

Word Intelligibility of Language Varieties in the Netherlands and Flanders under Minimal Conditions

Renée van Bezooijen and Rob van den Berg

1. Introduction

How much do speakers of standard Dutch understand of autochthonous language varieties with which they are not familiar? A rather obvious question in a language area where so much variation is found in such a small region as the Netherlands and Flanders. A question nevertheless which has only been approached till now by asking listeners to indicate their subjective and global impression of the degree of intelligibility of dialect fragments on a judgment scale. Examples of this type of research are Boets and De Schutter (1977), Van Hout and Münstermann (1981), and Van Bezooijen (1994). Experimental research quantifying the intelligibility of autochthonous language varieties objectively has not yet been conducted, in contrast to other types of language varieties, such as computer generated Dutch (e.g. Van Bezooijen and Pols 1993), allochthonous Dutch (Van Heuven and De Vries 1983), and pathological Dutch (Maassen 1985).

In this article an experimental study is described set up to assess how much speakers of standard Dutch understand of various autochthonous language varieties spoken in the Netherlands and Flanders under minimal conditions, i.e. with virtually no information from the context. More specifically intelligibility was assessed for varieties spoken in the Dutch province of Groningen, the Dutch province of Limburg, the Belgian province of West-Flanders, and the Dutch province of Friesland. The last variety was included as a representative of Frisian, which is generally considered a separate language from Dutch.

Intelligibility was assessed by means of an auditory translation task which focussed on the intelligibility of nouns referring to concrete objects from everyday life. This type of words is known to exhibit a great deal of variation among languages and language varieties. The nouns presented may differ from standard Dutch, the language of the listeners, at all linguistic levels: phonetically, phonologically, morphologically, lexically, and semantically. Intelligibility was expressed as the percentage of nouns translated correctly into Dutch, and consequently the study

can be seen as a study of word intelligibility. Context information was minimized by presenting the listeners with short, isolated speech fragments and by limiting the number of content words in each fragment to one, i.e. the target noun. So, the listeners had (virtually) only bottom up information to decide on the meaning of the target words.

It is hard to formulate hypotheses as to the outcomes of this study. From the subjective judgment study by Van Bezooijen (1994), in which listeners were asked to *estimate* intelligibility of speech fragments presented, it could be inferred that the varieties spoken in Bedum and Tielts are very (and equally) difficult to understand. 25 students from Nijmegen University gave both varieties a rating of 2.3 on a 10-point judgment scale where 1 indicates completely unintelligible and 10 perfectly intelligible. The subjective intelligibility studies by Boets and De Schutter (1977) and Van Hout and Münstermann (1981) are incomparable to the present study, both with respect to the type of listeners participating and the varieties presented, and can therefore not be used as a basis for predictions.

A second approach to estimating intelligibility could be based on objectively determined linguistic distances. It seems reasonable to assume that a variety will be harder to understand as it differs more from the variety spoken by the listener. There have been several attempts to quantify linguistic distances between varieties spoken in the Dutch language area. However, the studies by Hoppenbrouwers and Hoppenbrouwers (1988) and Nerbonne et al. (1996) are unfit to predict dialect intelligibility as they are restricted to the phonological level. We expect lexical differences to play an important role in dialect intelligibility.

The only study in which all relevant linguistic levels have been explicitly taken into account is the one by Van Bezooijen (1996). In this study distances were calculated between the varieties of Bedum and Tielts (also included in the present study) and standard Dutch. Tielts appeared to deviate more from standard Dutch than Bedum at all levels: phonologically, morphologically, lexically, and semantically. This result suggests that the Tielts variety will be more difficult to understand than the Bedum variety (note that this prediction disagrees with the prediction based on the subjective intelligibility data of Van Bezooijen (1994) in the preceding paragraph). No data are available to predict the relative intelligibility of the Frisian and Limburg varieties.

2. Method

2.1 Varieties

In this study the intelligibility of four language varieties from the periphery of the

Netherlands and Flanders was examined. The four varieties were chosen with a view of maximal geographical and linguistic distance, both from each other and from standard Dutch. Each variety was represented by three speakers. The speakers of the Groningen dialect were born and raised in Bedum (10 km north east of Groningen city), those from Limburg in Valkenburg (10 km east of Maastricht), those from West-Flanders in Tielt (25 km south east of Bruges), and those of the Frisian language in Grou (15 km south of Leeuwarden). The location of these places is shown in Figure 1.

All speakers were female. Information on their age and socio-economic status (SES) is given in Table 1. The SES-index is the average value for education and profession. It ranges between 1 and 6, where 1 indicates primary education and unskilled labor and 6 indicates an academic degree and higher profession. It can be seen from Table 1 that the varieties are well matched, both with respect to speakers' age (on average around 33) and with respect to speakers' SES (on average somewhat below the middle of the scale).

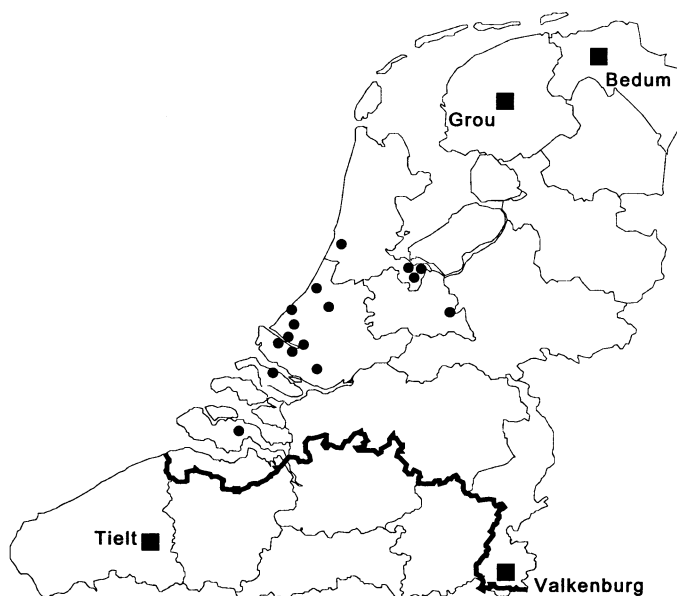


Figure 1. Location of the four varieties investigated (■) and the places of birth of the listeners (●)¹

Table 1. Speakers' age and socio-economic status (see text)

Variety	Age (mean and per speaker)	SES (mean and per speaker)
Groningen (Bedum)	35.0 (22, 39, 44)	2.3 (1.0, 1.0, 5.0)
Frisian (Grou)	33.0 (26, 29, 44)	3.0 (2.0, 3.5, 3.5)
Limburg (Valkenburg)	34.0 (29, 33, 40)	3.2 (2.0, 3.5, 4.0)
West-Flemish (Tielt)	33.3 (21, 35, 44)	3.2 (1.0, 3.5, 5.0)

2.2 *Speech material*

The speech material consisted of oral descriptions of five drawings representing everyday situations with all sorts of concrete objects. The drawings depicted a street (Figure 2), a house, a living room, a kitchen, and a bedroom. The speech material has been collected between 1991 and 1996 by five female interviewers, each speaking the same variety as the informants. Informants were selected on the basis of autochthony and the "dialectal authenticity" of their speech. Most of the informants were personally known to the interviewer, and the recordings took place at the informants' homes. The advantage of this approach is that (1) fairly informal and (more or less) spontaneous speech is obtained, (2) the speech material is comparable for the various varieties, and (3) lexical variation is maximal.



Figure 2. One of the drawings ('the street') described by the informants

From the descriptions of the drawings short speech fragments were selected which contained one noun referring to a concrete object and no other content words. Within each variety, each word was included only once, which means that identical tokens from different speakers were excluded. The total number of fragments meeting these criteria was 156. Segmentation was done by means of the computer.

As we wanted to reduce the amount of information contained in the context, a paper-and-pencil experiment was carried out to assess the predictability of the target nouns on the basis of the information present in the fragments. To that end all 156 fragments selected were presented to 8 subjects in their written form, translated into Dutch except for the target nouns (similar to the set-up of the intelligibility task, see Section 2.4). Subjects were asked to guess the words missing. Three responses were allowed. Fragments were discarded if one or more subjects gave the target noun itself or a semantically closely related noun either as a first, second, or third response. This applied to 19 fragments. So, 137 fragments were kept to be used in the intelligibility experiment: between 32 and 36 per variety and between 9 and 13 per speaker.

2.3 *Listeners*

16 listeners participated in the intelligibility study: 8 women and 8 men, between 18 and 50 years old, with an average of 30 years. All were students and staff members of the universities of Leiden and Nijmegen, all but two from language departments. All were born and raised in the provinces of Zuid-Holland, Noord-Holland, or Utrecht, i.e. in the western and central part of the Netherlands.¹ In this part of the country no dialects in the true sense, i.e. varieties with typical characteristics at all linguistic levels, are found, except in very isolated communities in the older generations. Everyone speaks a variety which is very close to the standard; one could speak at most of typical accents, the few differences from the standard language residing in pronunciation. In addition to the subjects not speaking any dialect, the provinces mentioned were chosen because they are not adjacent to the provinces where the speakers are from. In this way it can be sure that intelligibility is not based on a priori knowledge of the varieties investigated. The places of birth of the 16 listeners are indicated in Figure 1.

2.4 *Task*

Each listener heard four blocks of speech fragments, one for each variety. Within the blocks fragments and speakers were placed in random order. This means that the fragments were isolated from their original context, so that not only information

from the immediate context within the fragments was minimized but also any possible information from the wider context of the description of the drawings as a whole. The blocks were presented in four different orders, each order being presented to four listeners. So each variety was presented to four listeners in first position, to four others in second position, etc. The four blocks were preceded by five practice fragments in standard Dutch, so that the listeners could get used to the task.

The speech fragments were presented to the listeners over headphones, in a quiet environment. The listeners heard each fragment with in front of them the text written out in Dutch. In the written text the target noun, always referring to a concrete object on the drawing, was left open. The listeners' task consisted of translating the noun in question from the variety at hand into Dutch and noting the Dutch word on the empty place in the text. Fragments were separated by pauses of 6 sec in which the translation was to be written down.

3. Results

3.1 *Percentages correct translations*

The percentages target nouns translated correctly, averaged over 16 listeners, are 94 for the Groningen dialect, 58 for the Frisian dialect, 80 for the Limburg dialect, and 58 for the West-Flemish dialect. These data are shown graphically in Figure 3. Incorrect responses are of two kinds: wrong translations and missing translations.

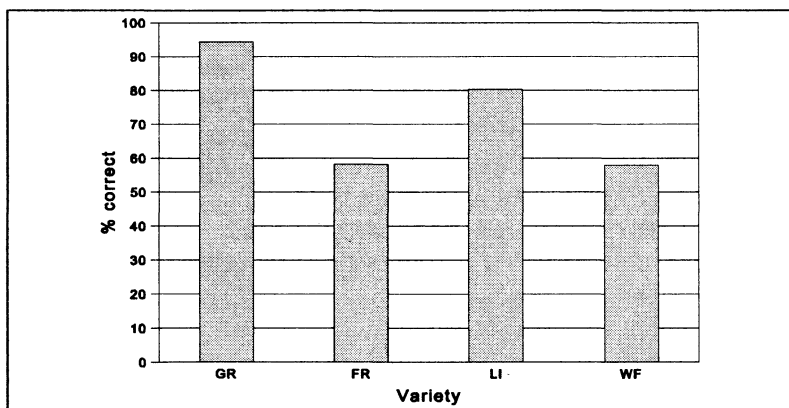


Figure 3. Percentages correctly translated nouns per variety. GR=Groningen, FR=Frisian, LI=Limburg, WF=West-Flemish.

Of course, to be able to interpret these results, one would like to know whether the mistakes made are truly due to problems caused by the varieties at hand and not to problems caused by the task itself. In other words, one would like to know what the performance of the subjects is like under ideal circumstances, without any “noise” caused by the fact that dialect fragments are presented rather than standard Dutch fragments. This type of data was obtained in another intelligibility experiment not reported on here (Van Bezooijen and Van den Berg, submitted). This experiment was set up along similar lines as the present experiment, the main difference being a greater amount of information present in the context, allowing top down information to be combined with bottom up information.

In this information rich experiment, a condition was included with three speakers of standard Dutch. It appeared that the percentage of incorrectly translated or missing target nouns was less than 2% for this variety. Evidently, the task itself was easy to carry out. We assume the same holds for the present experiment. We think that the errors yielded have been caused almost completely by properties of the varieties investigated and that the height of the percentages is a valid representation of intelligibility at the lexical level. This assumption is confirmed by the fact that no mistakes were made by the listeners with the practice stimuli, which were all in standard Dutch.

The picture emerging from Figure 3 is clear: the Groningen dialect of Bedum is easy to understand for speakers of standard Dutch, the Limburg dialect of Valkenburg takes an intermediate position, and the West-Flemish dialect of Tielt and the Frisian dialect of Grou are difficult to understand. An analysis of variance showed the factor “language variety” to have a significant effect on the intelligibility data ($F=75.8$; $df=3,12$; $p=.000$). This means that some varieties are significantly easier to understand than others. A post-hoc analysis (Tukey’s HSD) revealed that all pairs of varieties differed significantly from each other except for those of Frisian and West-Flemish.

3.2 *Explanations in terms of types of deviation from standard Dutch*

Of course one would like to know *why* the four language varieties yielded such widely diverging results, a difference of 36% between the dialect of Bedum and the dialects of Grou and Tielt. One would like to trace the responses of the listeners back to the characteristics of the varieties presented. We therefore set out to make a profile of each of the four varieties, distinguishing six categories of relations between the dialect target nouns and their standard Dutch equivalents.

- Category 1. The dialect word is identical with its standard Dutch equivalent (e.g. standard Dutch *dak* /dak/ 'roof' corresponds with West-Flemish *dak* /dak/)
- Category 2. The dialect word differs from its standard Dutch equivalent in one vowel (e.g. standard Dutch *rook* /rok/ 'smoke' is *rauk* /rouk/ in Limburg)
- Category 3. There is a difference in one consonant (e.g. standard Dutch *vrouw* /vrou/ 'woman' is *frau* /frou/ in Frisian)
- Category 4. There are differences in several phonemes but a language user can (be it sometimes with some effort) determine that the two words are related, i.e., that they are cognates (e.g. standard Dutch *haken* /hakən/ 'hooks' is *oakn* /ɔ:kn/ in West-Flemish)
- Category 5. There is a lexical difference: the form of the dialect word is (or seems from a synchronic point of view) unrelated to its standard Dutch equivalent (e.g. standard Dutch *laars* /lars/ 'boot' is *stevēl* /stevəl/ in Groningen).
- Category 6. There is a semantic difference: the form of the dialect word is identical with that of a standard Dutch word, but the meaning differs (e.g. West-Flemish *patat* /patat/ means 'potato' whereas standard Dutch *patat* /patat/ means 'french fries').

In the categorization, differences which we expect not to give rise to confusions with standard Dutch vowels and consonants have been ignored. This applies for example to non-standard Dutch realizations of /r/ and /χ/, aspiration of stops, more closed or open realisations of vowels, and audible nasalisation of vowels preceding a nasal consonant. This means that identity in category 1 should not be taken literally; there may be (and often are) phonetic differences between the dialect word and the corresponding word in standard Dutch. The degree of phonetic deviation may differ between the four varieties and within the varieties between the various speakers. Also, an extra schwa at the end of nouns, a characteristic of West-Flemish. e.g. *muzze* for *muīs* 'mouse', has not been taken into consideration.

The six categories have been ordered in such a way that we expect more intelligibility problems going from category 1 to category 5. Only the respective ordering of categories 2 and 3 is not clear beforehand. Moreover, we have no basis to predict the intelligibility problems caused by category 6.

In Figure 4 the profiles are given of the four varieties in terms of the frequency of occurrence of the six categories distinguished. The categories are mutually

exclusive, so the percentages for one variety add up to 100. Let's compare the profiles in Figure 4 with the intelligibility data in Figure 3, concentrating in a first approximation on the relative frequency of the 'easy' category 1, i.e. identity, and the 'difficult' category 5, i.e. lexical difference.

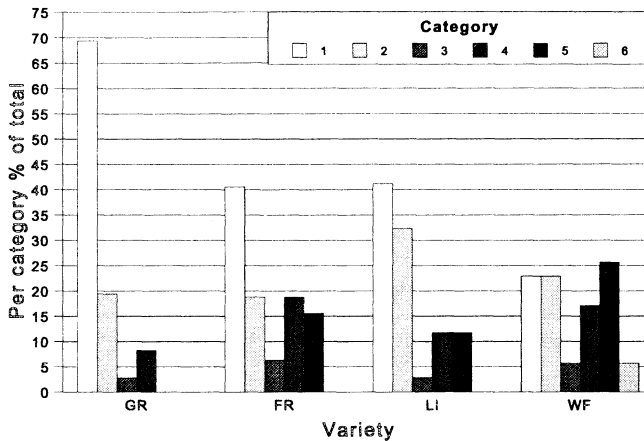


Figure 4. Percentages of target nouns in six categories of relationships with standard Dutch. 1 = identical, 2 = difference in vowel, 3 = difference in consonant, 4 = differences in several phonemes, 5 = lexical difference, 6 = semantic difference. GR=Groningen, FR=Frisian, LI=Limburg, WF=West-Flemish.

The Groningen dialect has relatively many instances (69%) of category 1 and no instances (0%) of category 5. From this one would predict that this dialect should be easy to understand for speakers of standard Dutch. And this is in fact what was found, less than 6% of the nouns were not translated correctly.

On the other hand, the West-Flemish dialect has relatively few instances (23%) of category 1 and relatively many instances of category 5 (26%). Together this can be assumed to lead to great intelligibility problems. And again this prediction is borne out by the results: 42% of the target nouns was not filled in by the listeners or translated incorrectly.

The frequencies of occurrence of categories 1 and 5 in the Limburg dialect take intermediate positions, with percentages of 41 and 12, respectively. This is in line with the intelligibility results, where the Limburg dialect is found to be more difficult to understand than the Groningen dialect (80% vs 94% correct), but easier to understand than the West-Flemish dialect (80% vs 58% correct).

Thus far, there seems to be a good agreement between the linguistic characteristics of the varieties judged (in terms of categories 1 and 5) and the ease with which they can be understood by speakers of the standard language. The case of the Frisian variety of Grou, however, seems to fit the overall pattern less well. As the Frisian variety appeared to be equally difficult to understand as the West-Flemish variety, one would expect their characteristics in terms of divergence to be comparable as well. However, as can be seen from Figure 4, this is not the case. Frisian has considerably more instances of category 1 (41% against 23%) and considerably fewer instances of category 5 (16% against 26%) than West-Flemish. So, these data point to West-Flemish being more difficult to understand than Frisian. If we consider the other categories, it can be seen that West-Flemish has some semantic differences which Frisian has not. The effect of this category is hard to estimate. The frequencies for categories 2, 3 and 4, which have to do with vowel and consonant differences, are very similar for Frisian and West-Flemish, so they do not compensate for the differences in the unequal distribution of categories 1 and 5.

Till now the approach taken to explore the relationship between the linguistic divergence of the varieties from standard Dutch and their intelligibility may be insightful but rather intuitive, and we therefore tried to quantify this relationship more objectively. This was done in the following manner. We determined separately for each of the six categories of linguistic divergence how many target nouns had been translated correctly, totalled over the four varieties. The resulting percentages were used as weighting factors and multiplied with the actual numbers of target nouns in the six categories for each of the four varieties. From this, expected overall percentages correct can be calculated and compared to the actual percentages observed. This approach yields the picture shown in Tabel 2. Again it is the Frisian variety which may be seen to be least in line with expectations, it is less intelligible than one would expect on the basis of its linguistic characteristics.

Table 2. Difference between observed and expected percentages of correctly translated target nouns

Variety	Observed	Expected	Difference
Groningen	94.4	87.1	+7.3
Frisian	58.2	71.1	-12.9
Limburg	80.3	73.9	+6.4
West-Flemish	58.0	60.3	-2.3

Closer inspection of the differences between percentages expected and observed for Frisian shows that the largest discrepancy is found for category 2, pertaining to a difference in the vowel. With respect to this category, intelligibility for Frisian is 35% lower than one would expect on the basis of the results for all four varieties combined. The percentages correct for category 2 are 36 for Frisian, 93 for Groningen, 80 for Limburg, and 68 for West-Flemish. These results suggest that whenever there is a vowel difference between a target noun and standard Dutch, the correspondence is less transparent to the listeners for Frisian than for the other varieties. This would mean that Frisian is relatively difficult to understand not only because of quantitative reasons, i.e. because of the *number* of nouns belonging to the various categories of deviation from standard Dutch, but also because of qualitative reasons, because of the *types* of deviations *within* particular categories, especially category 2.

Unfortunately, the number of target nouns within category 2 is too small (between 6 and 11 per variety) and too diverse to allow a valid comparison among the varieties. Just to give an idea, Table 3 lists for each variety one or two category 2 target words (if available) that have yielded percentages correct of less than 40 and more than 90. From this table it is not obvious at first sight what systematics could underlie the distinction between low and high intelligibility.

Table 3. Examples of category 2 target words (dialect — standard Dutch — meaning) with low (<40%) and high (>90%) intelligibility

	Groningen	Frisian	Limburg	West-Flemish
Low		rek-rok 'smoke' a:mər-emər 'bucket'	bəkər-bakər 'baker'	mʏzə-məys 'mouse' vʌzə-vas 'vase'
High	hu:s-høys 'house' kə:mər-kamər 'room'	hy:zən-høyzən 'houses'	tu:rt-tart 'pie' bɪt-bet 'bed'	wɪŋkəl-wɪŋkəl 'shop' kəs-kas 'cheese'

4. Conclusion

This experimental study used a translation task to assess the relative intelligibility of four autochthonous varieties spoken in the Netherlands and Flanders for speakers of standard Dutch. The variety of Bedum in Groningen appeared to be easy to understand, the variety of Valkenburg in Limburg occupied an intermediate position, and the varieties of Grou in Friesland and of Tielt in West-Flanders appeared to be

— equally — difficult to understand. This result could be explained to a large extent on the basis of the linguistic differences of the varieties from standard Dutch. Only the Frisian variety was found to be more difficult to understand than its linguistic characteristics suggest. One of the problems seems to reside in a lack of transparency of the vowel correspondences between the Frisian target words and their standard Dutch equivalents. A separate experiment is needed, concentrating on this issue, to explore what ideas listeners have of plausible and implausible vowel correspondences.

Notes

1. One listener was born in in Goes (province of Zeeland) but apart from the first four years was raised in the province of Zuid-Holland.

References

- Bezooijen, R. van (1994) 'Aesthetic Evaluation of Dutch Language Varieties'. *Language and Communication* 14, 253–263.
- Bezooijen, R. van (1996) 'De Kwantificering van Geografische Taalvariatie: Een Verkennde Studie naar het Tielts en het Bedums'. In: R. van Hout en J. Kruijsen, eds. *Toonzettingen en Modulaties op een Thema*. Foris, Dordrecht, 9–21.
- Bezooijen, R. van and Berg, R. van den. *Taalvariëteiten in Nederland en Vlaanderen: Hoe Staat het met hun Verstaanbaarheid?* Submitted to *Taal en Tongval*.
- Bezooijen, R. van and Pols, L.C.W. (1993) 'Evaluation of Text-to-speech Conversion for Dutch'. In: V.J. van Heuven and L.C.W. Pols, eds., *Analysis and Synthesis of Speech: Strategic Research towards High-quality Text-to-speech Conversion*. Mouton de Gruyter, Berlin/New York, 339–360.
- Boets, H. and De Schutter, G. (1977) 'Verstaanbaarheid en Appreciatie. Nederlandse Dialecten uit België zoals Inwoners van Duffel die Ervaren'. *Taal en Tongval* 29, 156–177.
- Heuven, V.J. van and Vries, J.W. de (1983) 'Verstaan, Begrijpen en Waarderen van Buitenlandse Uitspraak'. *Tijdschrift voor Tekst- en Taalkunde* 3, 180–191.
- Hoppenbrouwers, C. and Hoppenbrouwers, G. (1988) 'De Feature Frequentiemethode en de Classificatie van Nederlandse Dialecten'. *Tabu* 18, 51–92.
- Hout, R. van and Münstermann, H. (1981) 'Linguistische Afstand, Dialekt en Attitude'. *Gramma* 5, 101–123.
- Maassen, B.A.M. (1985) *Artificial Corrections to Deaf Speech: Studies in Intelligibility*. Dissertation Nijmegen University.
- Nerbonne, J., Heeringa, W., Hout, E. van den, Kooi, P. van der, Otten, S. and Vis, W. van de (1996) 'Phonetic Distance between Dutch Dialects'. In: G. Durieux, W. Daelemans and S. Gillis, eds., *CLIN VI, Papers from the sixth CLIN Meeting*. Center for Dutch Language and Speech, Antwerpen University, 185–202.