

The long and winding road
from production data to the
psychological reality of
linguistic variation

Frans Gregersen

The DNRF LANCHART Centre
University of Copenhagen

Overview

- Chapter 1: Historical introduction
- Chapter 2: Types of variation
- Chapter 3: Three approaches to variation
- Chapter 4: Explaining variation
- Chapter 5: Coda: Theory, oh theory, and the dream of an ecumenical mundane linguistics

Chapter 1

THE HISTORY OF LINGUISTICS IN THE NORDIC COUNTRIES

Long lines

- From historicism in the past to a fragmented present
- Fragmentation breeds theoretical tolerance
- Too much so!
- The attempt here is to open an interparadigmatic discussion of VARIATION and the social characterization of speakers and speech events
- This necessitates a historical backdrop

The structuralist legacy

- Structuralism in the Nordic countries, the battle with the language historians (the neogrammarian paradigm): [...]until 1965 most Nordic linguists worked within the framework of Neogrammarian historical-comparative linguistics.” (Hovdhaugen et al. 2000: 306)
- Hjelmslev and glossematics in Denmark

From structuralism to what?

- The path from structuralism and onwards
- Two possibilities:
- Autonomy (and primarily American structuralism) as the cue: on to **TG**
- Anti-autonomy as the cue: on to **sociolinguistics** or **functionalism**

TG in the Nordic countries

- Hovdhaugen in Norway 1969
- Telemann in Sweden 1969 (diss.)
- Per Anker Jensen in Denmark 1979
- Thráinsson in Iceland 1979
- Kalevi Wiik and Auli Hakulinen in Finland
- Which country is the exception? Denmark?
Finland? Norway? Iceland?

The anti-autonomous tier

- Anti-autonomy focused on **variation**:
- **Sociolinguistics**
- Anti-autonomy focused on structure, or rather categories and paradigms, as derived from function, i.e. **semantics as the primary reason for structure**:
- **Functionalism**

The result

(I already KNOW that this is grossly simplified so please don't tell me!)

- **Formalists**
- **Functionalists**
- **Sociolinguists**
- **But what about all the others! Allow me to simplify, please.**

Data and paradigms

Paradigm	Aim	Data type	Linguistic focus
Recent theoretical linguistics	Explanatory adequacy: UG and the various possible structures of a language L	introspective data supplemented with judgmental data; primarily written	Syntax and Semantics
Functionalism	To explain language structure as a function of functional needs	No special data format; the data vary with the researcher	Grammar: Morphology and syntax on a semantic base
Sociolinguistics	Explaining Language variation and change	Observational and behavioral data; mostly spoken	Phonetic, morphological, syntactic and (rarely) semantic variation

A common problem?

- Variation is a theme in all paradigms although not in the same way
- My contention is that variation is problematic or peripheral for both theoretical linguistics and functionalism (albeit for different reasons) and central, but perhaps under-theorized, in sociolinguistics

Chapter 2

TYPES OF VARIATION

WAYF: The LANCHART study

- Replicating previous studies of spoken Danish in Vinderup, Odder, Vissenbjerg, Næstved (2 studies), Køge and Copenhagen
- Panel studies: Re-recording the same informants (and trend studies: recording informants of the same kind)

- Normcentrene
- Data-indhentningssteder



Phonetic variation

- This is central sociolinguistic stuff!
- Some central issues:
- How to define the relationship between the structural analysis and the sociolinguistic: Do they build on one another or are they based on completely different logics?
- The limits of phonetic variation
- The codification of phonetic variation: auditory vs. acoustic phonetics

Exemplification

Let me exemplify with the short (a) in Danish!

The history of (a) in Danish

1817: J.H. Bredsdorff, a man of many talents, proposes a phonetic alphabet which clearly indicates one value of (a), whether long or short, probably phonetically [ɑ]

1850: Rydquist identifies a *long* Copenhagen (a) which is more front and higher than the Swedish counterpart

Brink og Lund 1975 argue that the specific Copenhagen dialect evolves precisely during the 1850s

The splits: long and short and the internal split of the short (a)

- Around 1850 **long** (a:) becomes higher and more front
- Around 1888 this variant [a], or even [æ], is found also in the short (a) before a (j) in words like 'mig' (*me*), 'dig' (*you*), 'sig' (*(your)self*)
- Around 1930 Otto Jespersen states the rule of **complementary distribution** for the **short** (a) variants:
 - AN [æ] is found before alveolar consonants and nil
 - AM [ɑ] is found before labials and velars

The final split of the short (a)s

- 4 main variants of short (a) may be distinguished:
- AN (before alveolars and nil) is found in two variants: AN1 [ɛ] and AN2 [æ]. **There is agreement in the literature that the AN1 was, or is, a characteristic of the Working Class (WC), in particular the males**
- AM (before labials and velars) is also found in two variants: AM3 [a] and AM4 [ɑ]. **In the literature there is agreement that the AM4 was, or is, characteristic of the WC, in particular the males**

The issues exemplified

- *Is this one sociolinguistic variable or two?*
- The structuralist says one short (a) but the sociolinguist says two variables.
- *Could we then define subsets of the two?*
- We have tried to, but no, not even the nasals are the key to variation
- *How can we study it?*
- We have done both auditory and acoustic analyses

Not a social variant but an individual feature!

- **The AM4 has all but disappeared in the speech community!**
- It seems to be characteristic that very few individuals have produced an extreme proportion of the sum total of AM4s
- In the Odder data set, for example, **only two persons** have an AM4 at all. One has one instance while the other one has **13** in the S1 and **5** in the S2.
- Of all the 19 AM4s in the Odder recordings, one person has produced 18!

Conclusion: a change has happened

- Starting in the 1850s or later, but at least around 1888, the original value of the (a) was raised. The result may be seen now: AM3 is more fronted than the original quality of AM4.
- The old quality of the (a), the AM4, has disappeared as a social feature
- **AM4 is now a feature of specific individuals and thus no longer a social marker with an attached stereotype; it no longer functions as a sociolinguistic variant**

What about the AN?

- AN is still a stable, situationally very sensitive, variable but we actually do not know what it means and suspect that it may mean various things
- Conclusion: the two A-variables are sociolinguistically so different that they do not make up one category

Chapter 8

THE (RU) VARIABLE (AND ITS POSSIBLE CONNECTION TO (ENG))

The (RU) variable

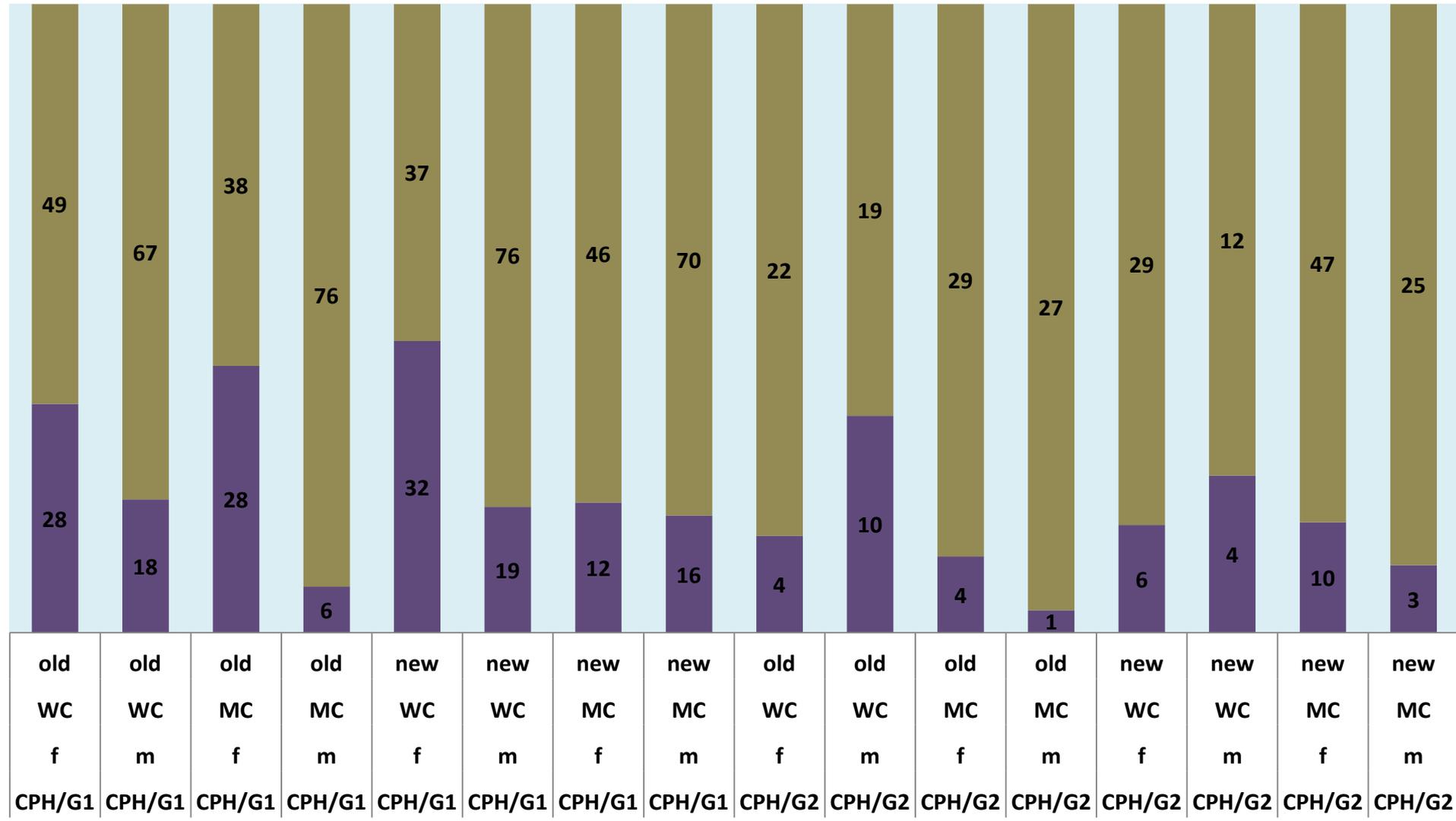
- R in Danish has in general had the effect of lowering the neighbouring vowels
- In this case we study the variable u after an r, i.e. the (RU) variable
- RU has four variants:
 - standard [u]
 - non-standard lowered [o]
 - lengthened [o:]
 - and an in between variant

Simplification

- To keep the (RU) variable simple, along the lines of the other variables we collapse the in between and the lowered non-standard variant
- We exclude the lengthened variant

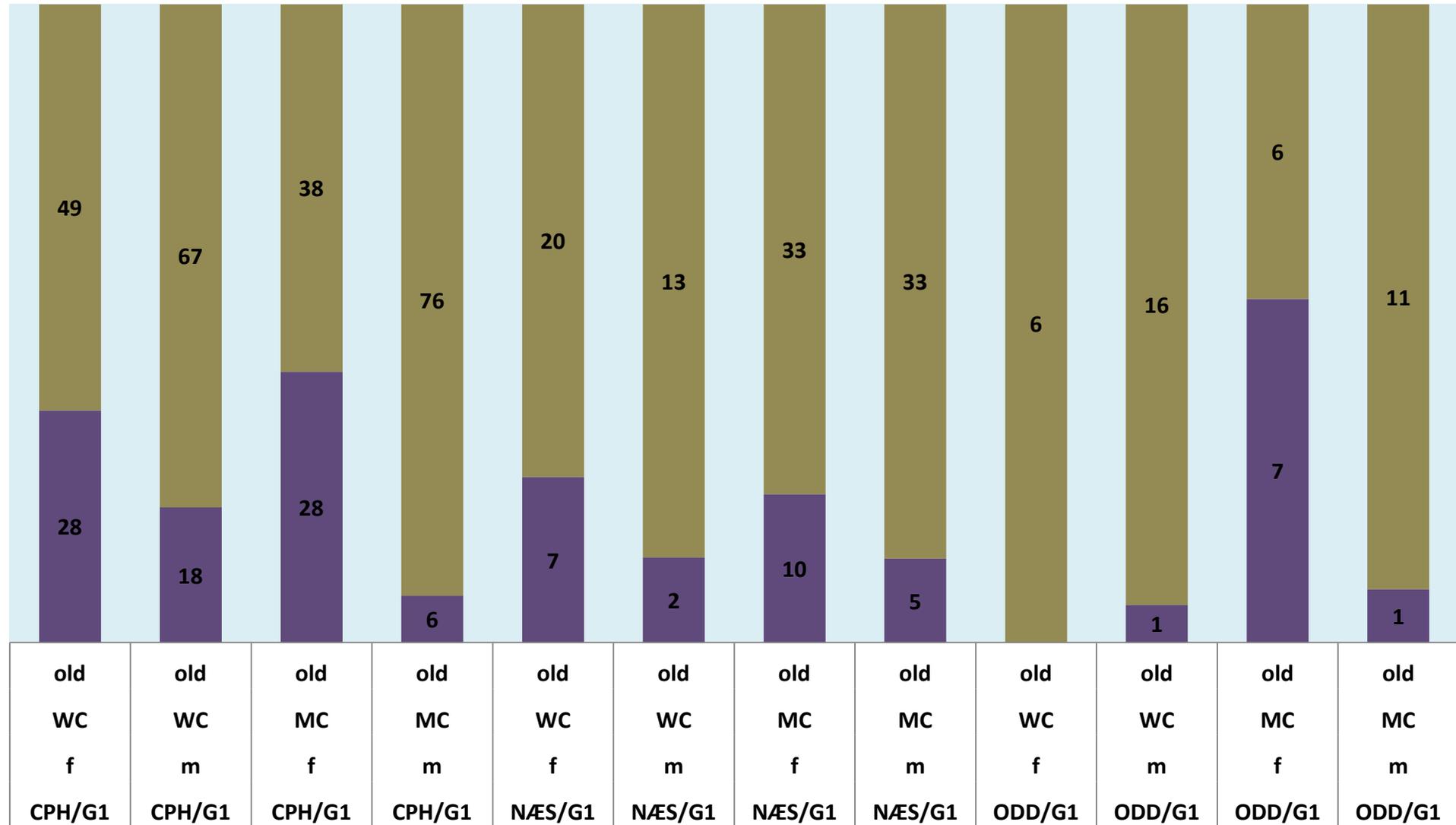
The CPH G1 and G2 in real time (RU) variable

stand.u nstand. o



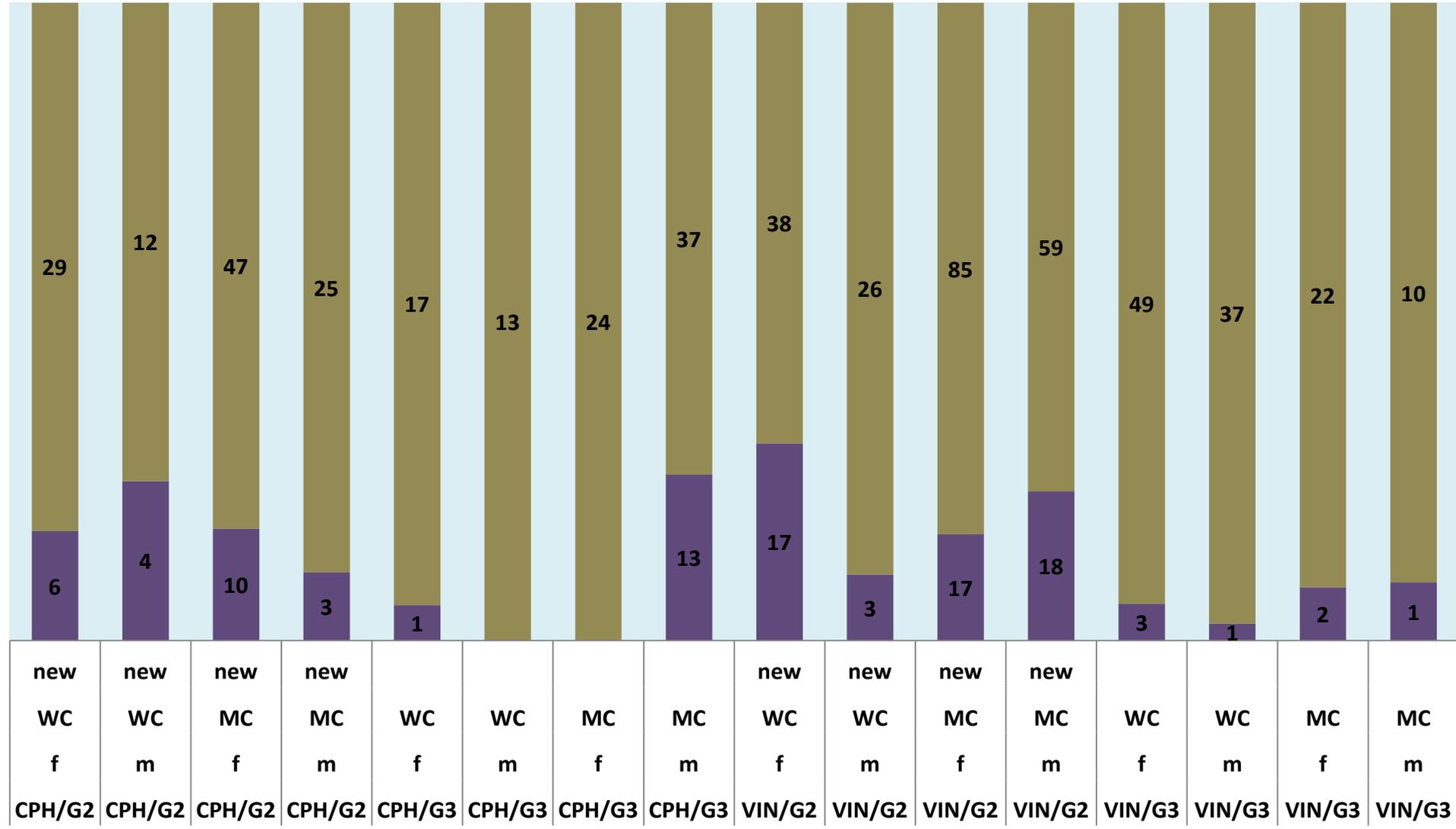
The CPH, NÆS and ODD G1 in S1; (RU) variable

stand.u nstand. o



The CPH G2 and G3 and the VIN G2 and G3 in S 2; (RU) variable

stand.u nstand. o



Conclusion as to (RU)

- The (RU) is a winner, it is close to categorical in the youngest generations, a change has happened, a non-standard variant has become the standard pronunciation
- This might be a reason to test the (RU) pattern vis a vis the (ENG): Do the patterns look the same for each site?
- Please observe: The measure is percentages of the non-standard variant (i.e. for the (ENG) the raised ones, for the (RU) the lowered ones

Chapter 9

(RU) AND (ENG)

Morphological variation

- The analysis of gender in Danish and Dutch
- Gender inside the Nominal phrase: determiners and adjectives
- Gender outside the nominal phrase: anaphoric pronouns pronouns with vague reference or higher order entities and placeholders
- Two types of over-generalization in Danish, dependent on who the informants 'are', i.e. informants with Danish as L1 or informants with something else as their L1: Is this sociolinguistically relevant? Is it Danish?

Syntactic variation

- Complementizer present or not in *at*- (that)-clauses
- The word order in dependent clauses (the V2 problem)
- Two types of passive (s-passive and periphrastic passive, Heltoft and Falster Jakobsen)
- Two types of Conditionals (with or without complementizer)

- Issue no.1: How far can we come without semantics here?
- Issue no. 2: What kinds of data?

Exemplification

Conditionals:

1. Kommer han går jeg
2. Hvis han kommer, så går jeg

From the point of view of syntax, type 1 goes with yes-no questions and other types of V1, whereas type 2 groups with adverbial clauses

From the point of view of sociolinguistics type 1 is extremely rare in spoken language but not so in written (to the extent of being a style marker) and type 2 may have V2.

Why do we have this variation?

The V2 phenomenon

- The analyses of spoken Danish show that we have much more V2 than was thought before but also importantly that in the majority of cases (around 75 %), it is impossible to tell (because there is no SA)
- In Gregersen and Pedersen 2000 we distinguish between weak signals and strong
- If V2 is indeed a signal of the embedded clause being the main point of utterance then what about the other 75 %?

Chapter 3

THREE APPROACHES TO VARIATION

A Chomskyan approach

- Biology as the framework
- Parameters as one kind of operationalization
- Parameters and usage:
 - The naturalistic theory of language must speak not only to ways in which languages are the same (principles, UG) but also to ways in which languages can differ. A descriptively and explanatorily adequate naturalistic theory of language should have the resources available to it to describe any given I-language and, to do that, it must have the theoretical resources to describe any biophysically possible I-language.

Some differences between I-languages are, however, beyond the reach of naturalistic study, people can and do differ in how they pair 'sound' information with 'meaning' in their lexicons

A functionalist approach

- **Go for differences in meaning!**
- The s-passive (morphologically marked) and the periphrastic passive (Heltoft and Falster Jakobsen 1996):
- Her synges altid en salme (general/objective) *They always sing a psalm here*
- Der blev sunget en salme og grædt (specific subjective) *(a psalm was sung and there was weeping)*
- But what about phonetic variation? Depending on your conception of meaning this can either be treated much the same way or not at all.

A sociolinguistic approach

The Labovian paradigm:

- Variation is ubiquitous
- Some of the variation has social meaning
- This meaning is connected to the social groups in the society which uses the language(s)
- Intra-individual variation: style
- Inter-individual variation: lects
- In both cases: quantitative patterns of variants belonging to a set of sociolinguistic variables

Between Schylla and Charybdis

or the devil and the deep blue sea

- The interpretation of variation:
- **SCHYLLA: voluntarism:** anybody can use any pattern of variants to express any *persona*
- **CHARYBDIS: determinism:** nobody can use any other pattern than they are positioned to do; we are at the mercy of our speaker variables
- Sailing between Schylla and Charybdis:
Identifying limits to repertoires but wide ones

Style and lect again

- In a fit of methodological masochism...
- The difference between first and second interview, the first one at work and the second at home:

Table 2: (æ) variable: results for KLT

	ε	in-between	æ	N	raising %	p (Fisher)
Interview 1	1	1	40	42	5	
Interview 2	13	1	25	38	36	0.0005***

Chapter 3

EXPLAINING VARIATION

Variation

- **Variation thrives on the asymmetry between production and reception.**
- Since our comprehension potential (or receptive competence) encompasses a lot more items than our productive potential (or competence), we are able to **decode** (or comprehend) features (or items) which we do not **use** (produce) ourselves

Language attitudes

- Tore Kristiansen has shown how linguistic items invariably are connected with stereotypes of users
- Exemplar theory, or usage based theories in general, fit this conception perfectly in that they stipulate that the brain is a huge repository of **all items heard, connected with the speech event they were used in**

Macro-variation

- I have talked about internal variation inside the same system
- But what about variation *between* systems but *within* the same community
- What kind of delimitation of the community is the relevant one: CoP; nation state; communicative network; 'speech community'?
- Which conception of community?

Variation in and between languages

- What if microsociolinguistics had to give up the notion of a language and replace it with the notion of the human linguistic pool containing all possible linguistic items
- If that happened we would be in a better position to explain the use of linguistic items connected with different 'languages' on a par with the use of variants within one and the same language

Direction

- Variation may be directional either at present or in time; this means that the variation is going somewhere (AM 4 disappearing). The s-curve.
- Or it may be stable or non-directional at present or in time; this means that the variation is there to be used. Accomodation to a moving target.

Variation in interaction

- Variation in interaction is for the audience and is always noted by the audience whether they are aware of this or not
- In this way a specific persona (or identity) is projected to the audience and noted by the auditors
- But this persona depends on our previous experience

History

- Individuals are individuals simply as a consequence of their history, we are each of us unique in this sense
- History is sedimented in the individual as layers of experience, including linguistic experience
- Thus, our various histories determine which stereotypes we (are able to) produce and interpret

Integrating Micro and Macro 1

- When an individual in an interaction within a specific domain uses a specific linguistic feature from the feature pool s/he projects an identity which evokes stereotypical attitudes
- An example: The use of features stereotypically connected with English in a domain such as the university is typically connected to internationalisation and success but it does of course depend on *which (kind of) English features*

Integrating micro and macro 2

- One way to integrate MACROSOCIOLINGUISTICS and MICROSOCIOLINGUISTICS is to focus on variation and to abolish the notion of a difference between variation between languages and within languages. This is a programme with a lot of promise - but also a lot of dangers lurking ahead

Some hidden monsters

- Which variation means something?
- The question of language ideology and language practice, consciously offered judgements and subconsciously offered ones
- The representation of variation: exemplar theory
- What processes are the representations subjected to in the mind?
- How are stereotypes used in perception of others? (Knock outs and significant features)

Chapter 5

THEORY, OH THEORY!

Dreaming back to a time...

- All histories of linguistics celebrate the era when it was possible to read all that was published in the discipline (Hovdhaugen et al. 2000: 479)
- Why?
- Because we should strive for general explanations of the general phenomenon **language in all its guises and all its uses**

Dialogues between paradigms

- A Kuhnian view of paradigms makes dialogue impossible
- But we are gathered here today...
- We seem to believe in the idea of dialogue
- How can we further it?
- By stating in a Popperian manner how a colleague, any colleague, could falsify our results!
And by laying down ground rules for the relationship between empirical work and theoretical work - if we can agree on this!

References

- F. Gregersen, J. Normann Jørgensen and J. Spindler Møller: Sideways, paper submitted to *Language Variation and Change* 2012
- Henriksen et al. (eds.): *Studies in the development of linguistics in Denmark, Finland, Iceland, Norway and Sweden*, Novus 1996
- A. Holmberg: Parameters in minimalist theory: The case of Scandinavian, *Theoretical Linguistics* 2010
- Hovdhaugen et al: *The History of Linguistics in the Nordic countries*, Academia Fennica 2000
- T. Juel Jensen og T. Karoli Christensen: Promoting the demoted, *Lingua* 2014
- F. Karlsson: Svenskans beskrivning 1-29, Falk et al. (utg): *Svenskans beskrivning* 30, 2010, 37-59.
- M. Maegaard et al: Accomodation to a moving target, *Journal of Sociolinguistics* 2013
- H. Bruun Pedersen and S. Kappelgård: Det stærkeste køn, MA thesis, UCPH 2012

acknowledgements

- Thanks to the DNRF for the grant to the LANCHART Centre 2005-2013
- Thanks to my colleagues Torben Juel Jensen, Tanya Karoli Christensen, 👍 J. Normann Jørgensen and Janus Spindler Møller and my students Hannah Bruun Pedersen and Sofie Kappelgård who have all contributed essential examples to this lecture without knowing it