores (between 1.7 and 2.7).

In the "prosody only" version, RP and the Liphook dialect form one group. The four varieties from the periphery form two overlapping groups: Rhondda is the most divergent variety, Glasgow takes an intermediate position, and Belfast and Newcastle are the least divergent varieties.
5.2. The role of prosody compared with the role of verbal information in the perceived divergence of the six language varieties

![Bar chart showing distances to RP for different versions.](image)

<table>
<thead>
<tr>
<th>Location</th>
<th>Original</th>
<th>Verbal Info Only</th>
<th>Prosody Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liphook</td>
<td>15%</td>
<td>40%</td>
<td>27%</td>
</tr>
<tr>
<td>Rhondda</td>
<td>40%</td>
<td>30%</td>
<td>27%</td>
</tr>
<tr>
<td>Belfast</td>
<td>27%</td>
<td>39%</td>
<td>39%</td>
</tr>
<tr>
<td>Glasgow</td>
<td>39%</td>
<td>39%</td>
<td>39%</td>
</tr>
<tr>
<td>Newcastle</td>
<td>30%</td>
<td>30%</td>
<td>30%</td>
</tr>
</tbody>
</table>

**Figure 7.** Relative distances to RP of non-RP varieties. The percentages have been obtained by dividing the distances to perceived RP in the “prosody only” version by the distances to perceived RP in the “original” version, times 100.

In order to compare the perceived divergence in the three versions, we have followed the same procedure as in the Dutch experiment. In Figure 7 the relative distances to RP, i.e., the distances to perceived RP rather than to perfect RP (10 on the scale), are presented, as well as the percentages obtained by dividing the relative distances in the “prosody only” version by those in the “original version”.

The “original” and “verbal information only” versions were found to be significantly different from the “prosody only” version for all varieties. This means that the perceived distance to RP is always larger when verbal information is present than when it is not. Furthermore, with respect to the Rhondda and Belfast dialects there is a significant difference between the “original” and the “verbal information only” version. For these dialects it makes a
difference whether intonation is present in the speech fragments or not: the
distance to RP fragments is perceived as smaller when intonation is absent
than when it is present, a finding that had also emerged from the absolute
distances.

6. Comparison of the Dutch and the English results

The results of the two listening experiments described in Sections 4 and 5
show that both verbal and prosodic aspects of speech play a role in the
perceived divergence of Dutch and English language varieties, thus confirm-
ing every-day, pre-scientific intuitions. This appeared most clearly from the
fact that both in the "verbal only" and the "prosody only" versions (groups of)
varieties were systematically distinguished from each other.

Moreover, the results for the two languages agree in that in both cases
verbal aspects of speech were found to contribute more to the perceived
divergence of language varieties than prosodic aspects of speech. This could
be concluded from the fact that varieties were differentiated much more
clearly in the "verbal information only" version than in the "prosody only"
version. In the former version four (groups of) varieties (Dutch experiment) or
three (groups of) varieties (English experiment) were distinguished without
overlap. In the latter version only two groups of varieties were distinguished
without overlap: Standard Dutch from non-standard Dutch in the Dutch
experiment and RP and Liphook from the peripheral English varieties in the
English experiment. It could also be inferred from the fact that the removal of
intonation hardly affected perceived divergence whereas removal of verbal
information invariably led to a drastic decrease in perceived divergence.

The question of whether the relative importance of verbal and prosodic
aspects is identical for the two languages, is less easy to answer. First, the
part of the scale used in the "prosody only" version is somewhat larger in the
English experiment (2.7 scale positions) than in the Dutch experiment (2.3
scale positions). This difference is caused both by the English varieties from
the periphery being rated a little more divergent than the Dutch varieties
from the periphery and the RP variety being rated a little more standard than
SN. However, it is not clear whether this difference in range is meaningful, as
its statistical significance cannot be tested.

Second, averaged over all varieties, the percentages of the perceived
divergence attributable to prosody (i.e., the percentages included in Figures 5
and 7) are identical for the Dutch and English experiment, namely 30%. So,
speaking in general terms, this finding suggests no difference in the relative importance of prosodic and verbal aspects of speech for the two languages. However, if the percentages are averaged separately over the varieties from the periphery of the respective language areas, i.e., over Bedum, Maastricht, and Uithoorn for the Netherlands, and over Rhondda, Belfast, Glasgow, and Newcastle for Great Britain, a difference is found, the mean percentage being 34% for the English language area and 26% for the Dutch area. So, taking the peripheral varieties as a point of departure, prosody seems to play a more important role in the English than in the Dutch experiment. Of course, if the non-peripheral varieties are looked at, the opposite picture emerges: the percentage for Liphook is is only 15%, whereas the percentages for The Hague and ss are 39% and 35%, respectively. So, in this case one would have to draw the conclusion that prosody plays a less important role in the perceived divergence of English varieties than of Dutch varieties. At this point, the question of the comparability of the non-peripheral varieties selected for the two languages is of crucial importance. To what extent can one compare the Hague variety (an urban dialect or rather accent), ss (a standard language), and Liphook (a non-urban, non-standard variety)?

However, we think there is one result which points unambiguously to prosody playing a larger role in the English than in the Dutch language area, namely the fact that for two English varieties, namely Rhondda and Belfast, the removal of intonation had a significant, decreasing effect on perceived divergence, whereas for the Dutch varieties this was the case for none of the varieties.

7. General conclusions

The study reported here fits in with a history of methods to measure divergence between language varieties. A perceptual approach was adopted, using lay listeners. New is the use of electronic techniques to artificially separate the prosodic level of speech (i.e., intonation, tempo variation, and loudness variation) from the verbal level (i.e., syntax, lexicon, morphology, and segmental phonetics/phonology). The method used has proven to yield interpretable and systematic results. In two experiments conducted independently from each other in different language areas (English and Dutch) highly similar tendencies were assessed in the ratings of perceived divergence. In both cases both verbal and prosodic aspects of speech were found to contribute to the perceived divergence of language varieties, and in both cases the latter
aspect was found to be less important than the former. In addition, relatively speaking, prosody appears to play a somewhat larger role in the perception of English varieties than in the perception of Dutch varieties, at least those spoken in the periphery of the area.

Notes

1. This article is based on a doctoral dissertation by Charlotte Geoskens. For more detailed data we refer to this work. The research was supported by the Linguistic Research Foundation, funded by the Netherlands Organisation for Scientific Research, NWO, project number 300-173-024. We would like to thank Toni Rietveld for his statistical and methodological support and Dick Smakman for collecting the English speech material.

2. The mean rating for the standard speakers from the Netherlands and Belgium was 8.7 and 7.2, respectively, on a scale from 1 (not at all) to 10 (very).

References

Ealing, John H.

Goebl, Hans
1984 Dialetometrische Studien, Beihefte zur Phonetik 191: 1, 2, 3, Tübingen.

Geoskens, Charlotte
1997 On the role of prosodic and verbal information in the perception of Dutch and English language varieties. Dissertation, University of Nijmegen.

Hagen, Anton

Hermes, Dick J.

Hoppenbrouwers, Cor and Geer Hoppenbrouwers

Knops, Uts
Charlotte Gooskens and Renée van Bezooijen

Moulines, Eric and Werner Verhelst

Nerbonne, John, Wilbert Heerings, Erik van den Hout, Peter van der Kooi, Simone Otten, and Willem van de Vis

Séguy, Jean

Trudgill, Peter

Van Bezooijen, Renée


Van Hout, Roeland and Henk Müntemann

Willemsens, Roland