

Phonetic and grammatical variables in a project on real time change

Working paper for the IC meeting 4th – 6th June 2007

Torben Juel Jensen & Nicolai Pharao

Contents

Introduction.....	2
The grammatical variables	3
The phonetic variables	4
Changes across the life span.....	6
The grammatical variation	6
The phonetic variation	9
Comparison of grammatical and phonetic variation	13
Situation	15
The grammatical variation	15
The phonetic variation	15
Comparison of grammatical and phonetic variation	19
Context: Development during conversation.....	19
The grammatical variation	19
The phonetic variation	22
Context: Discourse context categories	26
Type of speech event.....	26
Activity type.....	26
Macro speech act.....	26
The grammatical variation	26
The phonetic variation	28
Interaction structure	30
The grammatical variation	30
The phonetic variation	33
Comparison of grammatical and phonetic variables.....	35
Genre.....	35
The grammatical variation	35
The phonetic variation	37
Comparison of grammatical and phonetic variation	39
References.....	40

Introduction

This report presents the interim results from the analyses of the grammatical and phonetic variation found in the interviews selected for the exploratory corpus of the LANCHART project, a sub-corpus consisting of a number of interviews and group conversations from the LANCHART corpus that have been selected for scrutiny (c.f. Gregersen: “An empirical solution to the problem of comparability”). These recordings were coded in their entirety for the variation of two grammatical variables and six phonetic variables, all of which are described below. This thorough investigation was carried out in order to trace in detail the progress of variation throughout a conversation.

The coding enables a detailed examination of the relationship between particular types of sequences of a conversation, e.g. voicing an opinion on a topic of general interest or disclosures of a more private nature, and the predominant choice of phonetic variants (for a full account of the discourse categories that have been registered in the exploratory corpus see the report by Gregersen (this bundle)). By coding the conversations for phonetic and grammatical variation, these exploratory studies also allow for a comparison of the course of variation at both levels of description. This allows us to examine to what extent variation at one level co-varies with variation at another. We take into account the following external factors:

Time: The subjects were recorded in the 1970ies or 1980ies and again in 2005/06. By comparing the two recordings of each speaker we will answer the following question:

Does a change in the proportions of the different phonetic variants correlate with a change in the use of the different grammatical variants?

Situation: Two settings: sociolinguistic interview and peer group conversations are represented. This allows us to examine whether or not the uses of the different variants are alike in the two situations. If there is a difference, does a change in the proportions of phonetic variants correlate with a change in the proportions of grammatical variants?

Context: We examine the distribution of variants in the course of each conversation from beginning to end. This allows us to answer questions such as:

Do speakers “switch” in their proportions of grammatical variants at the same points in time as they “switch” their proportions of phonetic variants? If speakers converge or diverge in their use of new variants during the conversation, do they do this to the same extent on both levels?

We have assembled the results from the studies of grammatical and phonetic variation here. The results will be given independently for each of the three types of investigation: 1) Change across the life span, 2) situation, and 3) context. In each section the results of the variation of the

grammatical variable is given first and then followed by the results of the study of phonetic variation. A brief summary describing the relationship between the grammatical and phonetic variation concludes each section.

The grammatical variables

The grammatical variables chosen for the exploratory study are “pronouns with generic reference” and “word order in dependent sentences”. The primary reason for choosing these variables is that they have been investigated previously, the first as part of the LANCHART project (c.f. Jensen: “Generic use of the second person pronoun in Danish – the spreading of a linguistic innovation”), the latter by Gregersen and Pedersen (Gregersen & Pedersen 1997 & 2000) as well as by a number of functional and formal grammarians, prominent among them Sten Vikner (Vikner 1999, 2001 & 2004).¹ Furthermore, and equally important, these variables may be frequent enough to allow statistical analyses of differences between particular conversations and even parts of conversations.

The variable “pronoun with generic reference” includes all Danish pronouns used to refer to an undefined person or group of persons in general. In the analysis of the exploratory corpus, though, only the variants *man* (the traditional generic pronoun originally developed from the noun *man(d)* ≈ English *man*) and *du* (second person singular) are included. Danish also has the generic pronoun *en* (≈ English *one*), and both the third person singular pronoun *den* (it) and the first person plural *vi* can be used with generic meaning, but to all appearances these pronouns do not play any significant role as generic pronouns in the varieties of Danish represented in the exploratory sub-corpus - at least not as regards generic pronouns functioning as grammatical subjects. The pronoun *man* can only function as grammatical subject, in other syntactic functions the pronoun is complemented by *en*. Therefore only pronouns functioning as grammatical subjects are included in the analysis.

The analysis of a larger part of the LANCHART corpus has established a change in the use of generic pronouns from the first to the second recording of three groups of 24 speakers each from Odder, Næstved and Copenhagen. The second person pronoun *du* has gained ground at the expense of *man* during the intervening 20 years at all three subgroups of speakers seen as wholes. At the same time the results show significant differences between the speakers when divided into subgroups on the basis of geographical location, class and gender. The results of the exploratory

¹ Due to problems with our search engine it has not, as of yet, been possible to analyse the variation in word order in dependent sentences.

study of particular speakers will be compared with the results of the relevant subgroups of speakers as regards speakers from Odder, Næstved and Copenhagen.

The phonetic variables

6 phonetic variables were chosen for this initial coding². They were chosen primarily because they have been investigated previously (all of them in Copenhagen, but some also in Odder and Næstved (cf. Gregersen & Pedersen (1991), Nyberg & Nielsen (1993), and Normann Jørgensen & Kristensen (1994)). At least two of the variables were hypothesized to be changes in progress in the original studies of linguistic variation in Copenhagen, namely the raising of [ɛ] to [e] before velar nasals (the EN variable) and the lowering of [u] to [o] after /r/ (the RU variable) - a surprising change in Danish phonology since no other high vowels are affected by neighbouring /r/, whereas all non-high vowels are lowered in this context. The rest of the variables involved the variation of the phoneme short /a/ in different contexts: before alveolars and nothing (AN), before labials and dorsals (and in words containing /r/ in post tonic syllables – the AM variable), the syllabic component of the diphthong [aj] (AJ) and /a/ after homosyllabic /r/ (RA). The variables are given in table 1 below.

² This is a somewhat small number of variables. Once the investigations into the types of discourse sequences that can be seen as the most representative of the kinds of interaction that take place within the sociolinguistic interview have yielded a principled way of limiting the amount and kinds of speech events to be compared, more variables will be added to the set.

Table 1 – Phonetic variables in the exploratory study

Variable name	Standard variant	Innovative variant
AN	[æ]	[ɛ̝]
AM	[ɑ]	[ɑ̝]
AJ	[ɑ]	[ɑ̝]
RA	[a]	[ɑ]
EN	[ɛ]	[ɛ̝]
RU	[u]	[o]

The designation “standard” and “innovative” assigned each of the variables here rests on the results from some of the previous studies. For the variables pertaining to short /a/ the raising, retraction and lowering in the respective segmental contexts are described as markers of social class in Copenhagen with all of these variants being typical of the speech of working class informants (cf. Gregersen & Pedersen

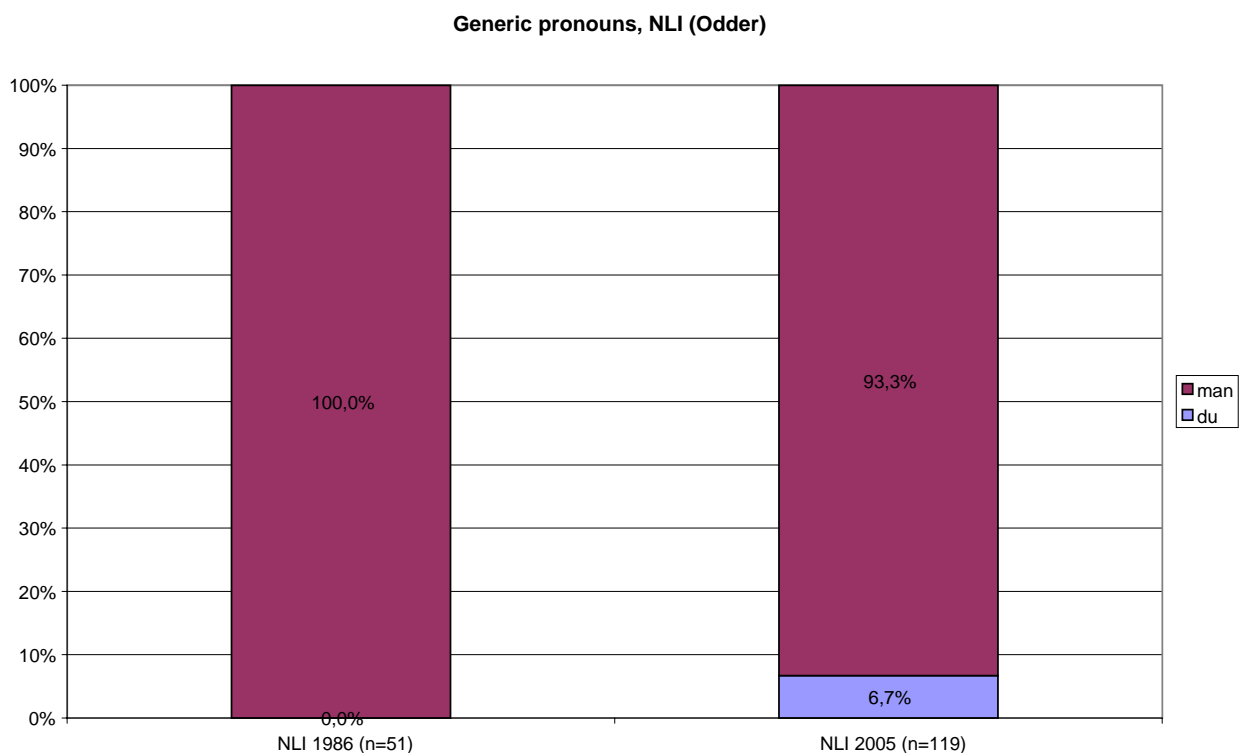
(1991)). The raising of /a/ before alveolars and syllable boundaries was found to be increasing in the speech of working class adolescents in Næstved (cf. Normann Jørgensen & Kristensen (1994)). The variables EN and RU are both particular to the Copenhagen study, although the raising of [ɛ] to [ɛ̝] characterising the EN variable has also been reported as an innovation in rural Zealandic varieties (Køster (1996)). The lowering of [u] to [o] after /r/ is described as an innovation in the speech of young Copenhageners in the original Copenhagen study and in a recent textbook on the phonetics and phonology of Danish (Grønnum (2005)).

Unfortunately, we do not have results from the entire group of speakers from each of the regions studied in the LANCHART project, so unlike the grammatical variation, we are not able to compare the analyses of individuals in the exploratory corpus to the phonetic variation found in the larger corpus.

Changes across the life span

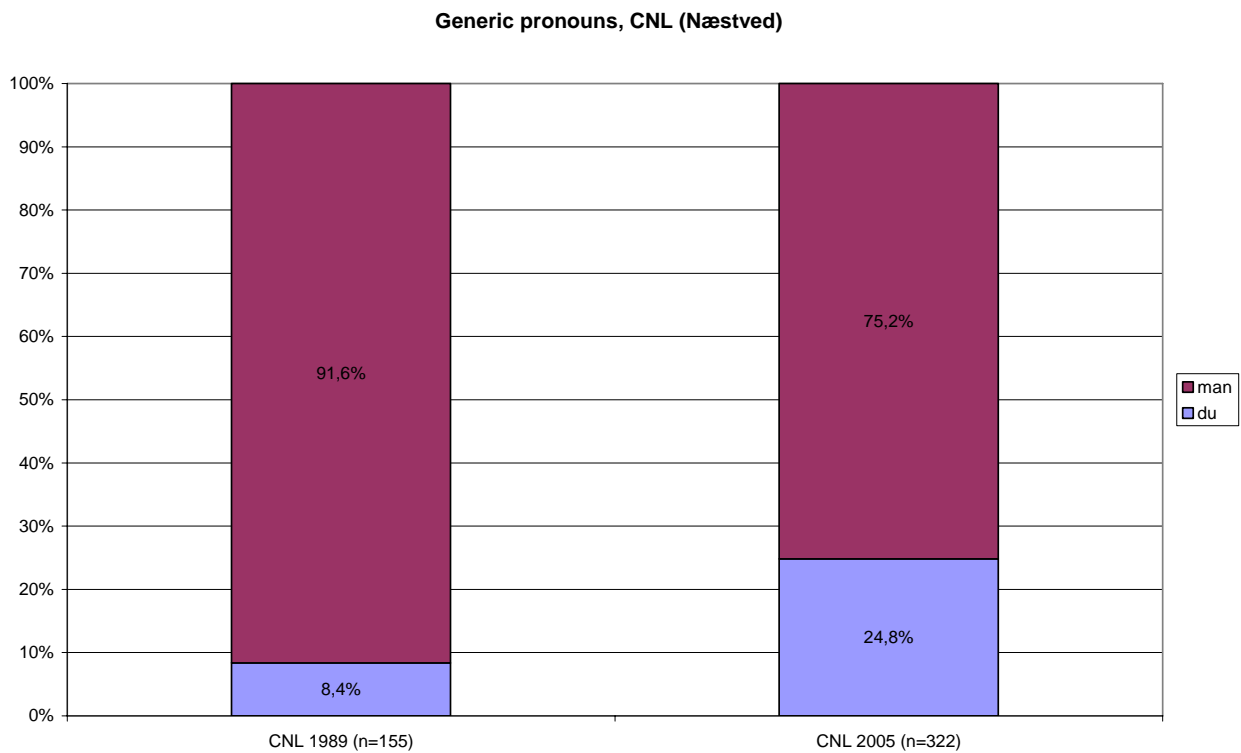
The grammatical variation

Five of the speakers represented in the corpus chosen for the exploratory analysis are recorded twice with approximately 20 years in between allowing us to evaluate whether or not they have changed their use of generic pronouns across their life span. One of them, TCH, uses generic pronouns too rarely to allow statistical analysis; the results for the other four speakers are shown below:



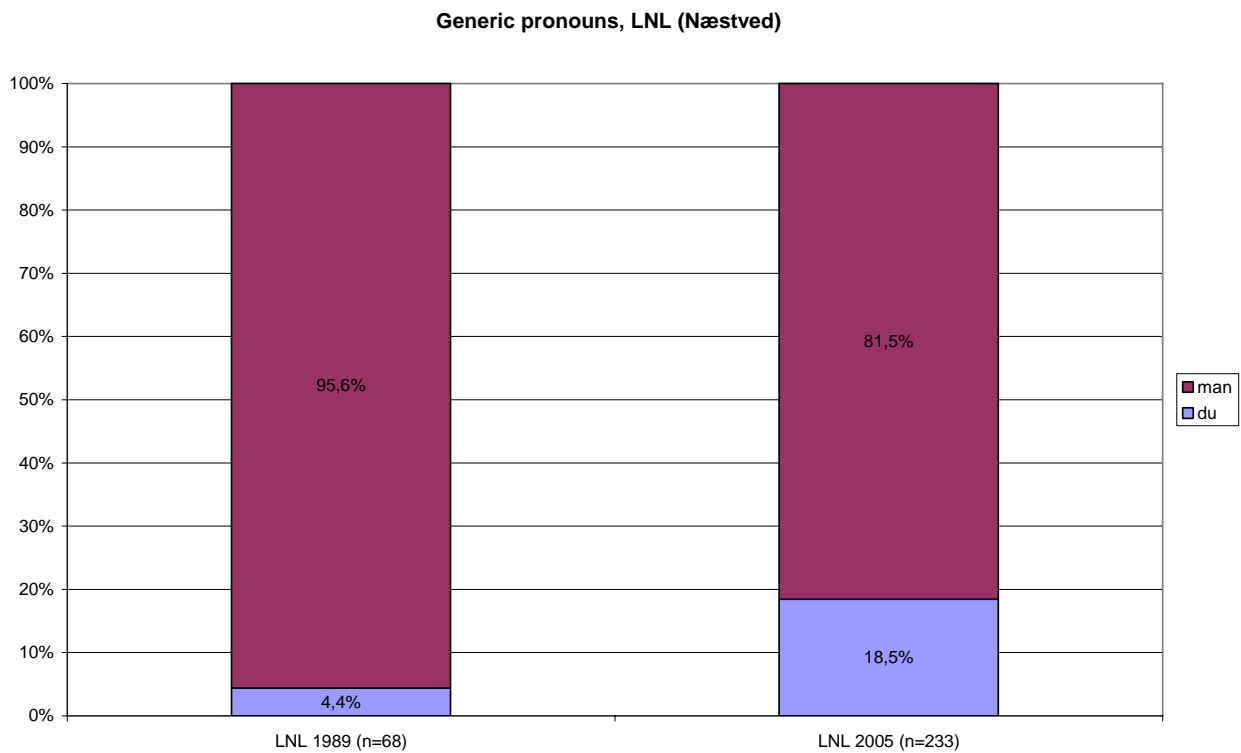
($\chi^2=2,25$; 1 d.f.; not significant)

The share of *du* rises from zero in the first recording to 7 % in the second, but the difference is *not* statistically significant. NLI's share of *du* is lower than the mean of the group of female working class speakers in Odder, both in the old and the new recording where the share of the group as a whole are 9 and 18 %, respectively.



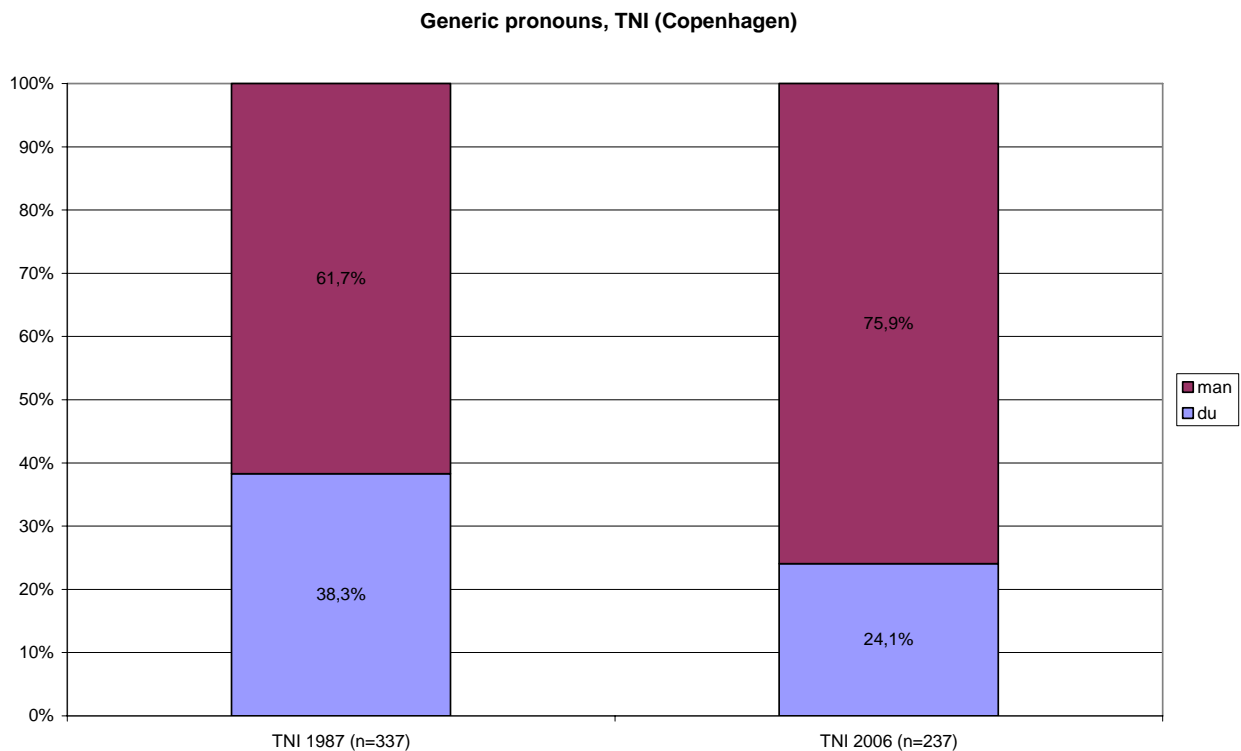
($\chi^2=17,02$; 1 d.f.; $p<0,01$)

The share of *du* rises from 8 to 25 %. This is the same tendency as in the group of working class men in Næstved seen as a whole, but CNL's share of *du* is lower than the mean of the group where the share of generic *du* in is 24 and 39 % in the old and new recordings, respectively.



($\chi^2=6,96$; 1 d.f.; $p<0,01$)

The share of *du* rises from 4 to 19 %, and in this respect LNL is *not* representative of the group of middle class women in Næstved seen as a whole where the share of *du* declines from 18 to 15 % from the old to the new recordings.



($\chi^2=12,21$; 1 d.f.; $p<0,01$)

TNI's share of *du* declines from 38 to 24 %; in the group of middle class men in Copenhagen seen as a whole the share of *du* is stable in the old and new recordings (28 and 27 % respectively).

The phonetic variation

The phonetic behaviour of three informants in the original and the new recordings are available for statistic analysis: two female speakers from Odder (TCH and NLI), both classified as working class, and one middle class male speaker from Copenhagen (TNI).

The Odder speakers

As far as the two female speakers are concerned the results are quite clear if somewhat disappointing: there are no statistically significant changes from the old to the new recordings with respect to any of the 6 phonetic variables. This may not be so terribly surprising since these speakers are from a region rather far from Copenhagen, and all the variables are hypothesized to originate there. Thus, these analyses reconfirm the results from last year's investigation of shorter passages of more speakers from this region, which indicated that the Copenhagen trait of raising of short /a/ had not spread. Apparently none of the Copenhagen traits have spread to Odder.

Both of the female speakers from Odder are classified as working class speakers. However, they belong to different age groups, with NLI being the oldest. As mentioned above neither of them show any change significant changes over time, but a comparison of the two interviews at the same point in time reveals a statistically significant difference ($p < 0.005$) for the AN variable in the distributions of the variants between the two speakers in the new interviews. There was no statistically significant difference between the two speakers with respect to any of the variables in the old recordings. The distributions show the same over all pattern for AN for both speakers, a preponderance of standard non-raised [æ], but in the new interviews the older speaker has a higher share of raised [æ̥] (the non-standard, innovative variant) than the younger speaker. Sketchy as a comparison of two individuals might be, it does indicate a possible development of raised [æ̥] as a marker of generations in the contemporary speech of Odder, and it will be interesting to see whether this pattern of distribution is found, when we examine the complete Odder-corpus.

The Copenhagen speaker

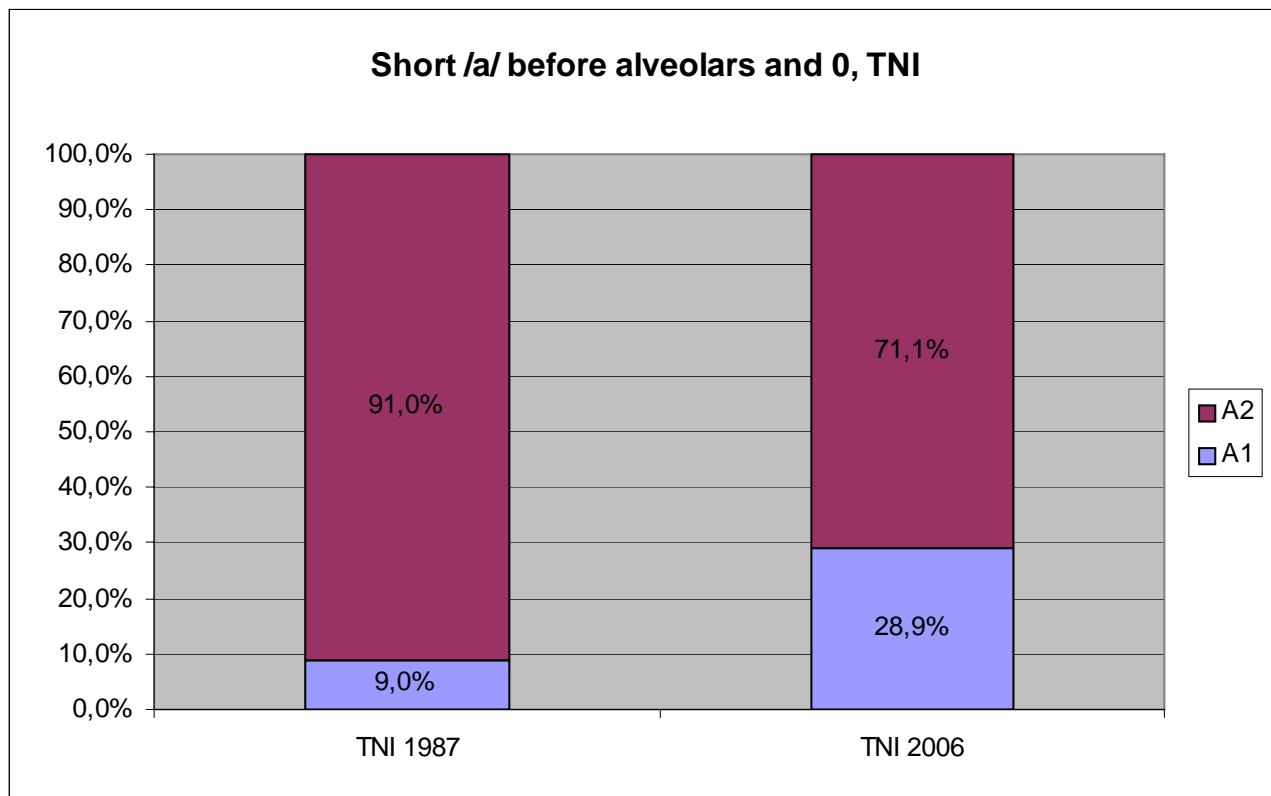
Table 2- Phonetic variables showing change in real time for the middle class Copenhagen speaker

AN	% innovative	$p <$	n
1987/old	9		624
2006/new	29	0.001	837
RA	% innovative		
1987/old	45		98
2006/new	67	0.01	63
EN	% innovative		
1987/old	28		173
2006/new	9	0.001	228

Not all of the variables show significant changes for the male Copenhagen speaker either, but half of them do. Of the 3 variables that show no change, one (RU, i.e. the lowering of [u] to [o] after /r/) is simply impossible to measure having so few occurrences in the original interview (and in fact none in the new one) that analysis be-

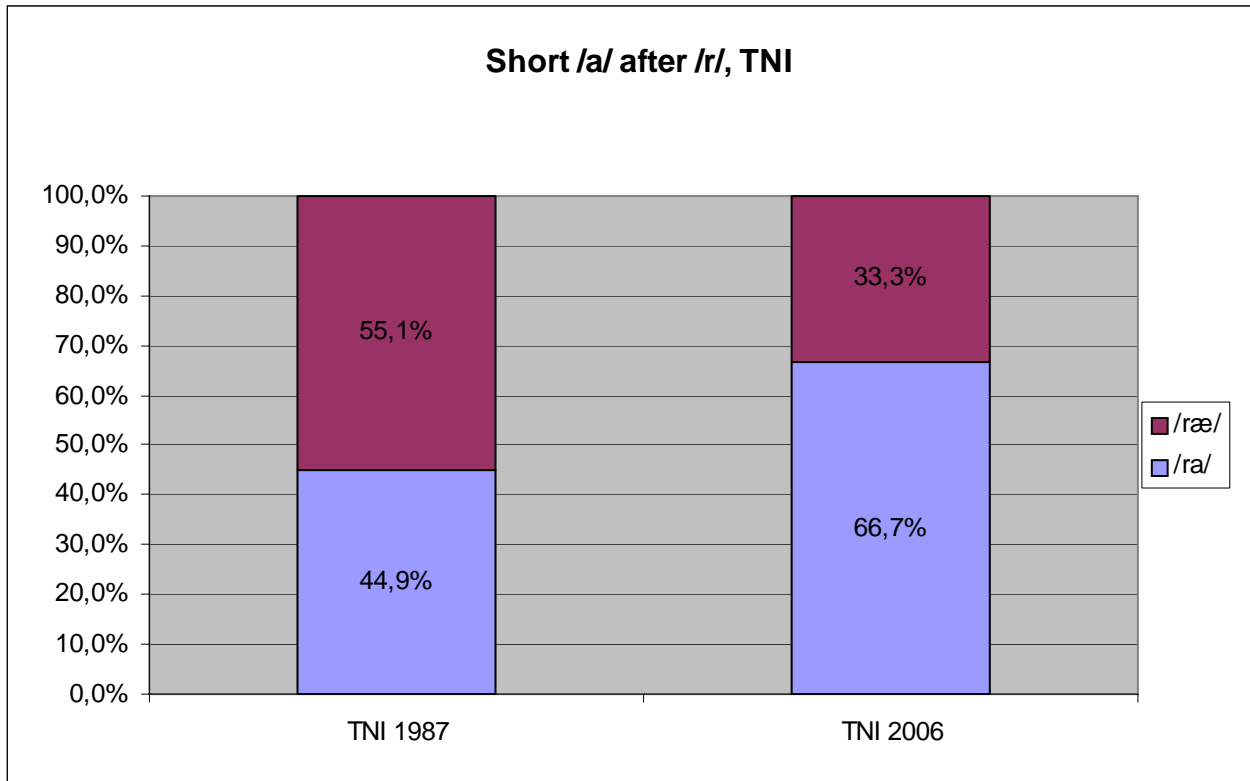
comes superfluous. The other two, AM and AJ, show no change over time but AJ does show change from interview to group conversation, (more on which in the next section). The three variables that do change over time are AN, short /a/ before alveolars and nothing, RA, short /a/ after /r/, and EN, the raising of /ɛ/ to [ɛ̥] before velar nasals. The first used to be sociolinguistically quite salient, with the marked variant having a popular name, the second is popularly known as a generational variable and a few potential mergers are commonly known (e.g. ‘ret’ and ‘rat’ (Eng.: ‘straight’ and ‘steering wheel’) respectively [ɾæd̥] and [ɾad̥] in the older norm, now completely

merged to [ɤɑ̃] in the younger generation). The EN variable is perhaps not very salient sociolinguistically speaking, although some speakers seem to be aware of it as a generational marker. The data for this middle class Copenhagen speaker indicate that raised [ɛ̃] may be a variant in decline. The results are shown below:



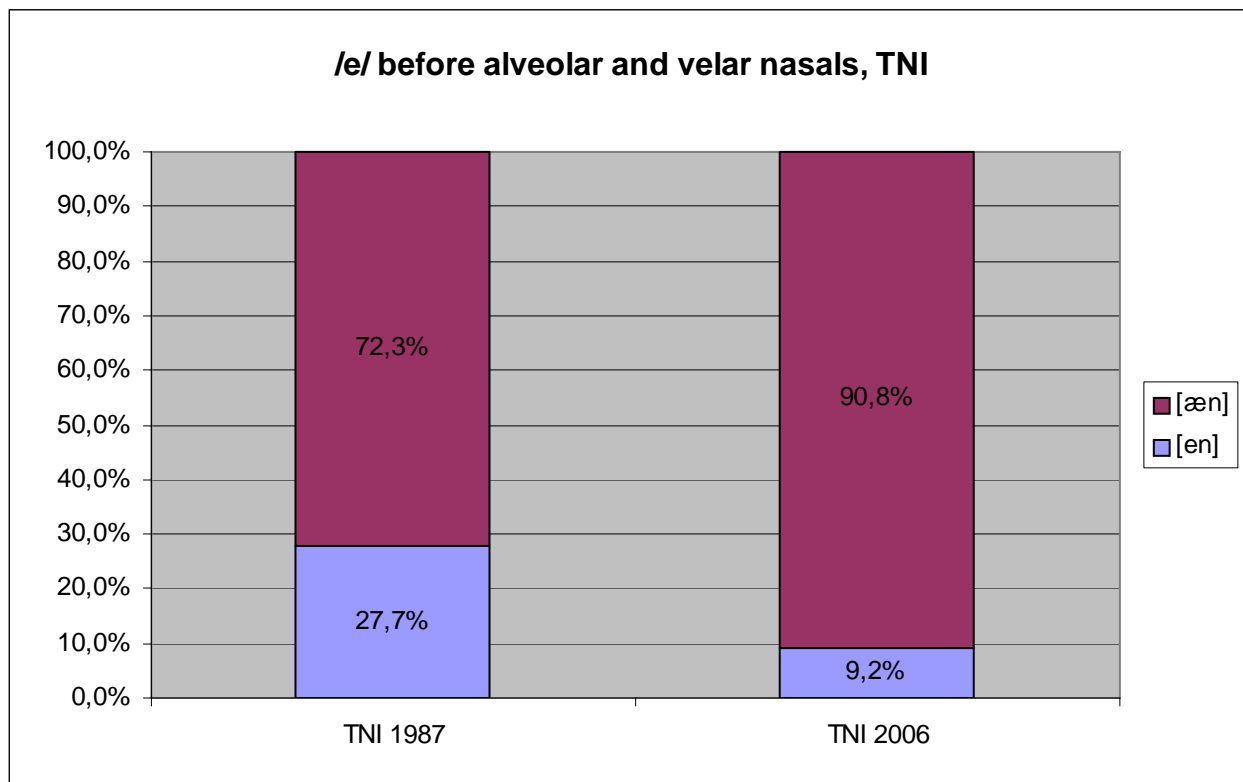
(A1 = [æ̃], A2 = [æ])

The difference is significant with $p < 0.001$ ($n=837$ in TNI 1987, $n=624$ in TNI 2006) showing an increase in the use of the raised variant of almost 20% from the time of the original recording to the new one.



(/ræ/ = [ɾæ]; /ra/ = [ɾɑ])

This difference is also significant, this time with $p < 0.01$ ($n=98$ in TNI 1987, and $n=63$ in TNI 2006). The data show an increase in the use of lowered (but not retracted) [ɑ] after /r/ of almost 22%.



([æ̃n] = [ɛ̃]; [ɛ̃n] = [ɛ̃])

The last significant difference between the original and new recording of TNI with $p < 0.001$ ($n=173$ in TNI 1987, and $n=228$ in TNI 2006), but this one showing a decrease in the use of the innovative variant, raised [ɛ̃] of 18.5 %. This change is in the opposite direction of the two other variables.

Comparison of grammatical and phonetic variation

Comparing the grammatical and phonetic variation, we get rather different impressions depending on which area we look at. In Odder only the grammatical variation of one of the speakers is possible to analyse, and although she exhibits a slight increase in the use of ‘du’ as a generic pronoun, the difference from original to new recording is not statistically significant. This matches the analysis of the phonetic variation for both speakers from Odder. None of them show any significant difference in the distribution of variants, i.e. no actual change in real time for any of the six variables. In Odder, then, grammatical and phonetic variation (as measured by this set of variables) would seem to follow each other – in these particular cases in *not* changing.

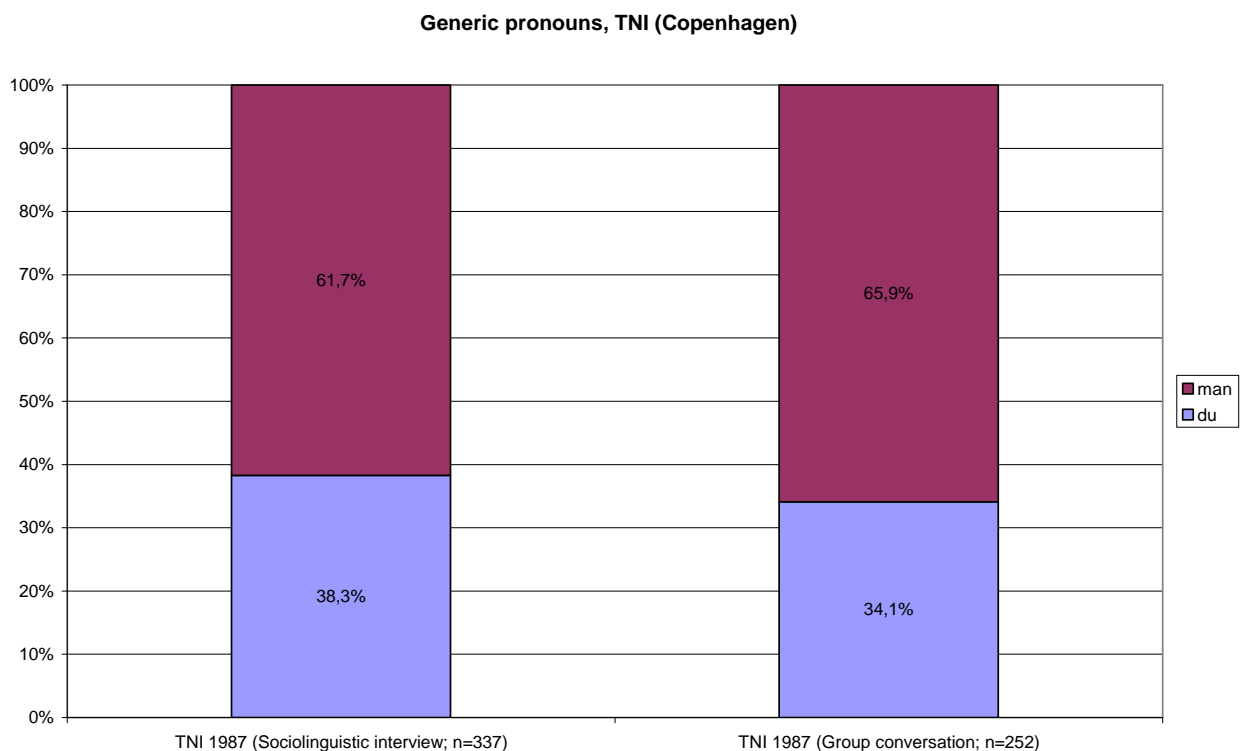
In Copenhagen we see a different picture. First of all, the speaker we have analysed does show evidence of change across the lifespan: he has a significant decrease in the use of ‘du’ from the old to the new recording. He also exhibits changes in pronunciation with significant increase in the use of

raised [æ] and [ɑ] after /r/, but a decrease in the raising of /ɛ/ before velar nasals. Following the classification of phonetic variants as innovations given above it would seem that the EN variable is declining and that the ‘generic pronoun’ variable is following it. Judging from these comparisons it would seem that grammatical and phonetic variation *may* co-vary, but that it doesn’t necessarily do so.

Situation

The grammatical variation

Five speakers represented in the corpus chosen for the exploratory analysis are recorded in two situations: sociolinguistic interviews and group conversations. Three of them, PEK, MTK and LHS, do not use generic *du* at all, neither in the old nor in the new recordings, and one, JBO, uses generic pronouns too rarely to allow statistical analysis. The results of the last speaker, TNI, are shown below:



($\chi^2=0,90$; 1 d.f.; not significant)

The share of *du* is lower in the group conversation than in the sociolinguistic interview, but the difference is not statistically significant.

The phonetic variation

Given the focus of this report, the co-variation of grammatical and phonetic variation, the analysis is confined to the male middle class Copenhagen informant TNI whose (lack of) grammatical variation as a function of situation is described above.

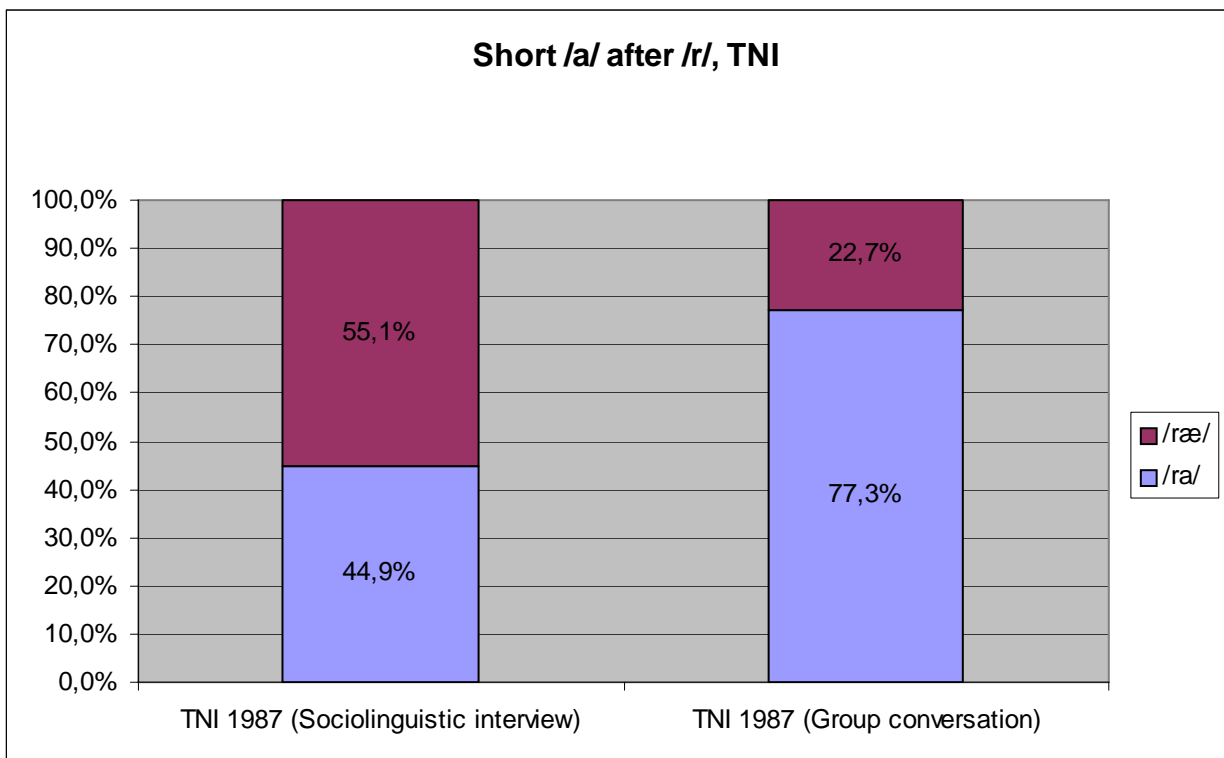
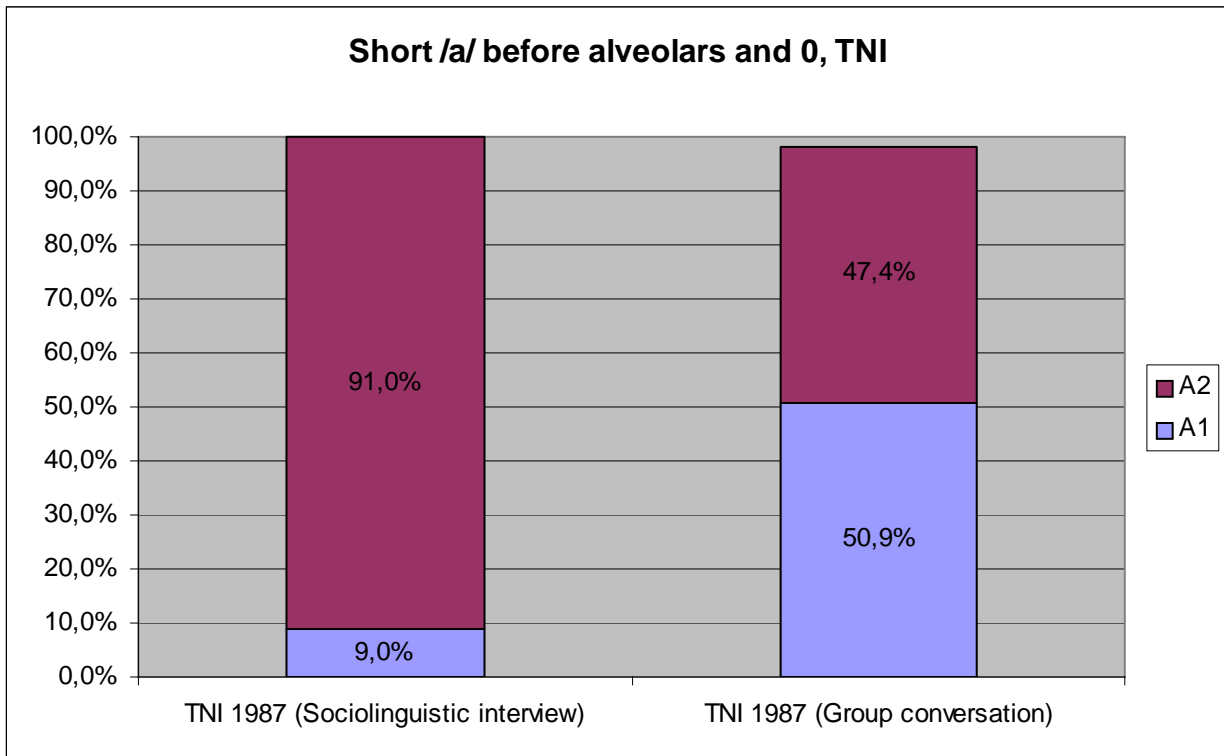
The data from the comparison is given below, limited to the variables that show statistically significant differences.

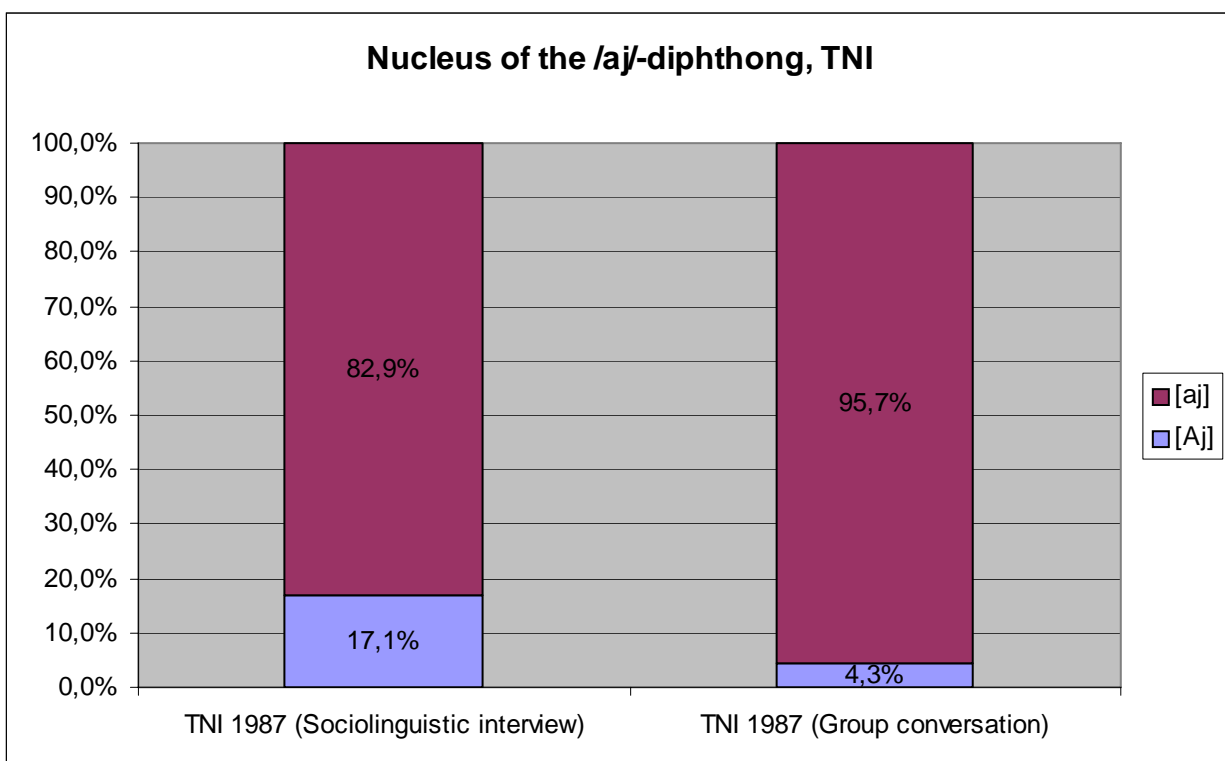
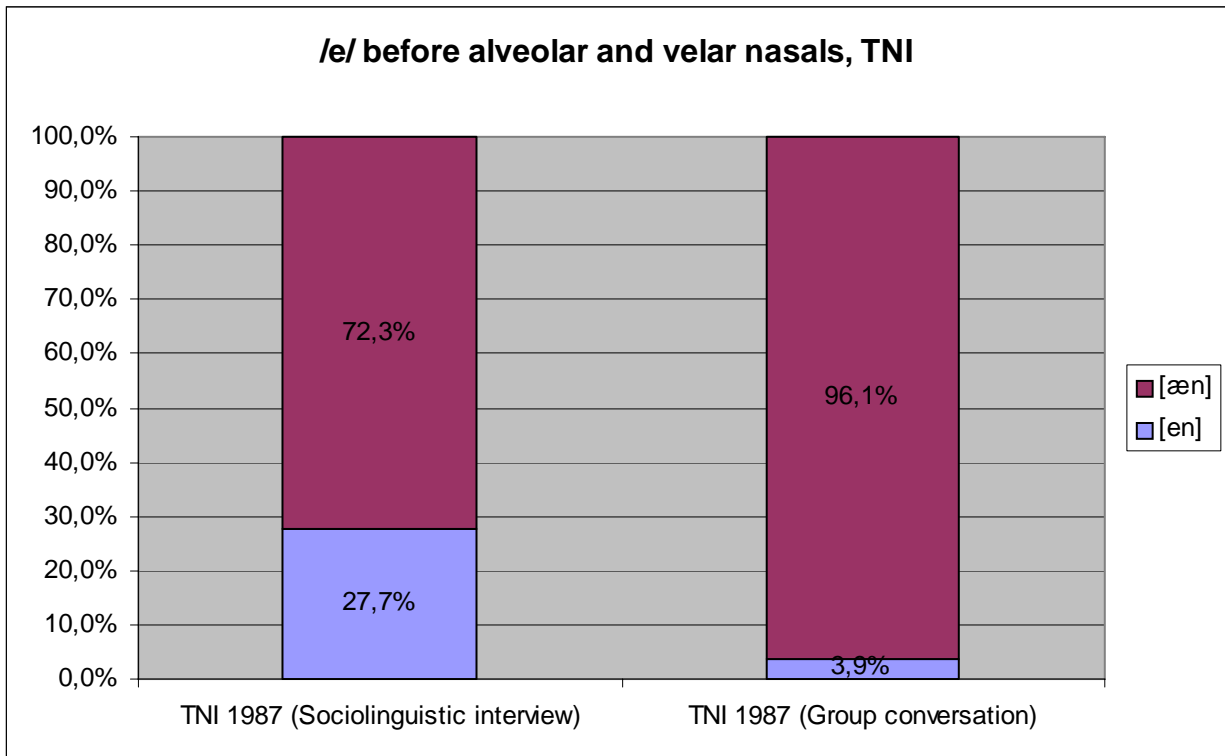
Table 3 – Comparison of interview and group conversation for the middle class male TNI

AN	% innovative	$p <$	n
int 1987	9		624
group 1987	51	0.001	557
AJ	% innovative	$p <$	n
int 1987	17		170
group 1987	4	0.01	70
RA	% innovative	$p <$	n
int 1987	45		98
group 1987	77	0.001	44
EN	% innovative	$p <$	n
int 1987	28		173
group 1987	4	0.001	155

Three of the four variables are the same as the ones that show change in real time for this speaker, namely AN, RA and EN. All three variables change in the same direction from interview to group as they do in real time, in line with the commonly held belief/hypothesis that the more casual mode of speech captured in group conversations is the kind of speech that has become standard (in other words, features from informal speech seep into formal speech – if indeed formality is the relevant parameter. We return to this in the context analysis below). However, one thing is worth noting in this

respect: whereas the AN and RA variables both show an increase in the use of innovative variants both in real time and as a function of situation, the EN variable shows a significant *decrease* (and a rather drastic one at that) as also noted above in the section on real time changes. The same is true of the AJ variable, which also shows a decrease in the use of the innovative variant when interview and group conversation are compared (recall that there was no change in real time for this variable). The results are displayed graphically below.





([aj] = [aj]; [Aj] = [ɔj])

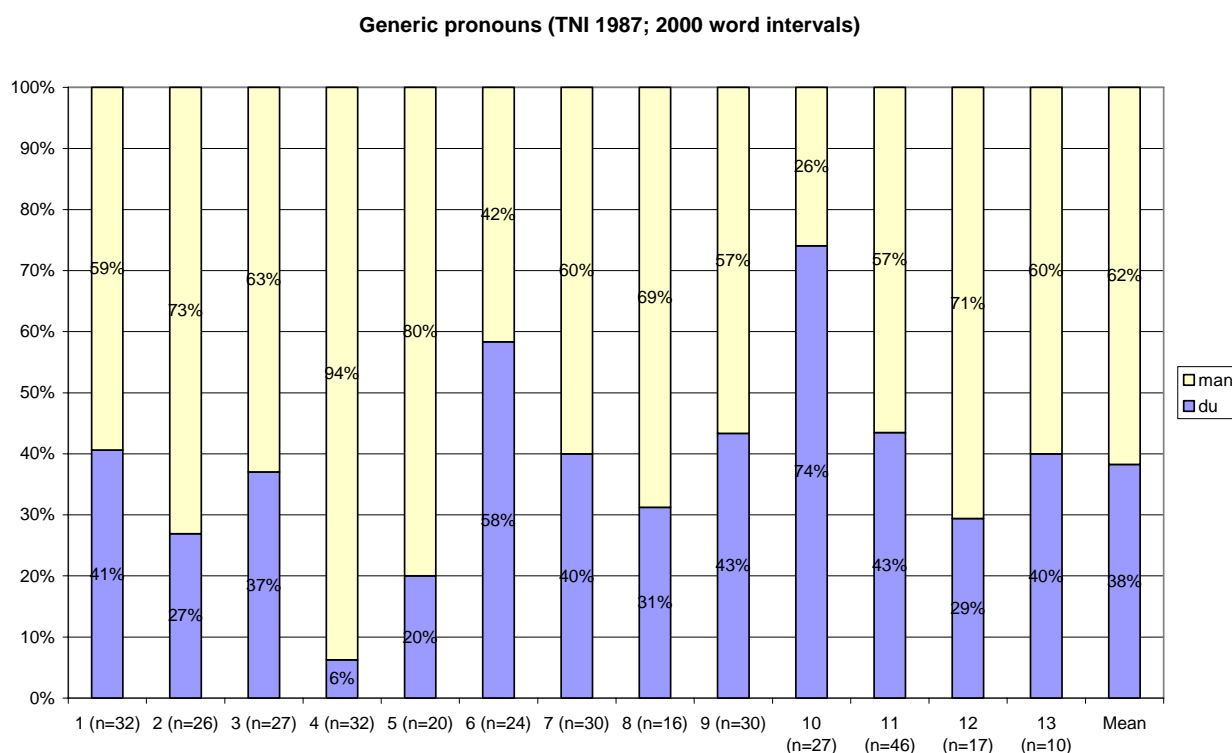
Comparison of grammatical and phonetic variation

It would appear that the difference between sociolinguistic interviews and group conversations affects pronunciation more than it affects the use of pronouns. Whereas the grammatical variable shows no statistically significant variation from interview to group conversation, 4 out of 6 phonetic variables all show significant differences. As before, in the comparison of recordings made 20 years apart, the changes in the variables do not all go in the same direction. This discrepancy is the first evidence we have of a *total* lack of correlation between grammatical and phonetic variation.

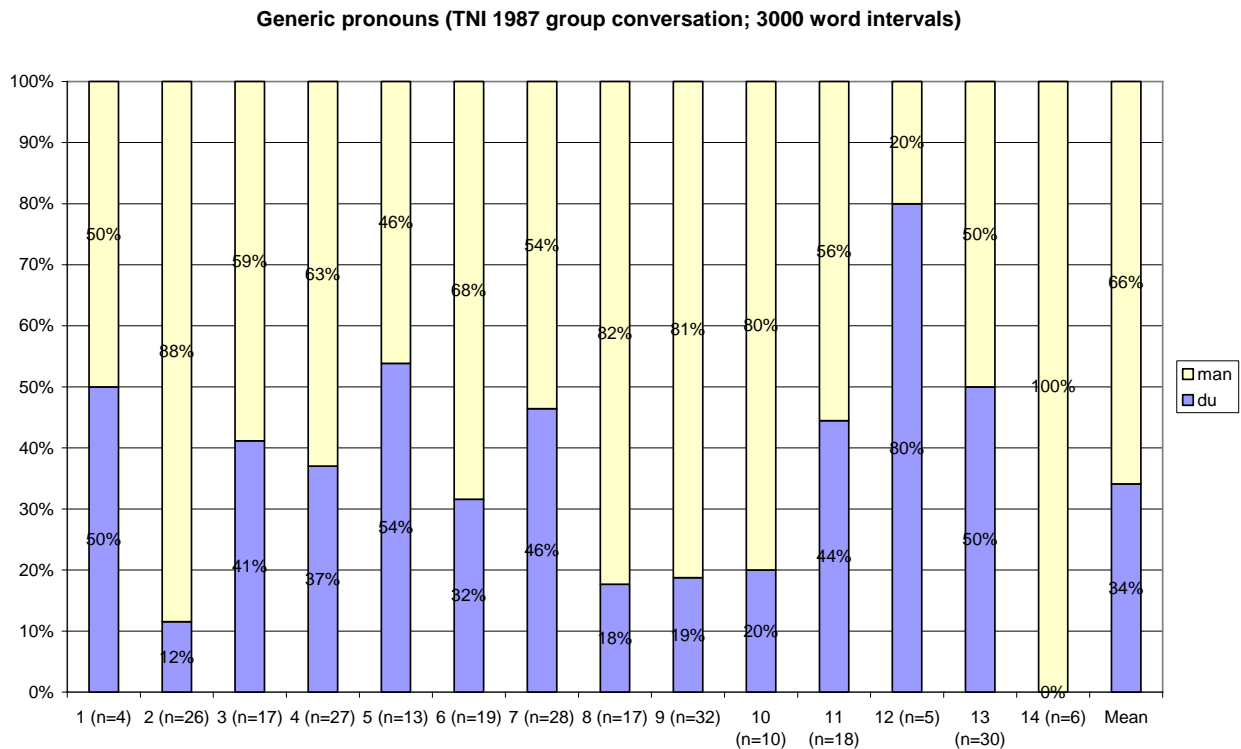
Context: Development during conversation

The three conversations with the speaker TNI, two sociolinguistic interviews and one group conversation, was split up from beginning to end in parts containing 2000 or 3000 words each. This allows us to examine the distribution of variants in the course of each conversation from beginning to end. The results are shown below.

The grammatical variation

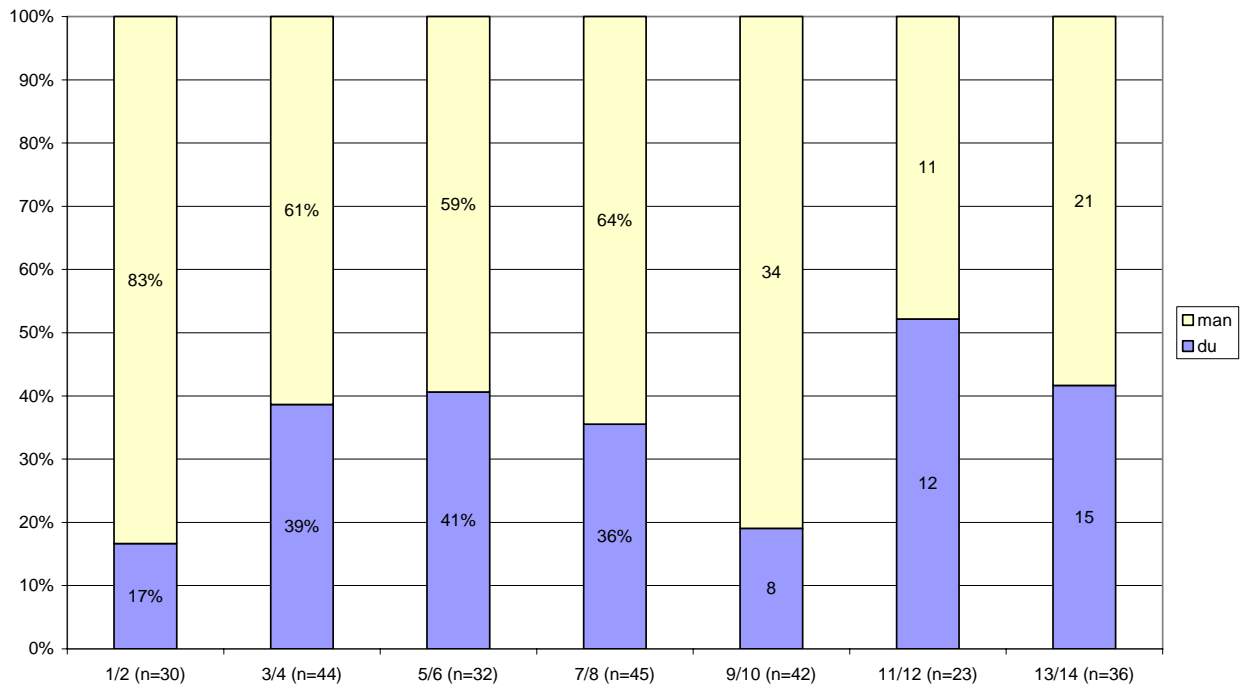


The distribution of generic *du* is highly inhomogeneous ($\chi^2=38,76$; 12 d.f.; $p<0,01$). Intervals 6 and 10 have relatively high shares of *du* while 4 and 5 have relatively low shares.

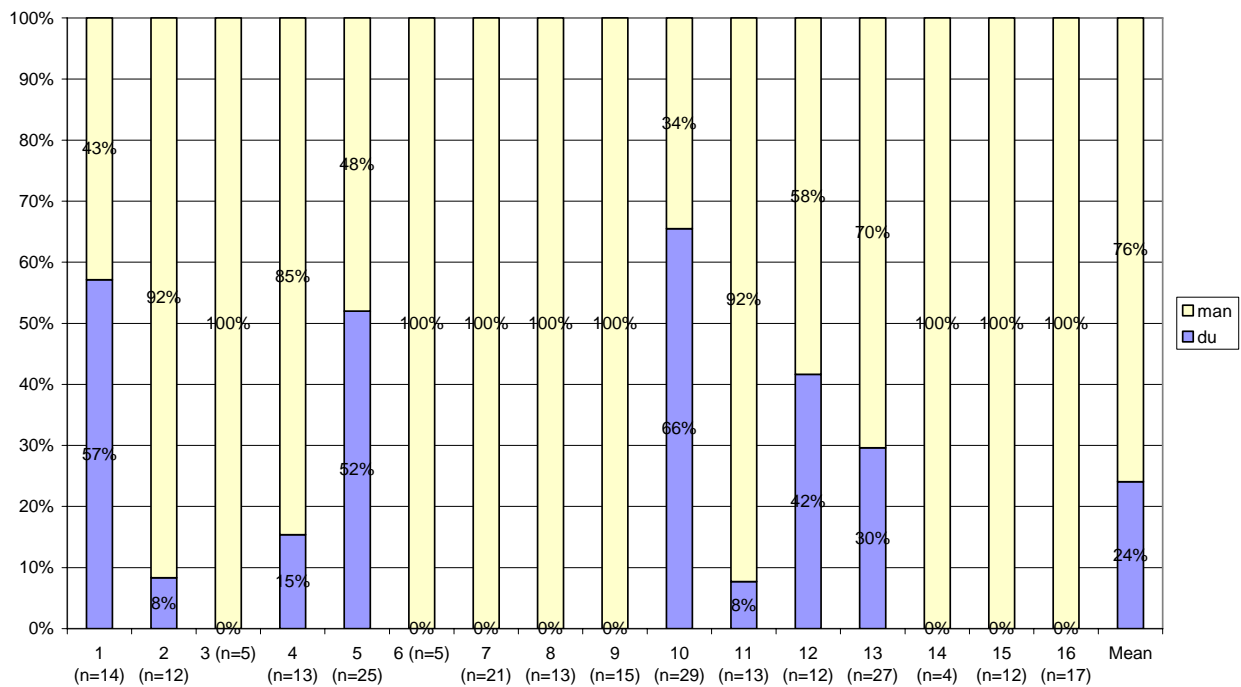


A chi square test of the distribution is not possible for the group conversation as too many of the expected frequencies of *du* in the intervals are less than five, but the tendency is relatively high shares of *du* in the intervals 1, 5, 7 and 11-13 and relatively low shares in 2, 8-10 and 14. A chi square test is possible if the intervals are collapsed in pairs, and the distribution of variants is significantly inhomogeneous ($\chi^2=13,60$; 6 d.f.; $p<0,05$). Interval 1/2 and 9/10 have a relatively low share of *du* while interval 11/12 has a high share:

Generic pronouns (TNI 87 group conversation)

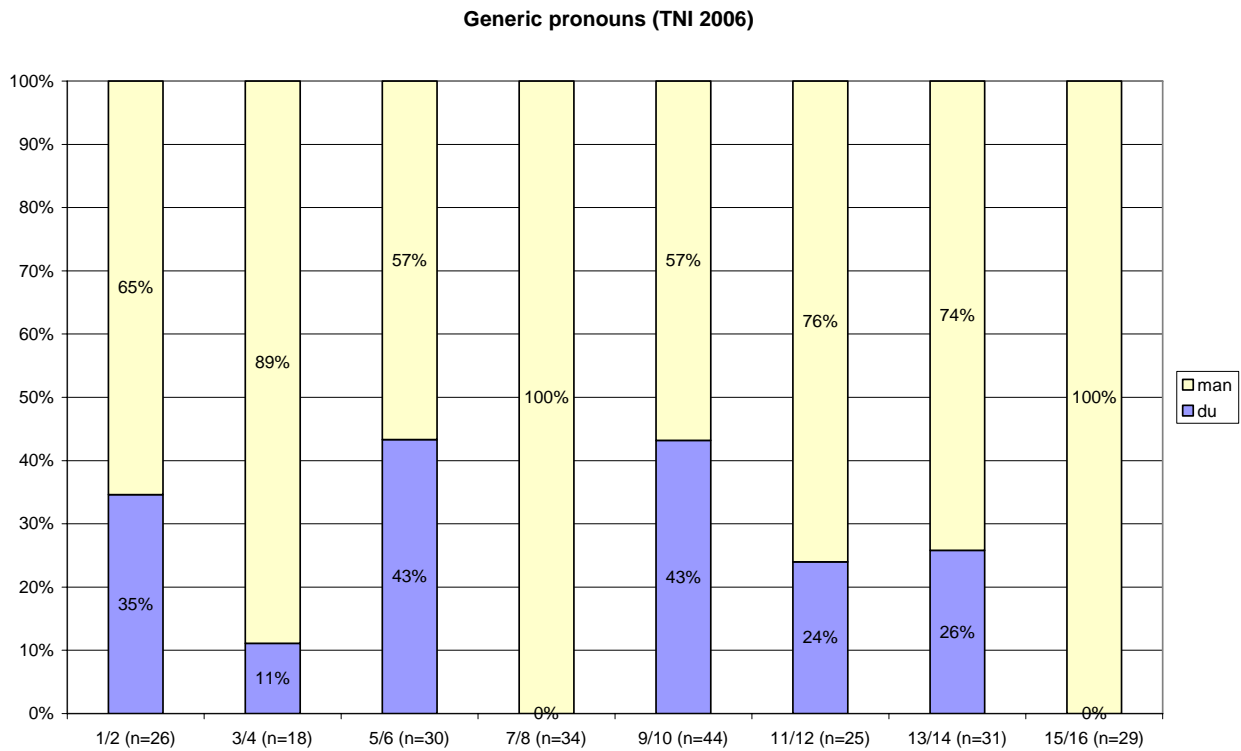


Generic pronouns (TNI 2006; 2000 word intervals)



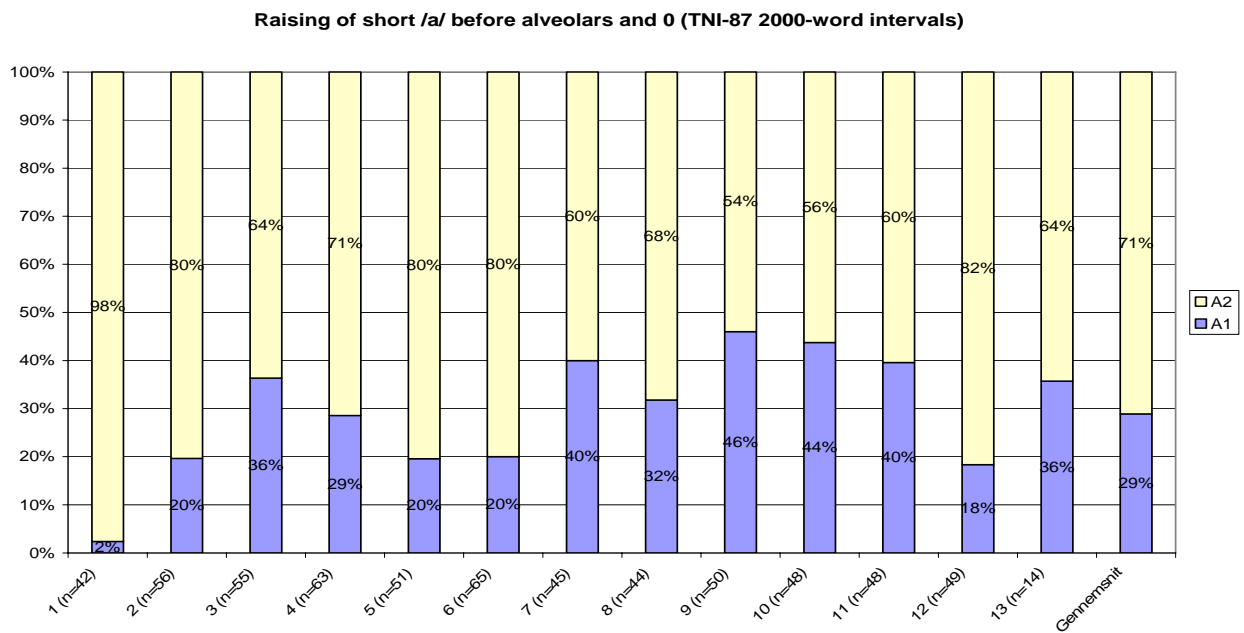
A chi square test is not possible as too many of the expected frequencies of *du* in the intervals are below 5, but the distribution of generic *du* is obviously very inhomogeneous: Generic *du* only

occurs in interval 1-2, 4-5 and 10-13. A chi square test is possible if the intervals are collapsed in pairs, and the distribution of variants is significantly inhomogeneous ($\chi^2=38,16$; 7 d.f.; $p<0,01$). Interval 7/8 and 15/16 have a relatively low share of *du* while interval 5/6 and 9/10 have high shares:



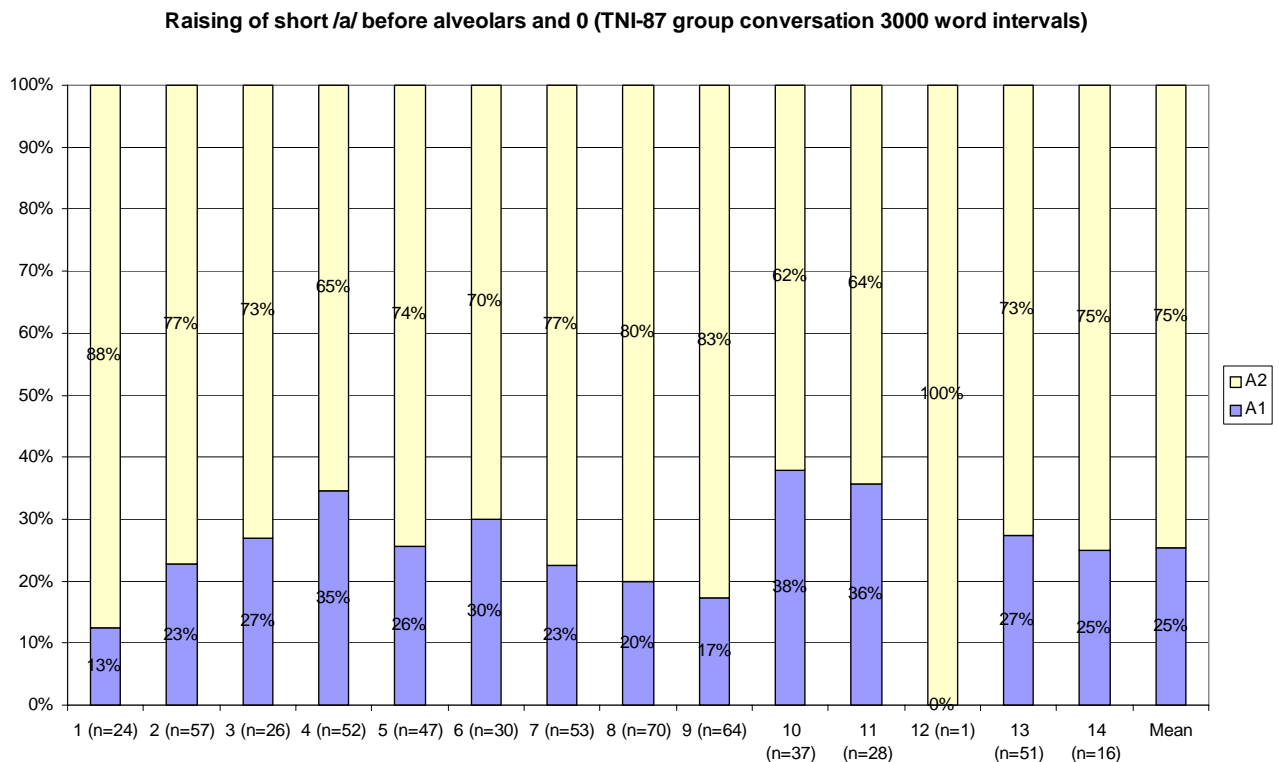
The phonetic variation

For the time being this analysis is confined to the variation found for the AN variable. First the results from the original interview:

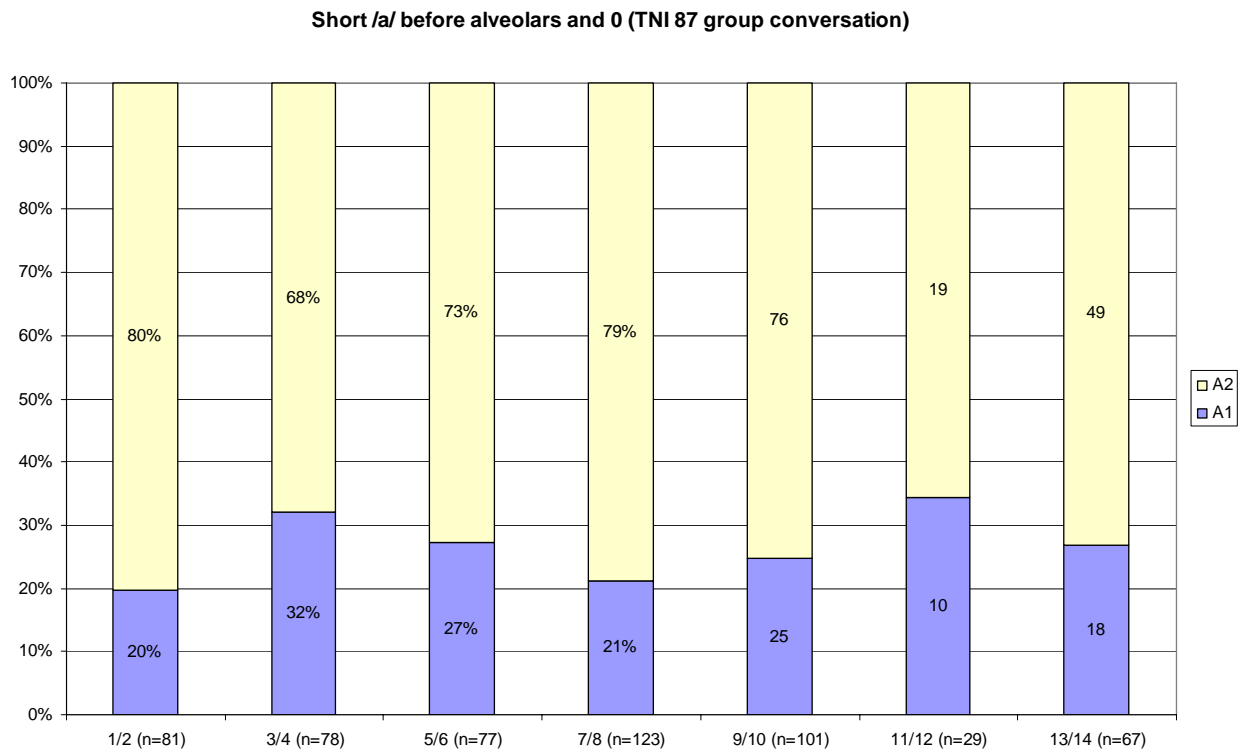


The distribution is quite inhomogeneous ($\chi^2 = 43,64$; 12 df; $p < 0,001$) with interval 1 containing by far the lowest share of raised [æ] and intervals 9 and 10 have relatively high shares of the innovative variant. For the remaining variables, the expected values are so small for each interval as to make χ^2 -testing impossible.

Next the results of the group conversation, again for the AN variable only:



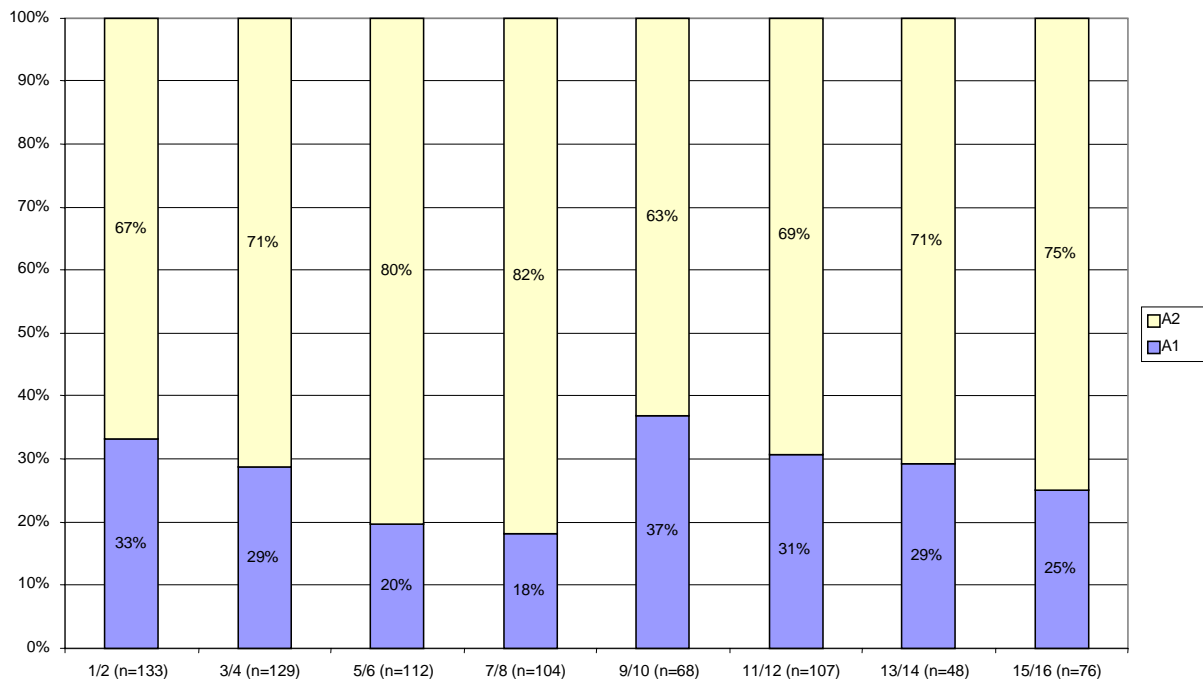
This distribution is not significantly inhomogeneous according to the chi square test, but the test isn't really possible since interval 12, which contains only 1 token of the AN variable, has an expected value of less than one, rendering the chi square test invalid. Instead the results of a pairwise collapsing of the intervals is given here:



Again, chi square testing shows a homogeneous distribution in spite of the apparent peaks at intervals 3/4 and 11/12.

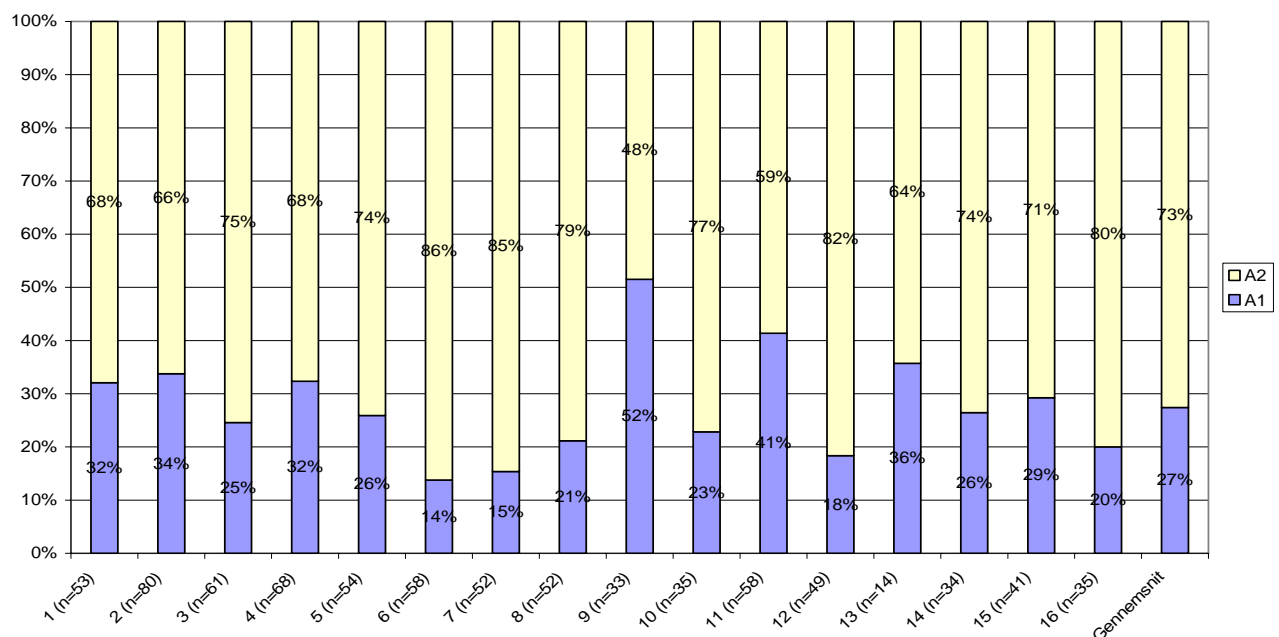
Although not strictly necessary the intervals have been collapsed for the 2006-interview in order to facilitate comparison with the grammatical variable. However, since the expected values do allow for the more fine-grained display, this is given after the display where the intervals have been collapsed.

Short /a/ before alveolars and 0 (TNI 2006, 4000 word intervals)



This collapsed view of the variation of short /a/ does not show inhomogeneity ($\chi^2 = 13,94$; 7 df; $p > 0,05$). But looking at the 2000 word intervals we find a rather more inhomogeneous distribution ($\chi^2 = 32,78$; 15 df; $p < 0,005$) with the intervals 6 and 7 showing a relatively low share of raised /a/, whereas intervals 9 and 11 have relatively high shares of the innovative variant.

Short /a/ before alveolars and 0, (TNI 06 2000 word intervals)



Context: Discourse context categories

The three conversations with TNI have been analyzed according to discourse contexts in five "dimensions": *type of speech event, activity type, macro speech act, interaction structure* and *genre*. The categorizations of the first four dimensions are exhaustive - all there is in the data has to belong to one or another of the categories – whereas the categorization into genres is selective: passages are only coded where relevant. This means that only between a fourth of and half the words uttered by TNI in the three conversations are coded as belonging to a specific genre.

Type of speech event

As TNI's speech in both the sociolinguistic interviews is exclusively categorized as "single person interview with informant unknown to the interviewer" and in the group conversation exclusively as "group conversation with informants unknown to the interviewer" this dimension has already been covered by comparing the sociolinguistic interview from 1987 with the group conversation (c.f. pp. 15).

Activity type

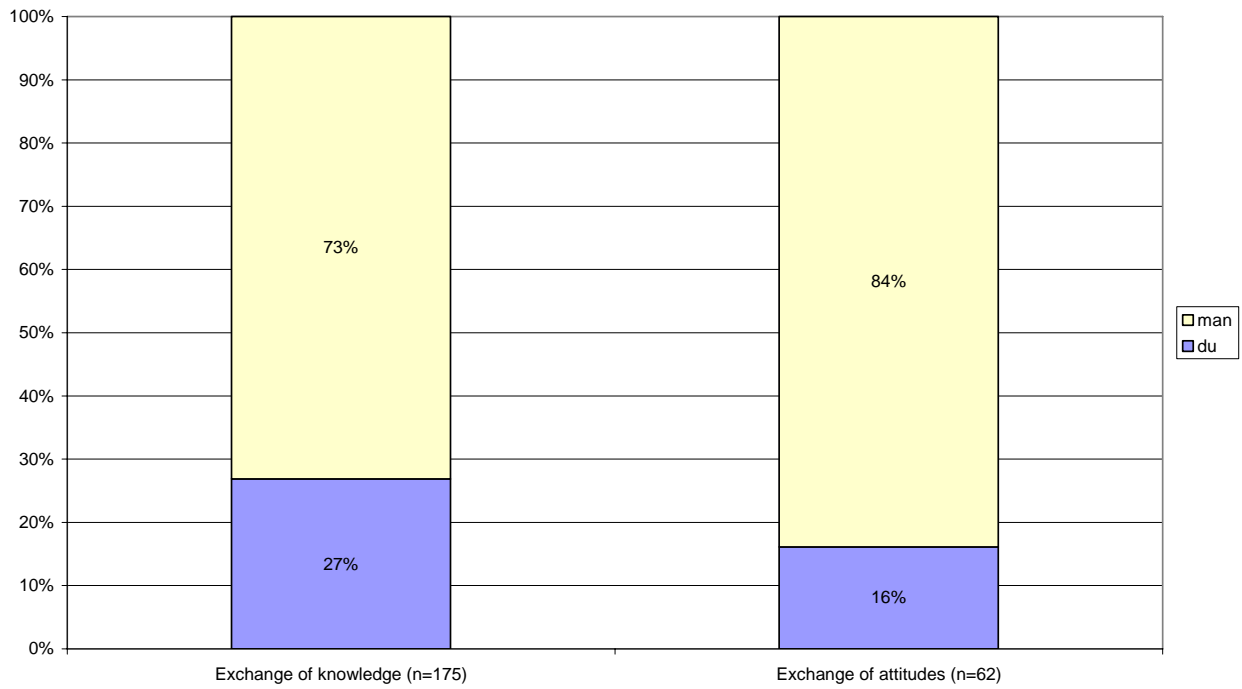
Almost all the generic pronouns uttered by TNI in the three conversations occur in passages categorized as "conversation", so it is not possible to evaluate whether there is any correlation between activity type and the use of generic pronouns. Since the purpose of this report is to compare grammatical and phonetic variation, the analysis of phonetic variation as a function of activity type is not given here, but will be presented in conjunction with the presentation of the discourse context "dimensions" (c.f. "An empirical solution to the problem of comparability").

Macro speech act

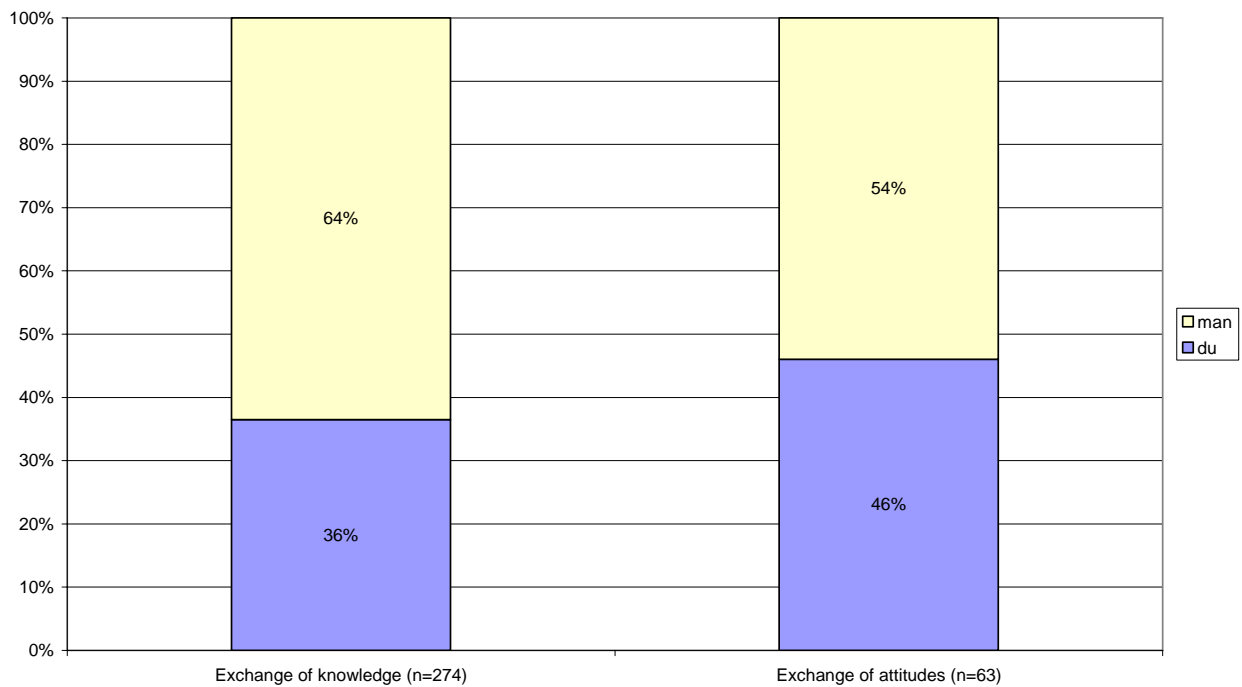
The grammatical variation

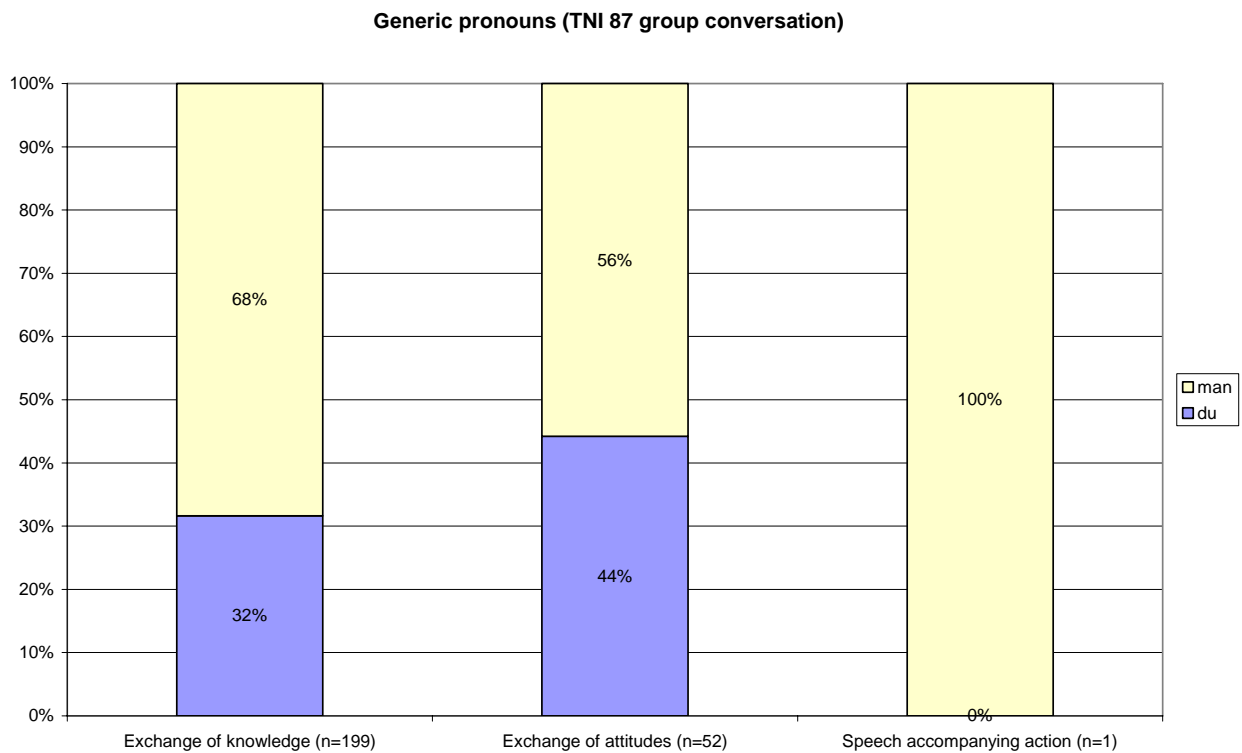
The generic pronouns in the sociolinguistic interviews almost exclusively occur in speech categorized as "exchange of knowledge" or "exchange of attitudes". There are *no statistically significant differences* between the distributions of variants in the two types of macro speech acts.

Generic pronouns (TNI 06)



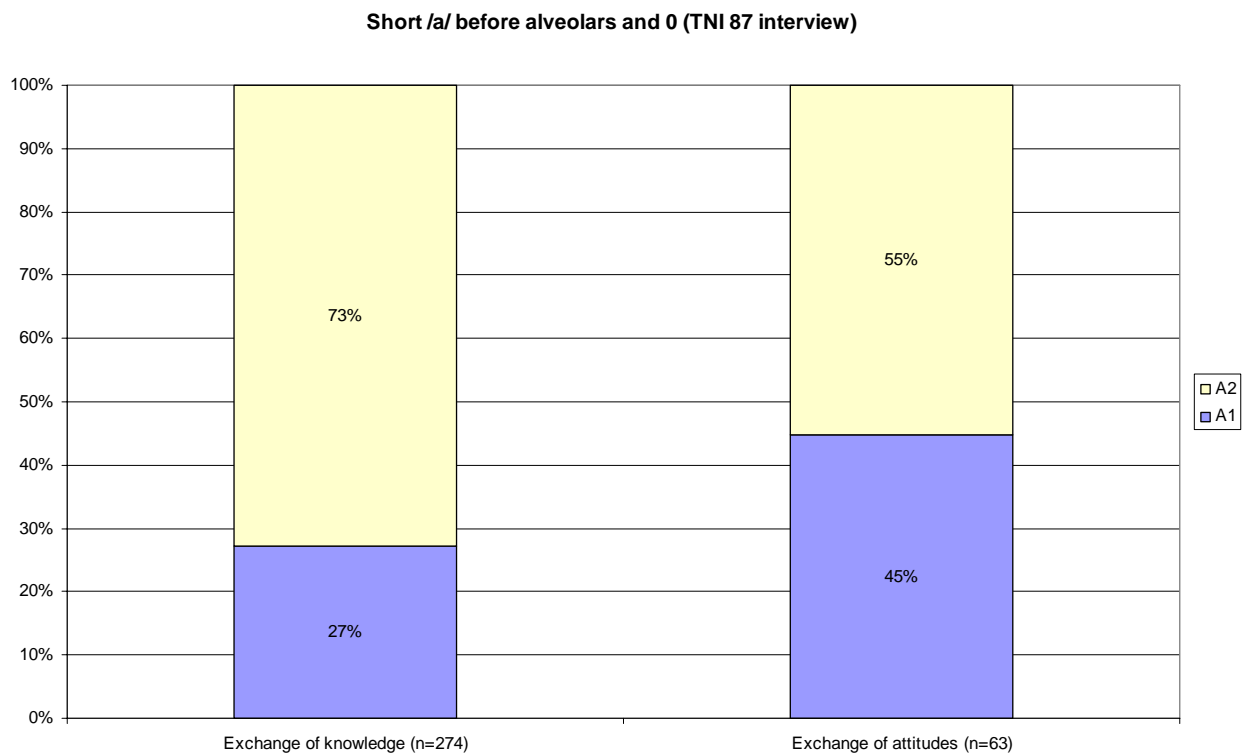
Generic pronouns (TNI 87)





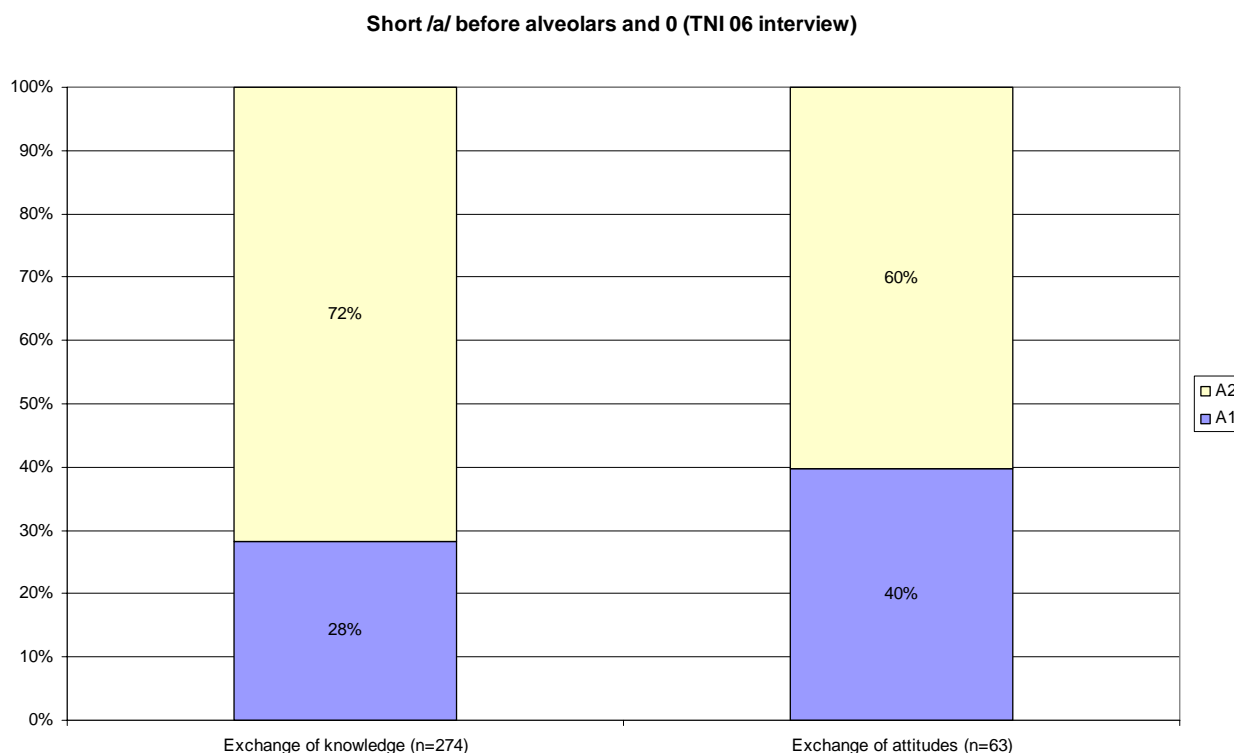
The phonetic variation

This analysis is limited to the AN variable and only the passages analysed for grammatical variation are considered here. First the original interview with TNI from 1987:



Contrary to the grammatical variable, the difference in the distribution of variants of AN with respect to the categories "Exchange of knowledge" and "Exchange of attitudes" is highly significant ($p < 0.003$) with 18 % more raised [æ]’s in passages involving the exchange of attitudes in the 1987 interview.

In the new interview from 2006 we also find a statistically significant difference in the distribution of variants of AN between the two categories ($p < 0.04$), this time with 12 % more raised [æ]’s in passages involving the exchange of attitudes.



However, for the group conversation there is no significant difference in the distributions with only a 2 % increase from passages classified as “Exchange of knowledge” to “Exchange of attitudes”, and the data are omitted here.

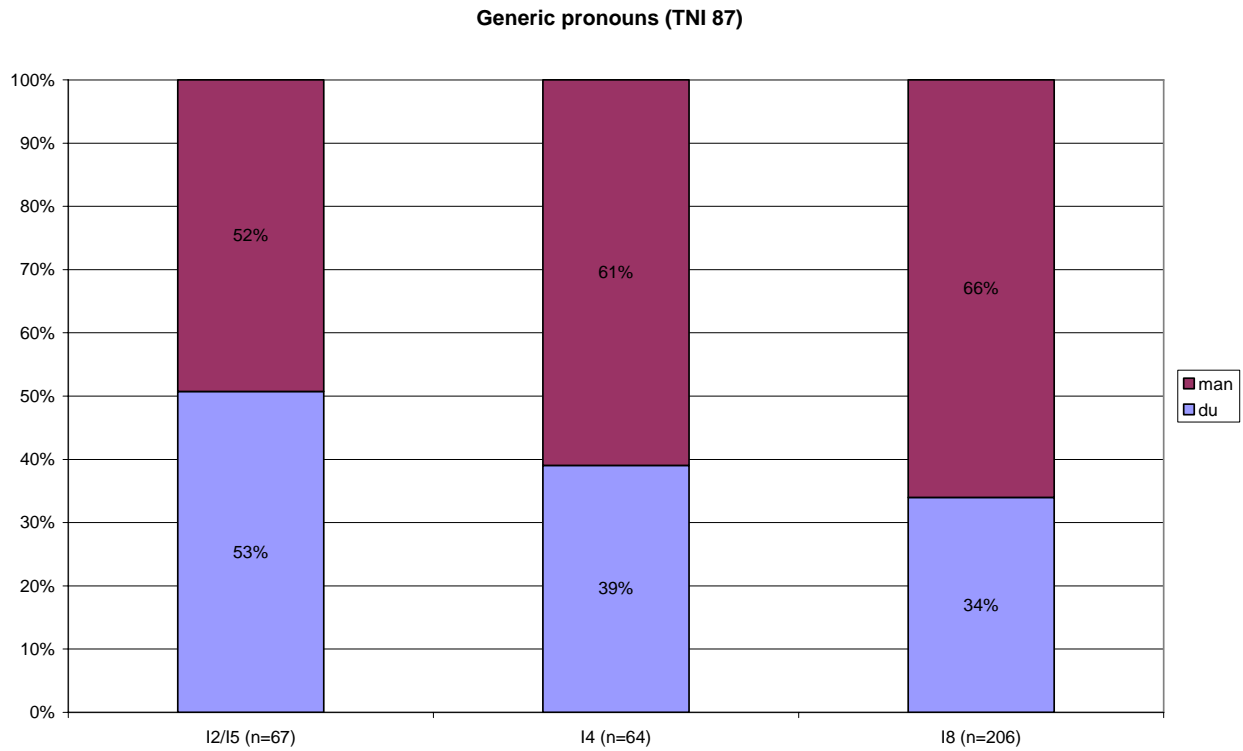
Interaction structure

The grammatical variation

Almost all the generic pronouns in the sociolinguistic interviews occur in passages categorized as “interviewer initiative with long response by informant” (I2), “no discernible structure” (I4) or “monologue” (I8). In the group conversation they most frequently occur in passages categorized as “no discernible structure” (I4) and “monologue” (I8). A chi square test is only possible if some of the categories are collapsed as the expected frequencies in too many cells in the contingency tables otherwise are below 5.

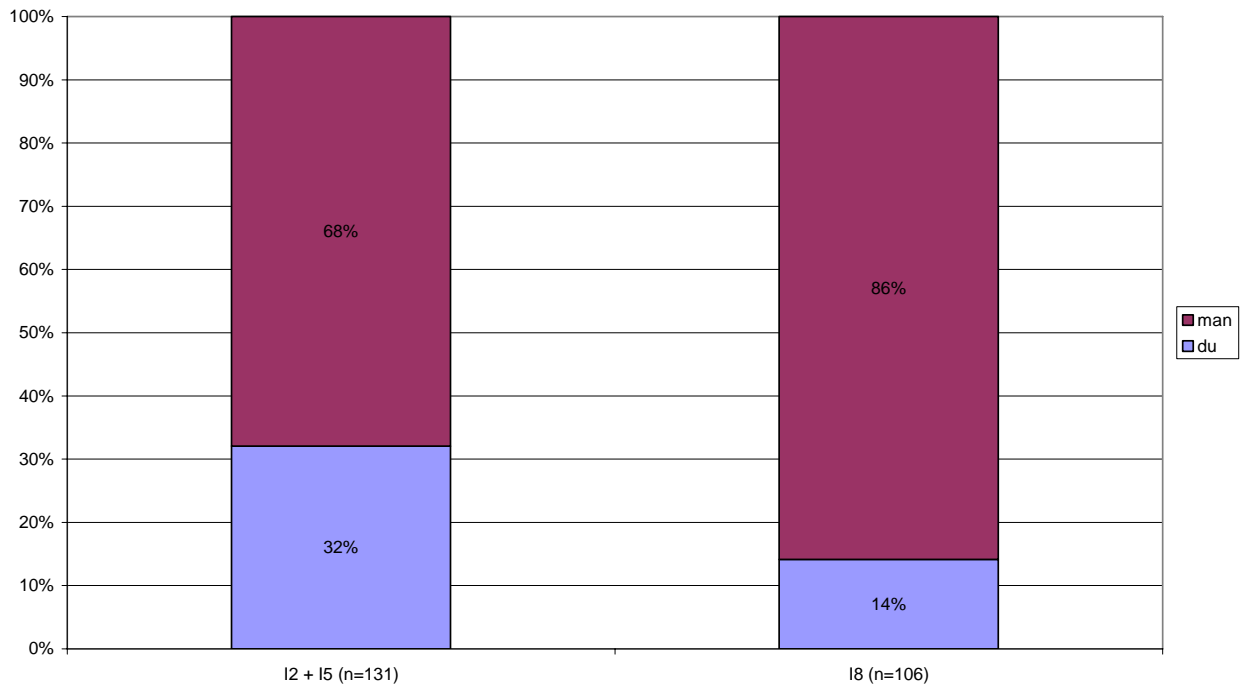
There seems to be a difference between the sociolinguistic interviews and the group conversation as regards the interaction of interaction structure and the use of generic *du*: In the sociolinguistic interviews there is a *higher* share of *du* in passages characterized by “interview structure” (one participant has the initiative, the other responds; I2, I5 and I7) than in passages dominated by one speaker (monologues; I8), whereas it is the other way around in the group conversation. The passages with no discernible structure and passages in which there is a fight for

the floor (I4 and I6, respectively) lie in the middle as regards the share of generic *du*, both in the sociolinguistic interviews (where the category is only represented in the old recording) and in the group conversation:



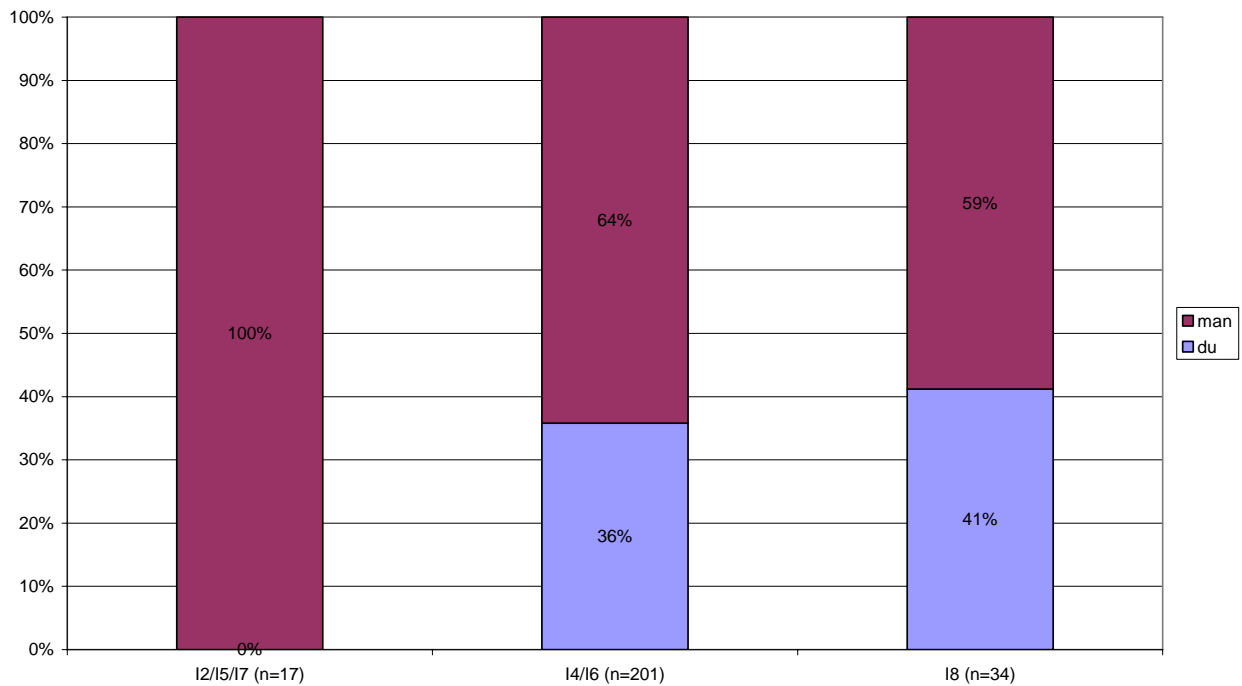
($\chi^2=6,38$; 2 d.f.; $p<0,05$)

Generic pronouns (TNI 06)



($\chi^2=9,33$; 1 d.f.; $p<0,01$)

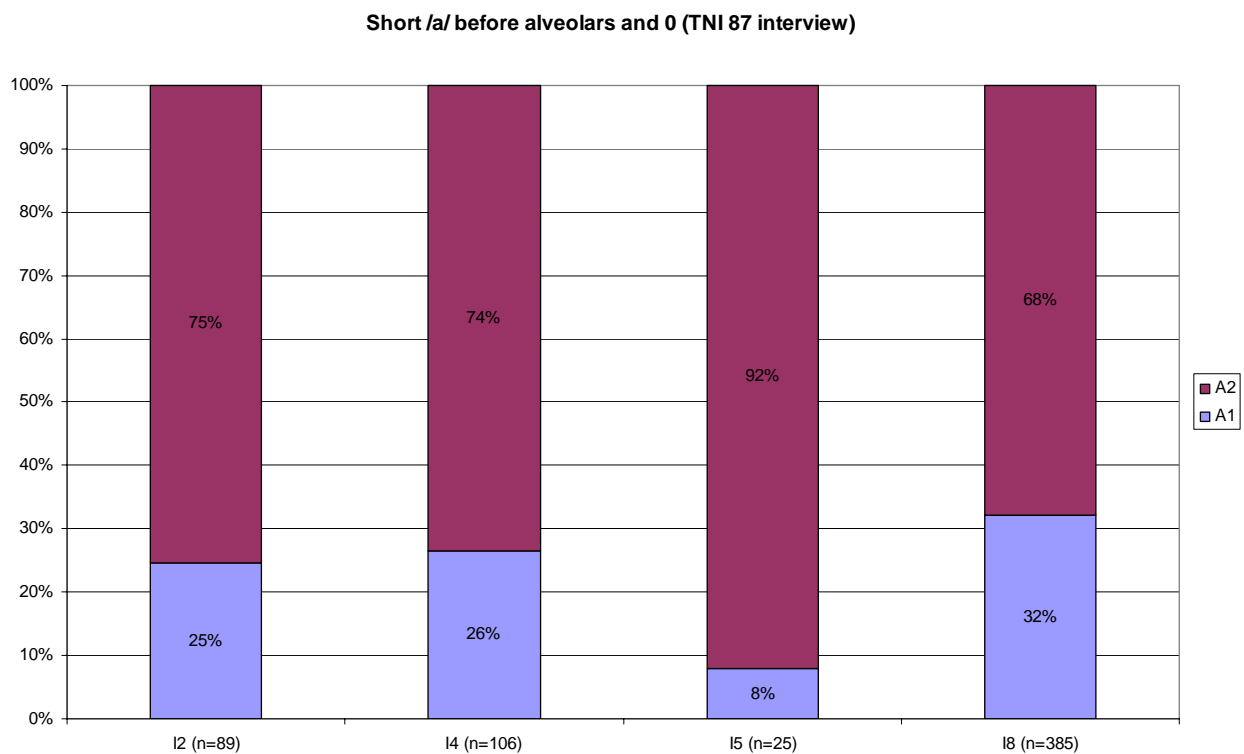
Generic pronouns (TNI 87 group conversation)



($\chi^2=9,82$; 2 d.f.; $p<0,01$)

The phonetic variation

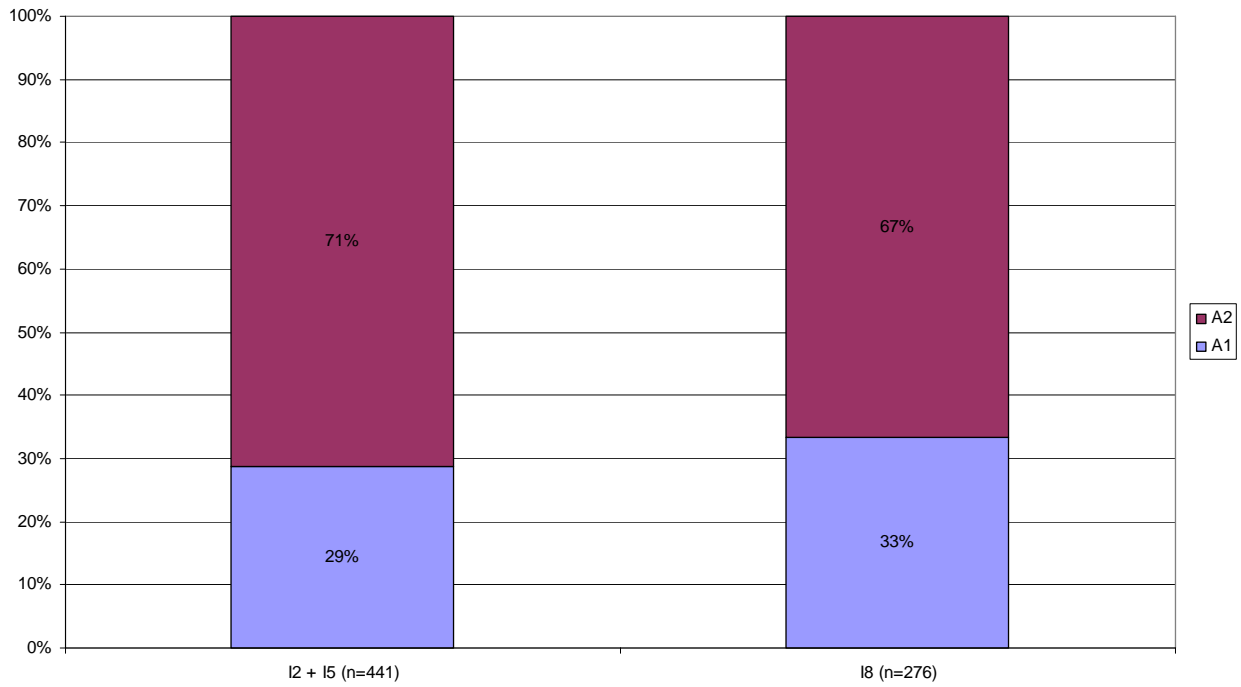
The analysis of interaction structure and phonetic variation also shows connections between the categories and the variation in the pronunciation of short /a/ before alveolars and syllable boundaries, but they are different from those found in the analysis of the grammatical variable. First we examine the data for interaction structures in the original interview. It is unnecessary to collapse the I2 and I5 intervals, so the picture is somewhat different from that seen in the analysis of the grammatical variable.



The distribution is inhomogeneous with a significantly lower share of raised [æ] in “I5-passages”.

Next we look at the new interview:

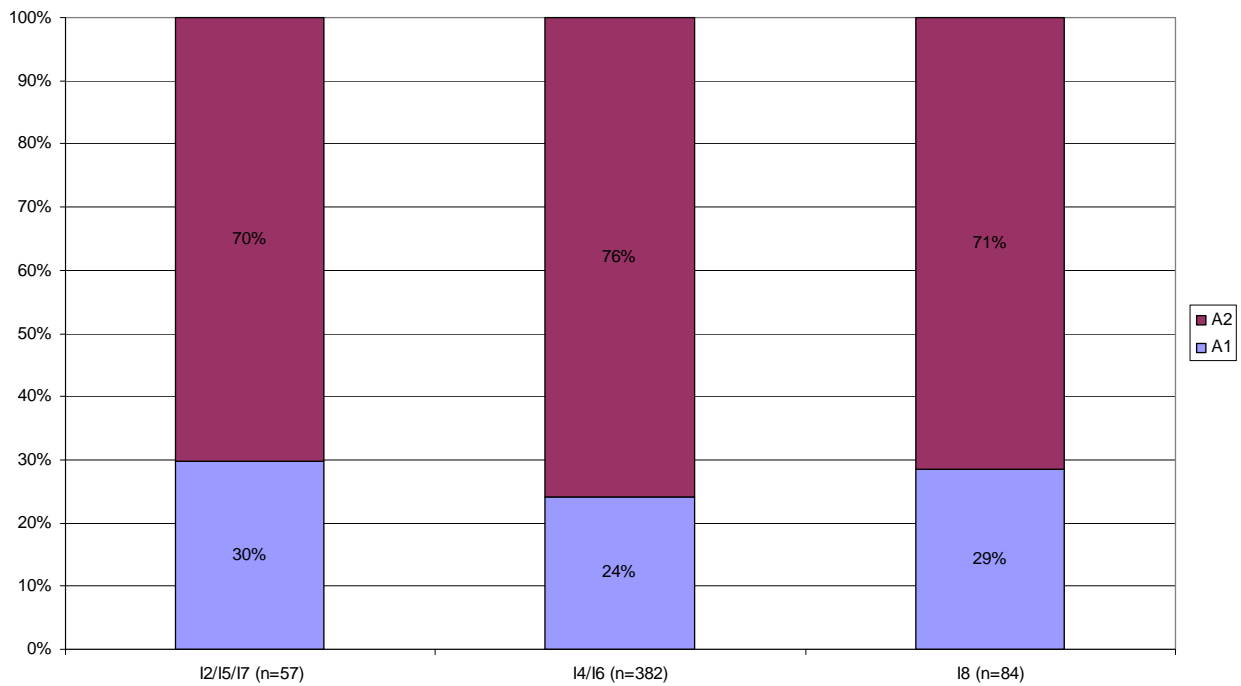
Short /a/ before alveolars and 0 (TNI 06)



Obviously this difference of 4 % is not significant ($p > 0.2$).

Finally, the group conversation:

Short /a/ before alveolars and 0 (TNI 87 group conversation)



For the AN variable as well the intervals must be collapsed and have been grouped according to the same principles as the ones used in the analysis of the grammatical variable. There is no statistically significant difference between the distributions in the collapsed categories ($p > 0.5$).

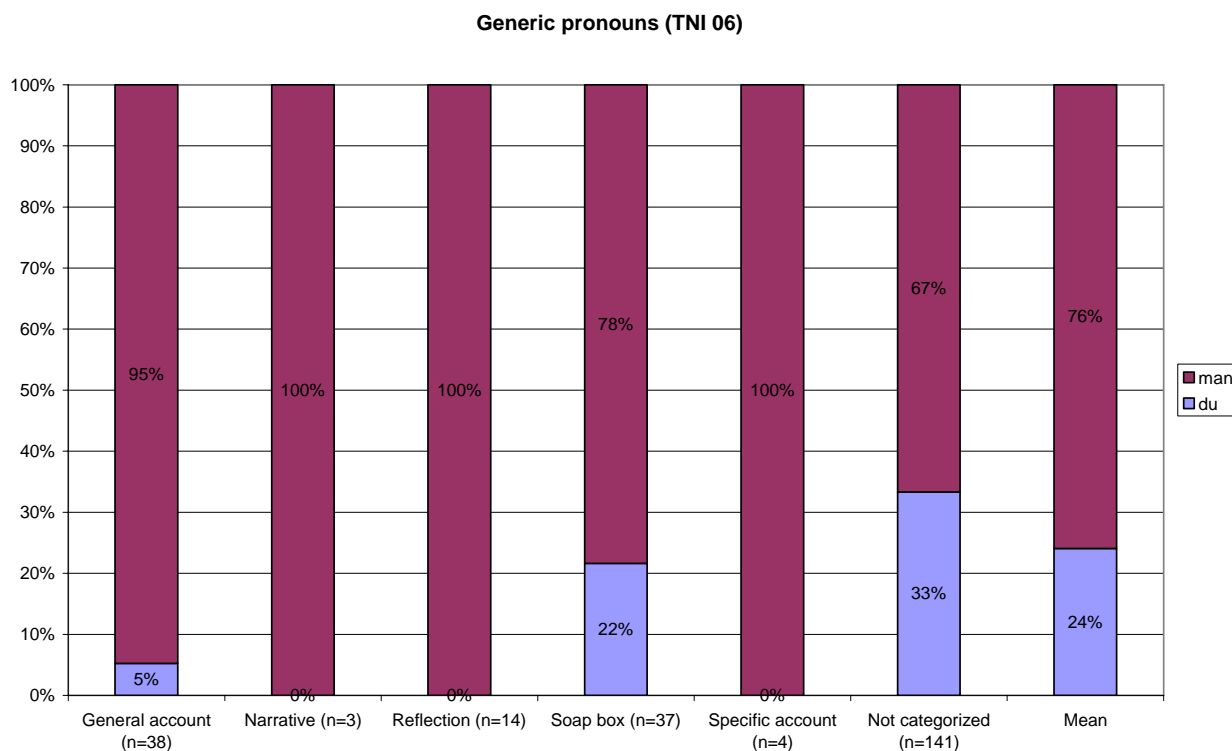
Comparison of grammatical and phonetic variables

There are only significant differences as regards the phonetic variable AN in the original sociolinguistic interview. In this conversation the grammatical variable and the AN variable show almost exactly opposite patterns: the share of *du* is highest in the I2/I5-passages and lowest in the I8 passages whereas the raised [æ] is most frequent in the I8 passages and least frequent in the I5 passages. This difference persists when I2 and I5 passages are collapsed: the share of raised [æ] in the collapsed I2/I5-category is much smaller than in the other two.

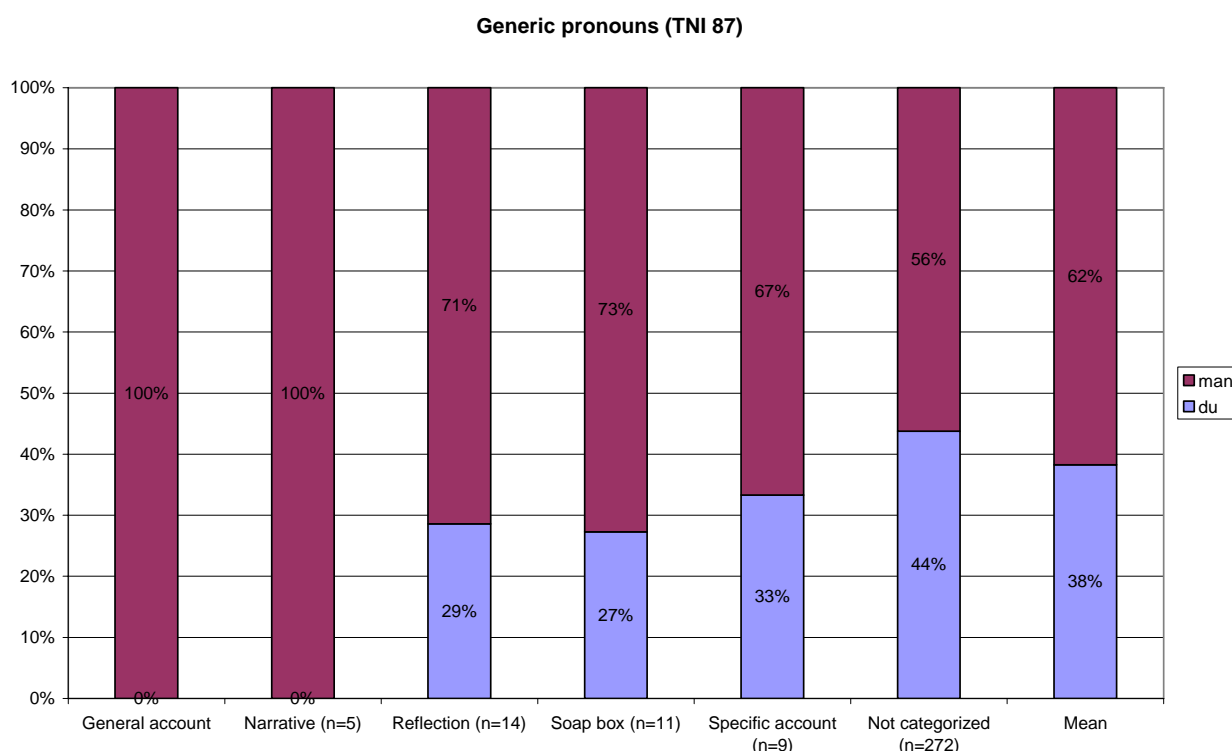
Genre

In many of the genre categories there are only very few occurrences of generic pronouns and it is not possible to perform a chi square test without conflating categories. We have therefore evaluated the tendencies statistically by comparing the most frequent genre categories with “everything else” i.e. by conflating all the other categories than the one in question.

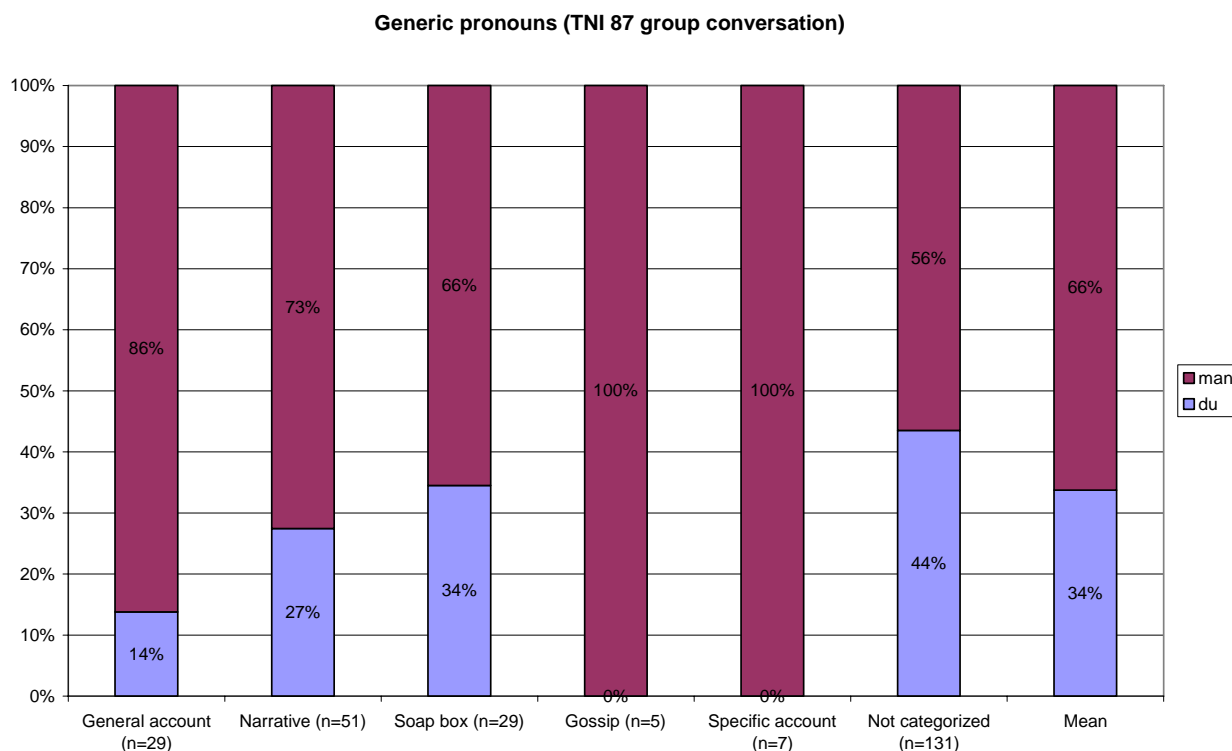
The grammatical variation



The share of generic *du* is significantly *lower* in the category ”general account” than in the rest of the conversation seen as a whole ($\chi^2=7,56$; 1 d.f.; $p<0,01$). The share of *du* is *higher* in passages not categorized into genres than in passages categorized as representing a specific genre, seen as a whole ($\chi^2=15,19$; 1 d.f.; $p<0,01$). It is not possible to evaluate the genres “narrative” and “specific account” because there are too few occurrences of generic pronouns in these passages, and there are no statistically significant differences between shares of *du* in “soap box” and “reflection” and the rest of the conversation seen as a whole.



The share of generic *du* is significantly *lower* in the category ”general account” than in the rest of the conversation seen as a whole ($X^2= 15,76$; 1 d.f.; $p<0,01$). The share of *du* is higher in passages not categorized than in passages categorized as representing a specific genre, seen as a whole ($X^2=16,68$; 1 d.f.; $p<0,01$). It is not possible to evaluate the genres “narrative”, “soap box” and “specific account” because there are too few occurrences of generic pronouns in these passages, and there is no statistically significant difference between shares of *du* in “reflection” and the rest of the conversation seen as a whole.

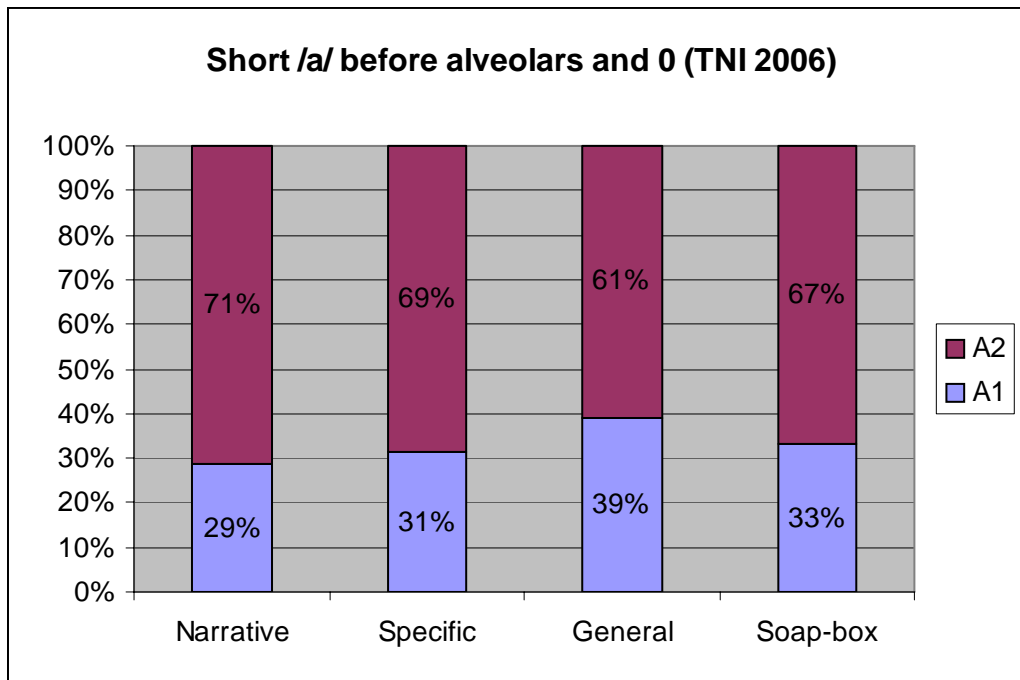


The share of generic *du* is significantly *lower* in the category ”general account” than in the rest of the conversation seen as a whole ($\chi^2= 4,86$; 1 d.f.; $p<0,05$). The share of *du* is higher in passages not categorized than in passages categorized as representing a specific genre, seen as a whole ($\chi^2=10,78$; 1 d.f.; $p<0,01$). It is not possible to evaluate the genres “gossip” and “specific account” because there are too few occurrences of generic pronouns in these passages, and there are no statistically significant differences between shares of *du* in “narrative” and “soap box” and the rest of the conversation seen as a whole.

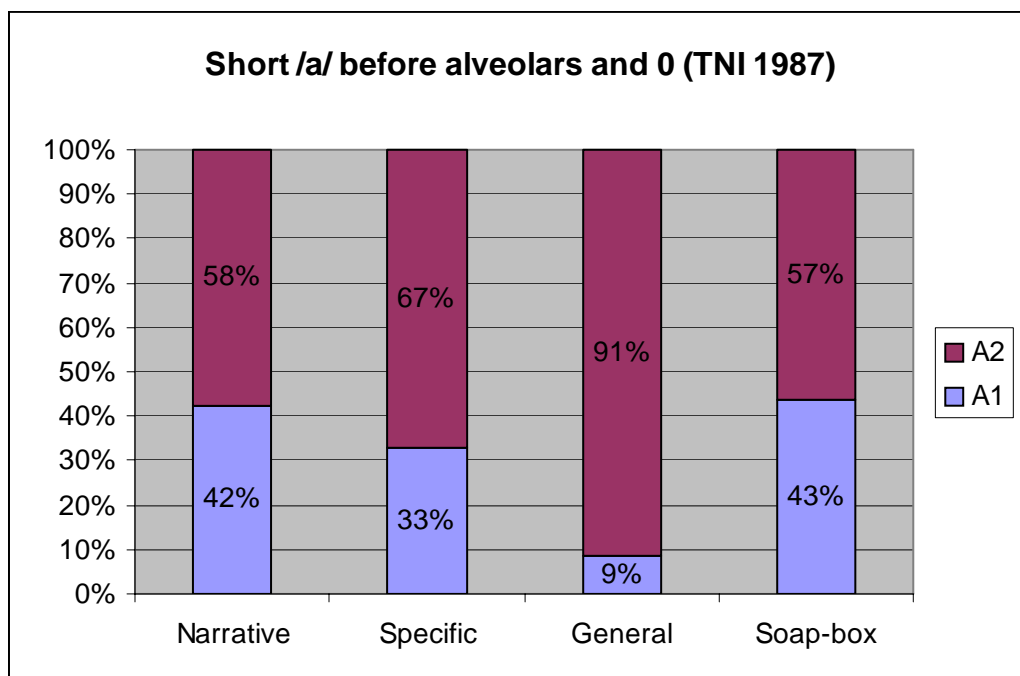
The phonetic variation

At the time of writing, the analysis of variation as a function of genre is limited to the four genres “Narrative”, “Specific account”, “General account” and “Soap-box” – no data are as yet available for the parts of the recordings which have been specified as being outside all of the four genres. In contradistinction to the grammatical variation we have a variable, AN, which is frequent enough to allow comparison of all four genres.

There are no significant differences for the AN variable between any of the 4 genre categories analysed in the new interview (in all cases $p> 0.05$). The distributions are shown below.



In the original interview the share of raised [æ] is significantly lower in passages classified as “general account” as compared to any of the 3 other genres “Narrative”, “Specific account” and “Soap-box” (in all cases $p < 0.01$), as shown in the graph below.



Comparison of grammatical and phonetic variation

For both the grammatical and phonetic variables, the category “general account” is the one that stands out from the rest in the original interview, with passages of this kind having significantly lower shares of *du* and of raised [æ]. In the new interview, the proportion of *du* is still significantly lower in “general account” passages, whereas the distribution of [æ] and [æ̥] is not significantly different from any of the other passages that have been analysed.

References

Gregersen, F. and I. L. Pedersen (1991): *The Copenhagen Study in Urban Sociolinguistics I*. C.A. Reitzel, Copenhagen, Denmark.

Gregersen, Frans & Inge Lise Pedersen (1997): Hovedsætningsordstilling i underordnede sætninger. In: *Danske Folkemål*, bind 39, Copenhagen, Institut for dansk dialektforskning.

Gregersen, Frans & Inge Lise Pedersen (2000): A la Recherche du Word Order Not Quite Perdu. In: Herring, Susan C. et al. (eds.): *Textual Parameters in Older Languages*, Amsterdam: John Benjamins Publishing Company.

Grønnum, N. (2005): *Fonetik og Fonologi – Almen og Dansk* 3rd edition, Akademisk Forlag, Copenhagen, Denmark.

Jul Nielsen, B. and M. Nyberg (1993): Talesprogsvariationen i Odder kommune II, *Danske folkemål* 35, pp. 249-348.

Kristensen, K. and J. Normann-Jørgensen (1994): *Moderne sjællandsk*. C.A. Reitzel, Copenhagen, Denmark.

Køster, F. (1996): Om æ-lukning foran nasal i sjællandsk og fynsk in Jul Nielsen, B. & I.L. Pedersen (eds.) *Studier i talesprogsvariation og sprogkontakt* (festschrift for Inger Ejkskjær), Copenhagen, Denmark, pp. 158-166.

Vikner, Sten (1999): Vo-til-Io flytning og personfleksion i alle tempora. *Íslenskt mál*, 19, 81-128.

Vikner, Sten (2001, unpublished): *Verb Movement Variation in Germanic and Optimality Theory*. vii + 269 pp. Habilitationsschrift, University of Tübingen.

Vikner, Sten (2004): Det finitte verbums position i dansk før og nu. Peter Widell & Mette Kunøe (eds.): 10. møde om udforskningen af dansk sprog. Århus.