

Stavelser og stavelsestab

Forskerskole Øst Torsdagsseminar 17. september 2009

Ruben Schachtenhaufen

rs.isv@cbs.dk

<http://schwa.dk>

Hvordan tæller man stavelser?

Hvis sprog nu opførte sig ordentligt, ville vi
forvente at

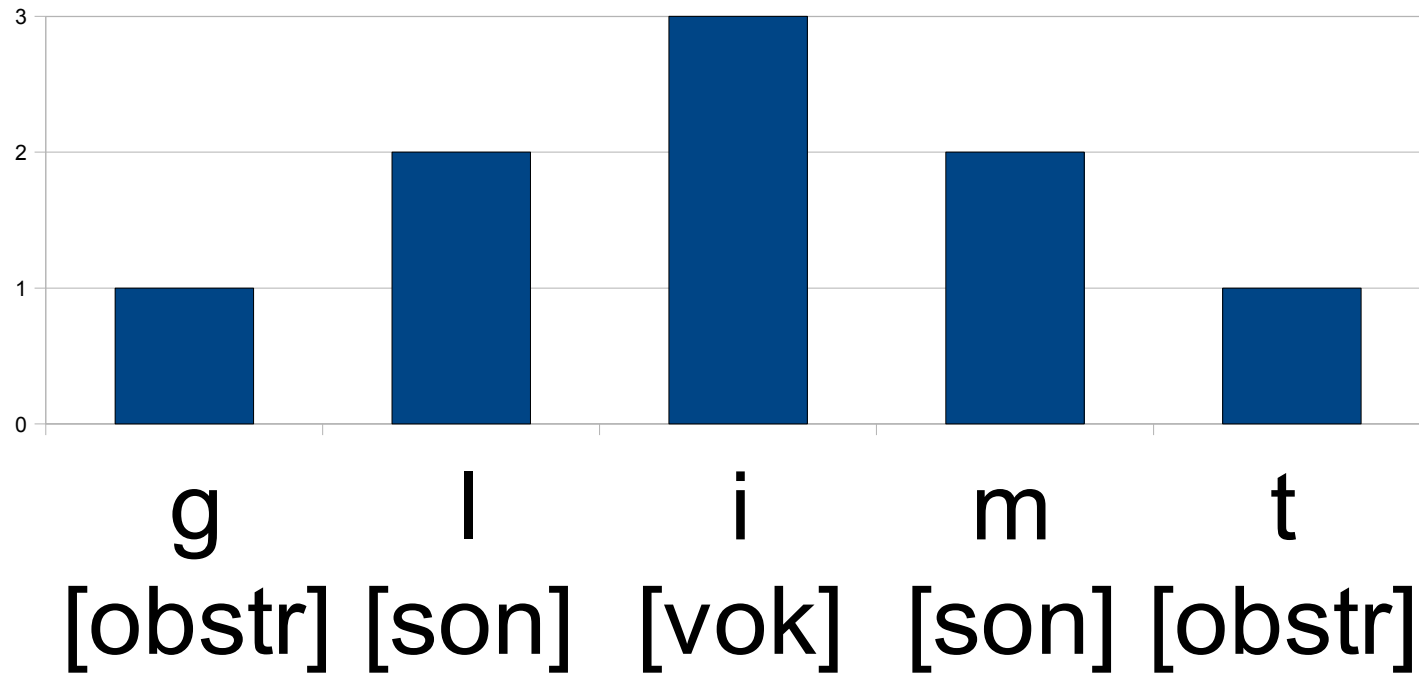
antal stavelser = antal vokaler

hverken mere eller mindre

Vokalerne ville i de fleste tilfælde være omgivet af
konsonanter

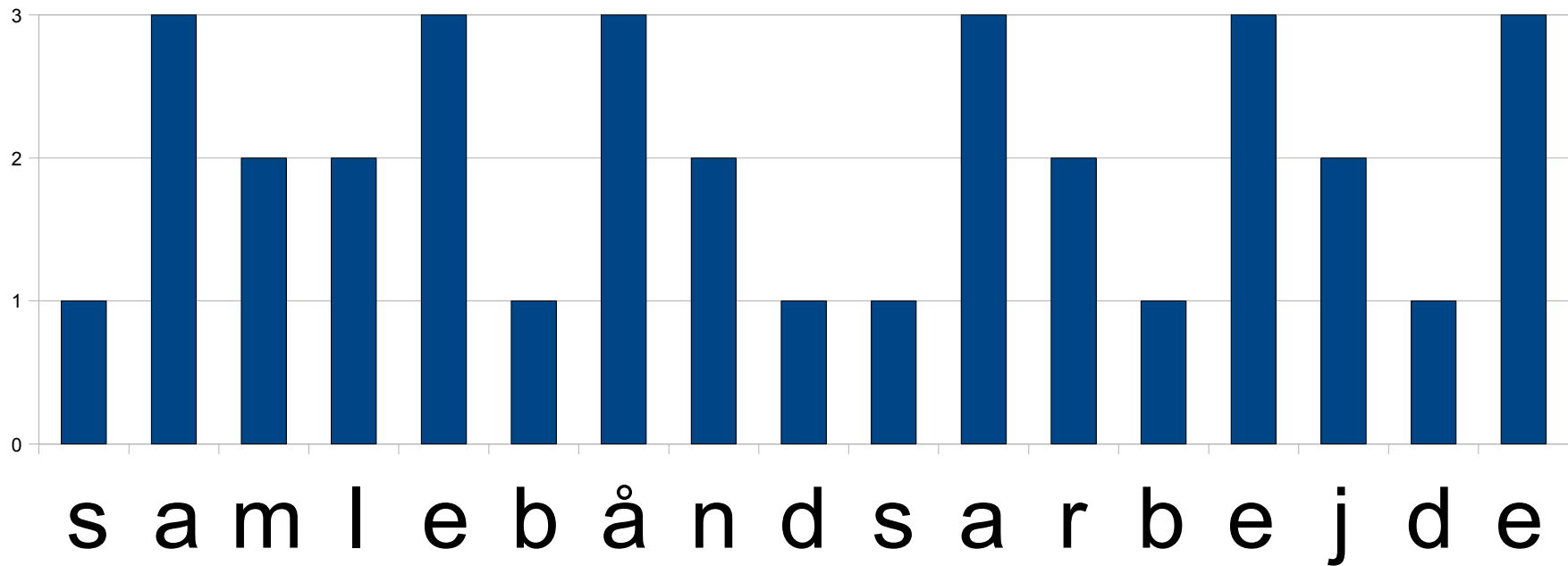
Vi ville således have en sonor kerne omgivet af
elementer af aftagende sonoritet

'glimt'
[g^hl^hem^hd]



”samlebåndsarbejde”
[ˈsamləˌbø̃nˀsaˌbø̃jˀdø̃]

6 sonoritetstoppe, 6 stavelser



Hvad er det særlige ved dansk?

Lynkursus i lydudviklingen i dansk

Lukkelyde > frikativer > approximanter (> Ø)

Ubetonet *i a u* > schwa

Schwa assimileres til sonore nabolyde
Eller falder bort

Lydudviklingen i dansk

li:i lɛ:ɛ k^hɔ:ʊ gæ:ð̥ lø:ʊ

”lige, læge, koge, gade, løbe”

p^has sɛd̥ seg̊ lø:b̥

”passe, sætte, sikke, løbe”

Sammenlign med svensk

Dansk	<i>gade</i>	[ˈgæ:ð̥]	<i>koge</i>	[ˈkʰo:ʊ]
Svensk	<i>gata</i>	[ˈgɑ:ta]	<i>koka</i>	[ˈku:ka]

Resultatet

Lange vokoide sekvenser

Mindre sonore stavelseskerner

Syllabiske kontoider

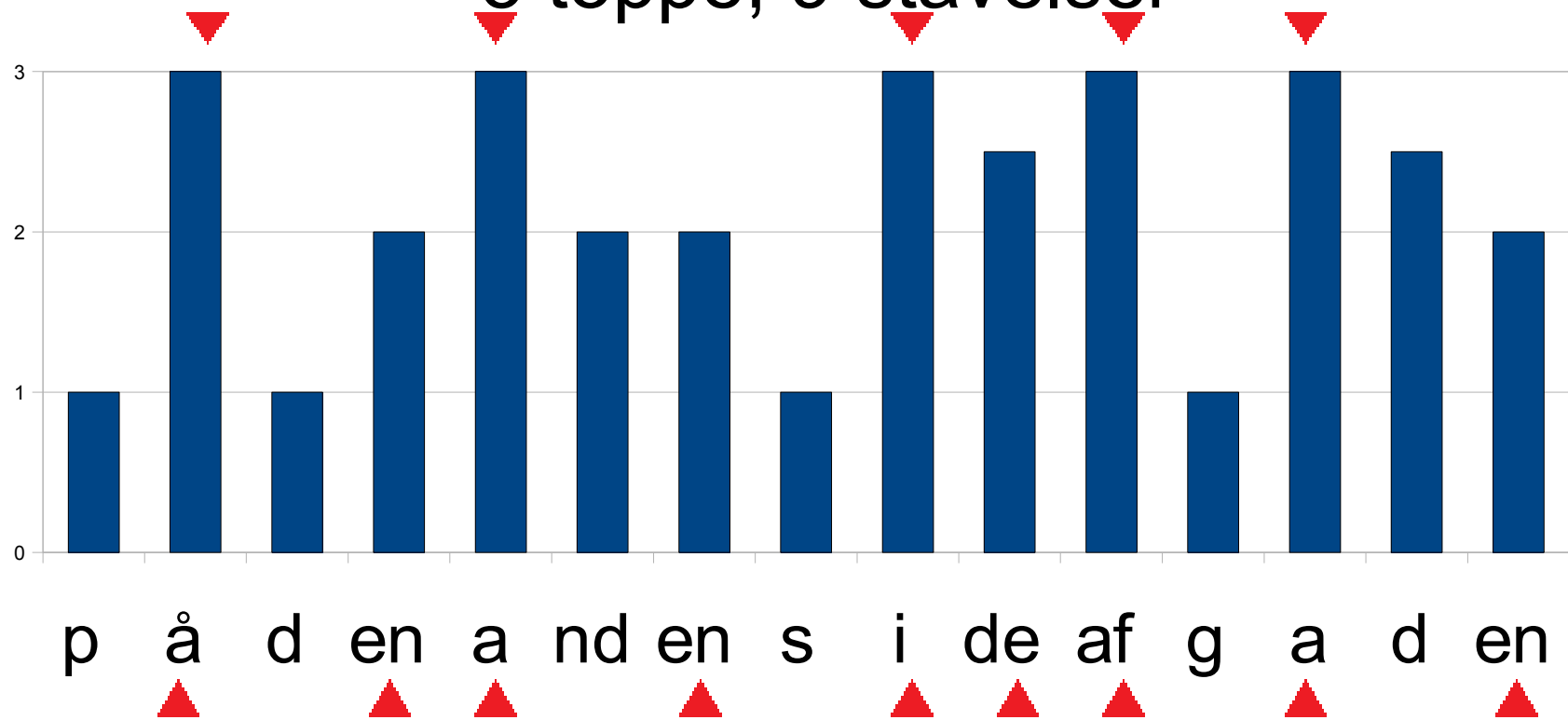
Ikke-syllabiske vokoider

Tvivi om syllabicitet

Problemer

”På den anden side af gaden”

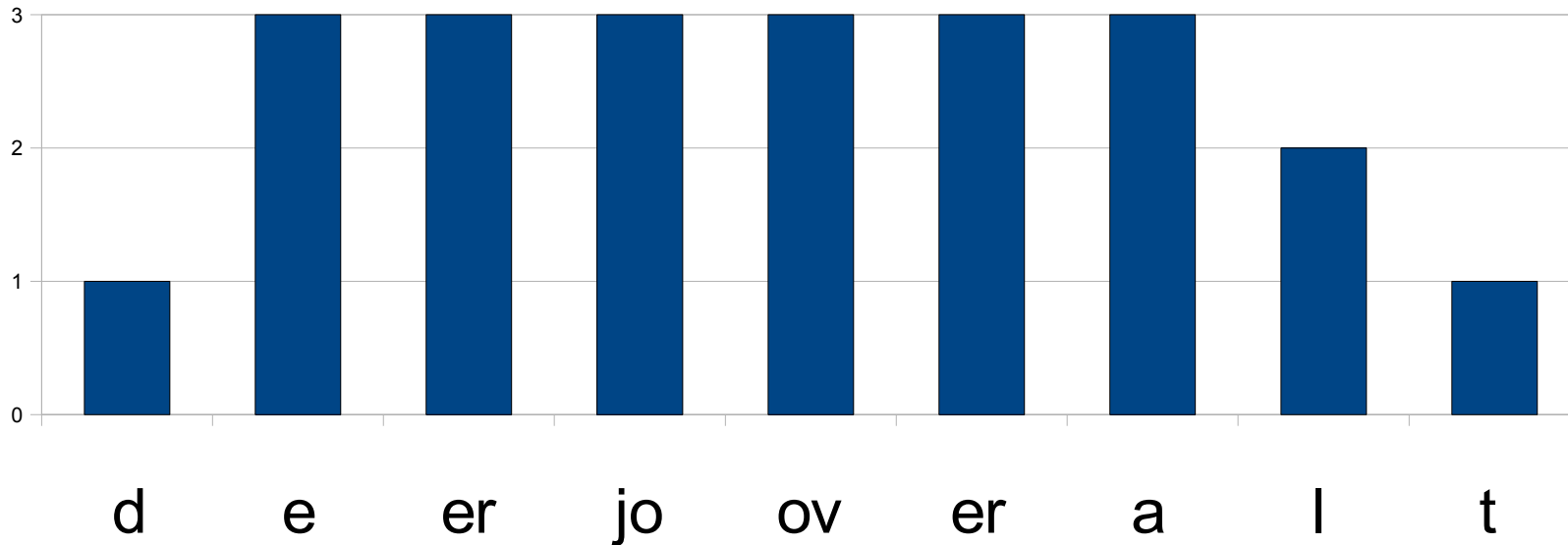
5 toppe, 9 stavelser



Problemer

”De er jo overalt”
[di 'æ o ɒb 'alʔd]

1 'top', 6 stavelser



Mange vokaler i træk

Var I i Italien?
[va i i i 't^sæ:ʔljn]

Der er arrangeret frokost kl. 12.
[da a aaŋæəð]

Det er hårdt at årelade en hest, men
det er hårdere at årelade en flodhest
['hɒ:ɒɒ ɒ 'ɒ:ɒlæ:ð]

etc.

Lange vokoide sekvenser

røget ørred

det varede året ud

få noget i øjet

det forvirrede ireren

Annotationsproblemer

Annotation af syllabicitet bliver ofte lidt arbitrær

ex. '(guld-/kul-/diamant-)minen' (from DanPASS)

mi:nn₁

mi:n₁n₂

mi:n₁

mi:n

minn₁

min₁

min

Annotationsproblemer

'lovede'

'lɔ:wəðə

'lɔ:wəð,

'lɔ:wʊð,

'lɔ:wðð,

'lɔ:wð,

'lɔ:ʊð,

'lɔ:ʊð

'lɔ:ðð,

'lɔ:cð,

'lɔ:ð,

Minimalpar

Syllabicitet *kan være* betydningsadskillende

ven – vende

sine – sigende

tordne – tordnende

bundne – bundene

cykler – cykel og

men ofte er der ingen forskel i udtalen

Stavelsestab i DanPASS-dialogerne

DanPASS

Danish Phonetically Annotated Spontaneous Speech Corpus

Map tasks

50.000 ord

22 talere

Omfattende annotation: Ortografi, part-of-speech, fonemskrift, lydskrift, tonegang, fraseintonation






Stavelsestab

75.442 fonologiske stavelser

70.501 anoterede stavelser

der sker et tab af 4.941 stavelser (6.6 %)

eksempler

-  [dsnø] *det er sådan et (ret stort område)*
-  [snɛðl] *(eller) sådan et eller (andet)*
-  [kmuð] *(når du) kommer ud af (stationen)*
-  [dsgɛg] *de skal ikke (røre hinanden)*
-  [βəg] *(du) behøver ikke (trække det endnu)*

To typer stavelsestab

vokalbortfald

[sʌdan ed] → [sned] 'sådan et'

[mo'sge:ʔ 'ʌsə] → [mo'sgʌs] 'måske også'

Vokal + vokal → diftong

[u o] → [w]

[i e] → [j]

[a] → [æ]

[sgu al[?]sɰ] → [sgwal[?]sɰ] 'skulle altså'

[gɔ i en 'bu:u] → [gɔjn 'bu:u] 'gå i en bue'

[du ha] → [duæ] 'du har'

Når vokal + vokal → diftong, opstår der nogle gange tvivl om hvilken vokal der mister syllabicitet

'diamantmine'
[diə'man[?]dmi:nə]
[dja'man[?]dmi:nə]

'biograf'
[biw'gʁɑ:ʔf]
[bjɔ'gʁɑ:ʔf]

Udfordringer for taleteknologien

Hvordan skal +/- syllabicitet repræsenteres symbolsk?

Hvordan genkendes +/- syllabicitet?

Hvordan syntetiseres +/- syllabicitet?

Akustiske cues for syllabicitet

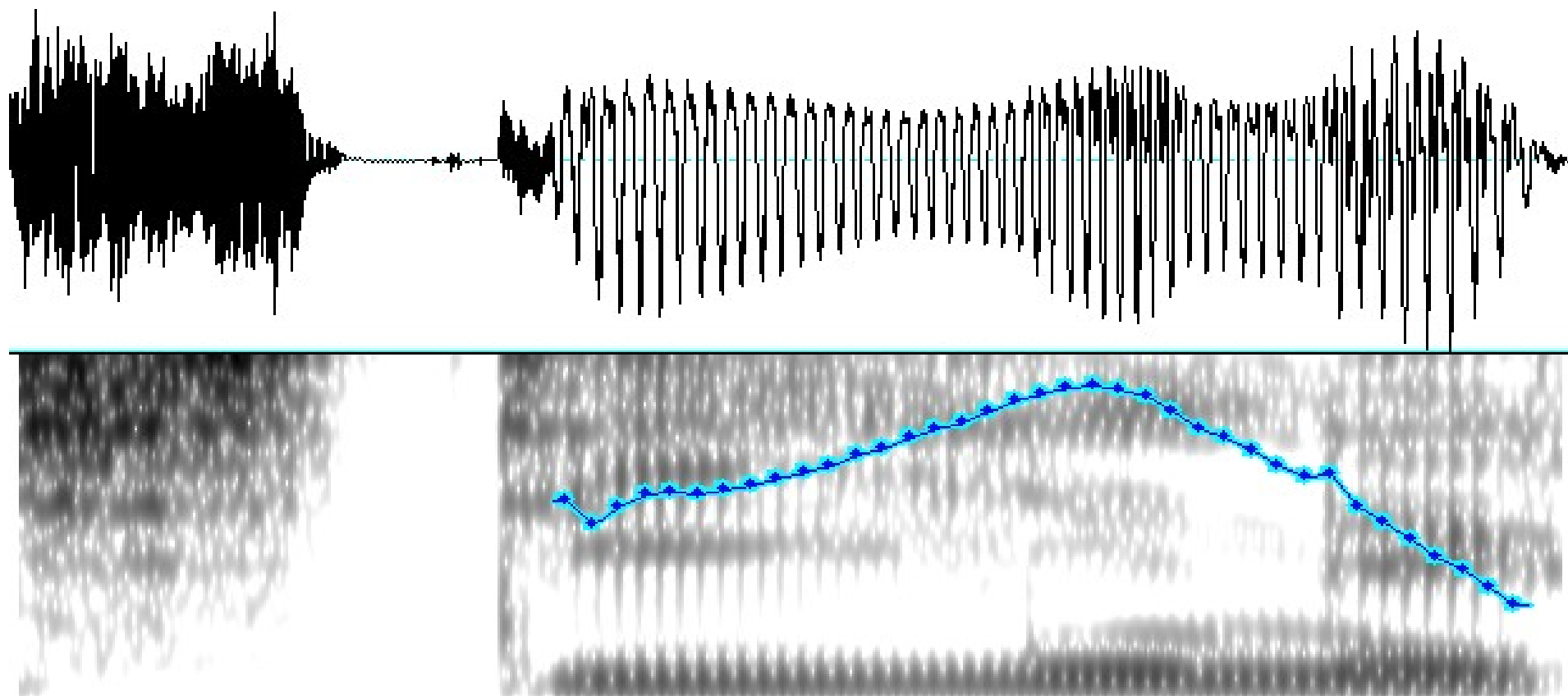
Grundtoneforløb

Intensitet

Varighed

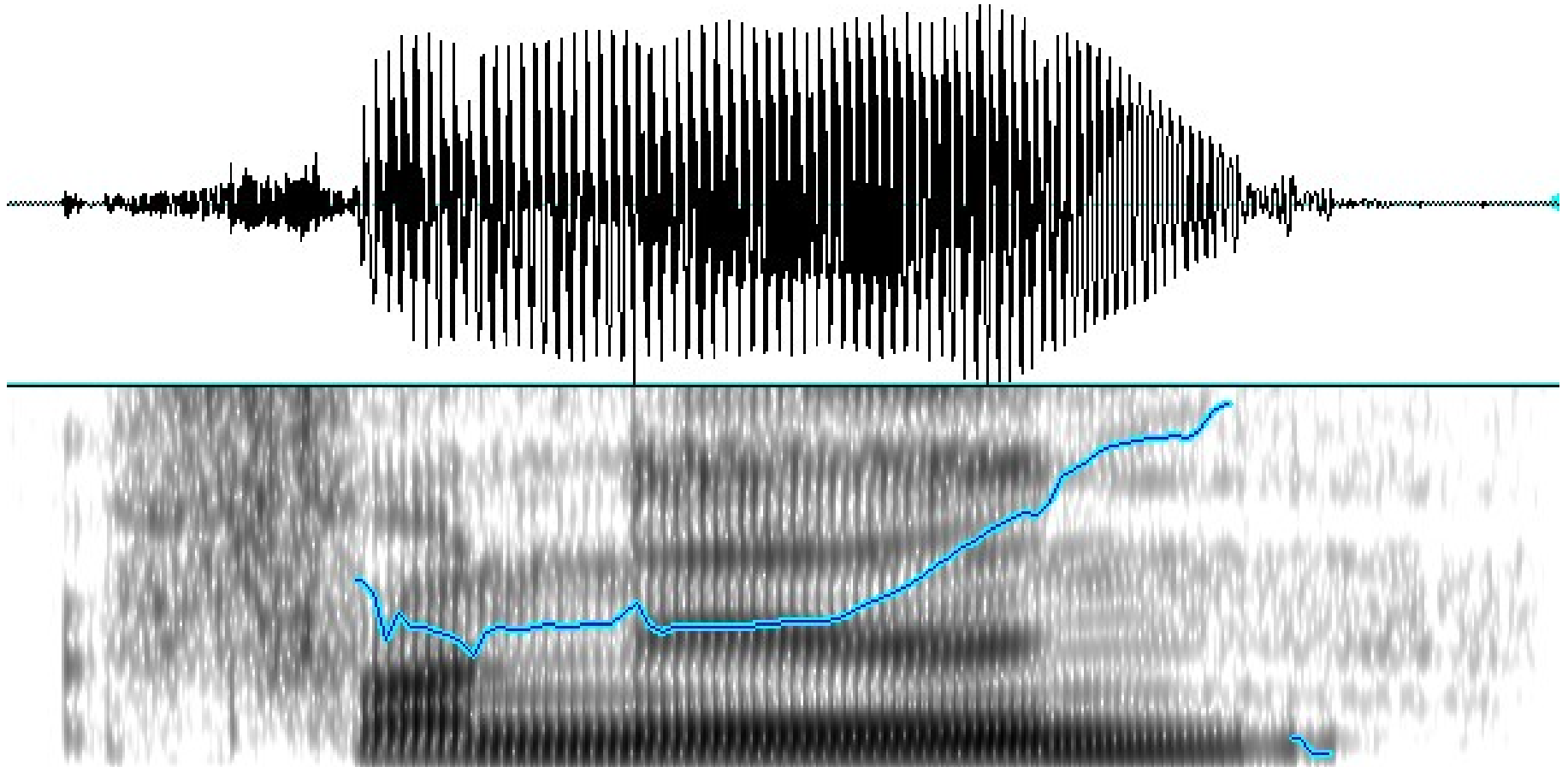
Stød

Det ideelle eksempel



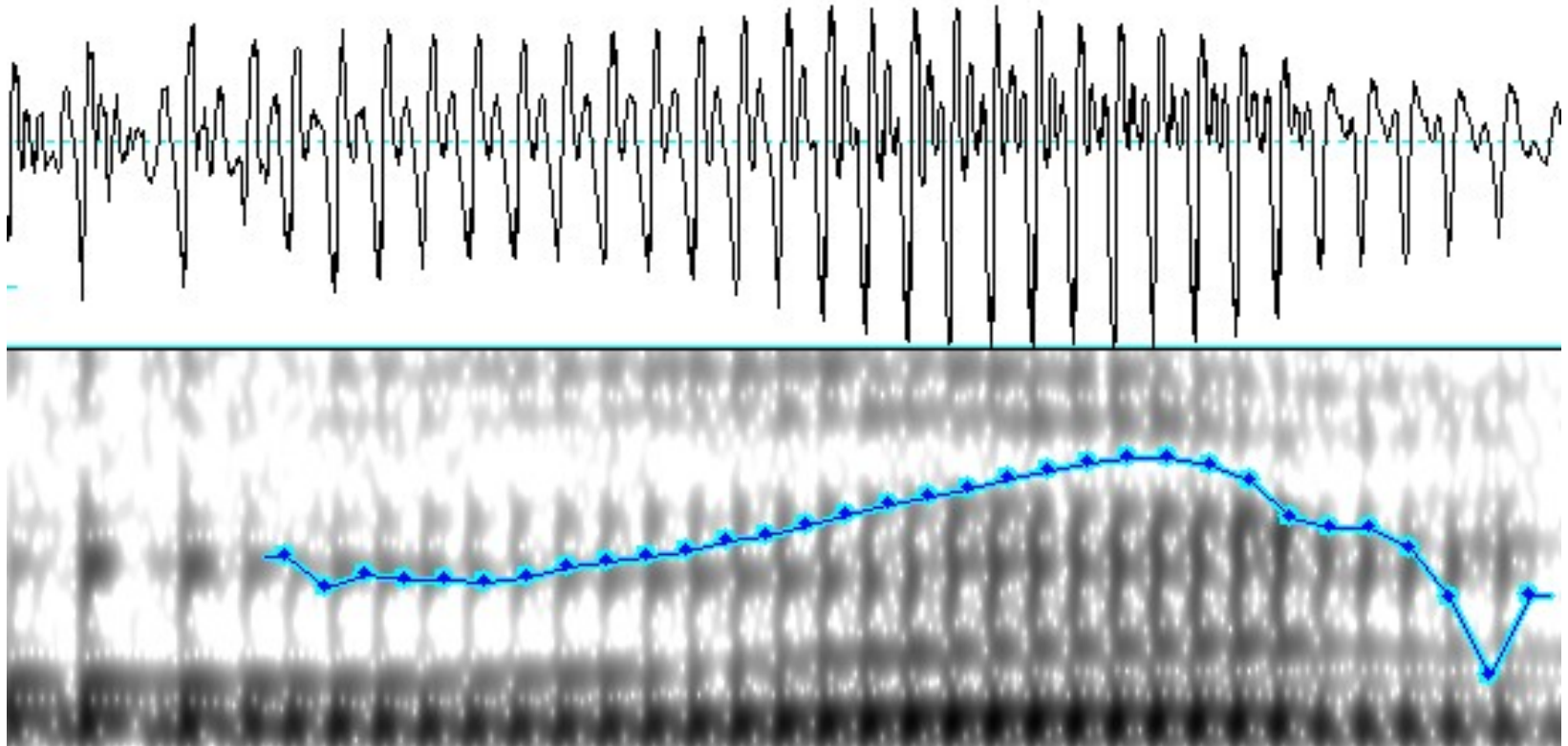
'stigende'

Cues: Grundtoneforløb



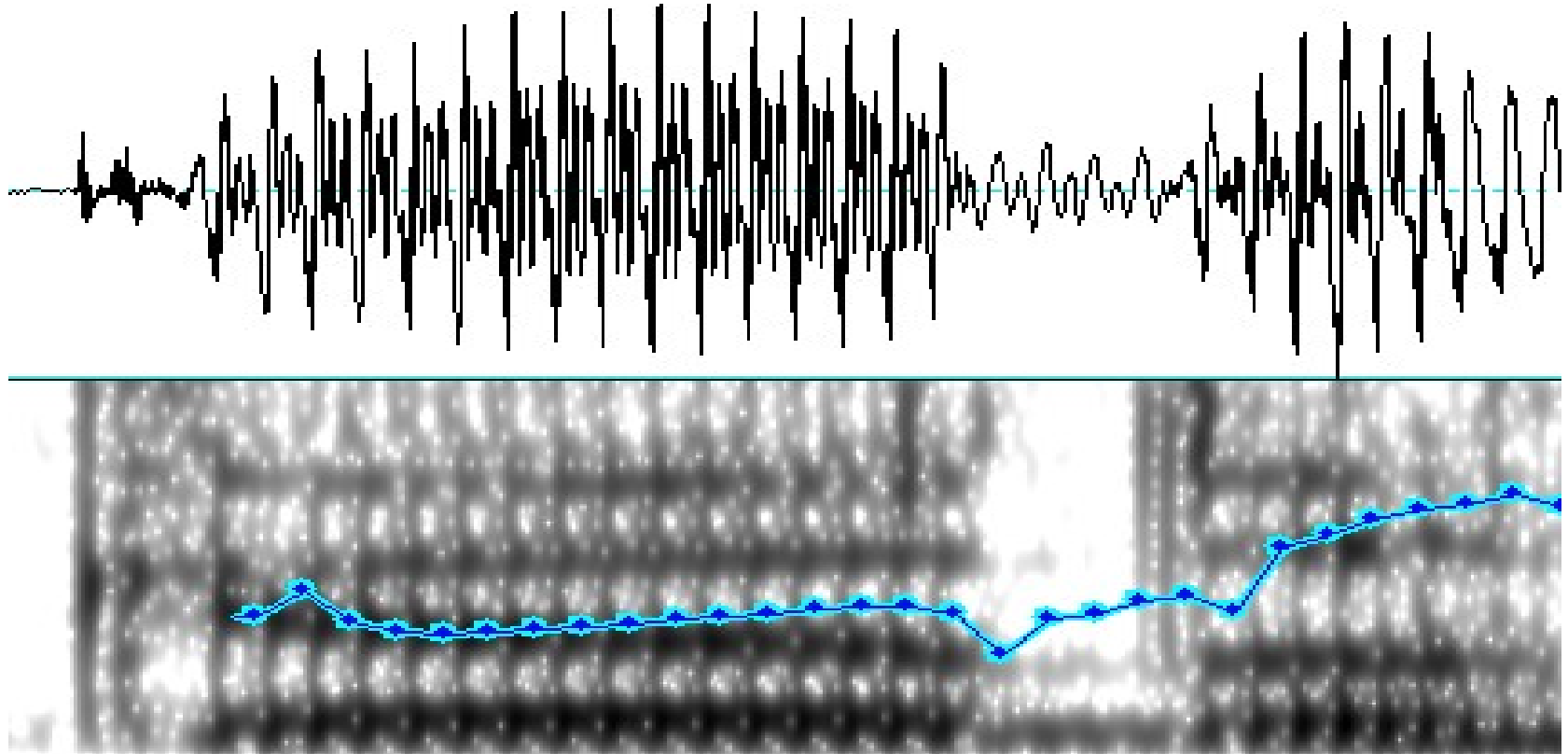
'forneden'
syllabisk n: [fɐ'ne:ðn]

Cues: Grundtoneforløb



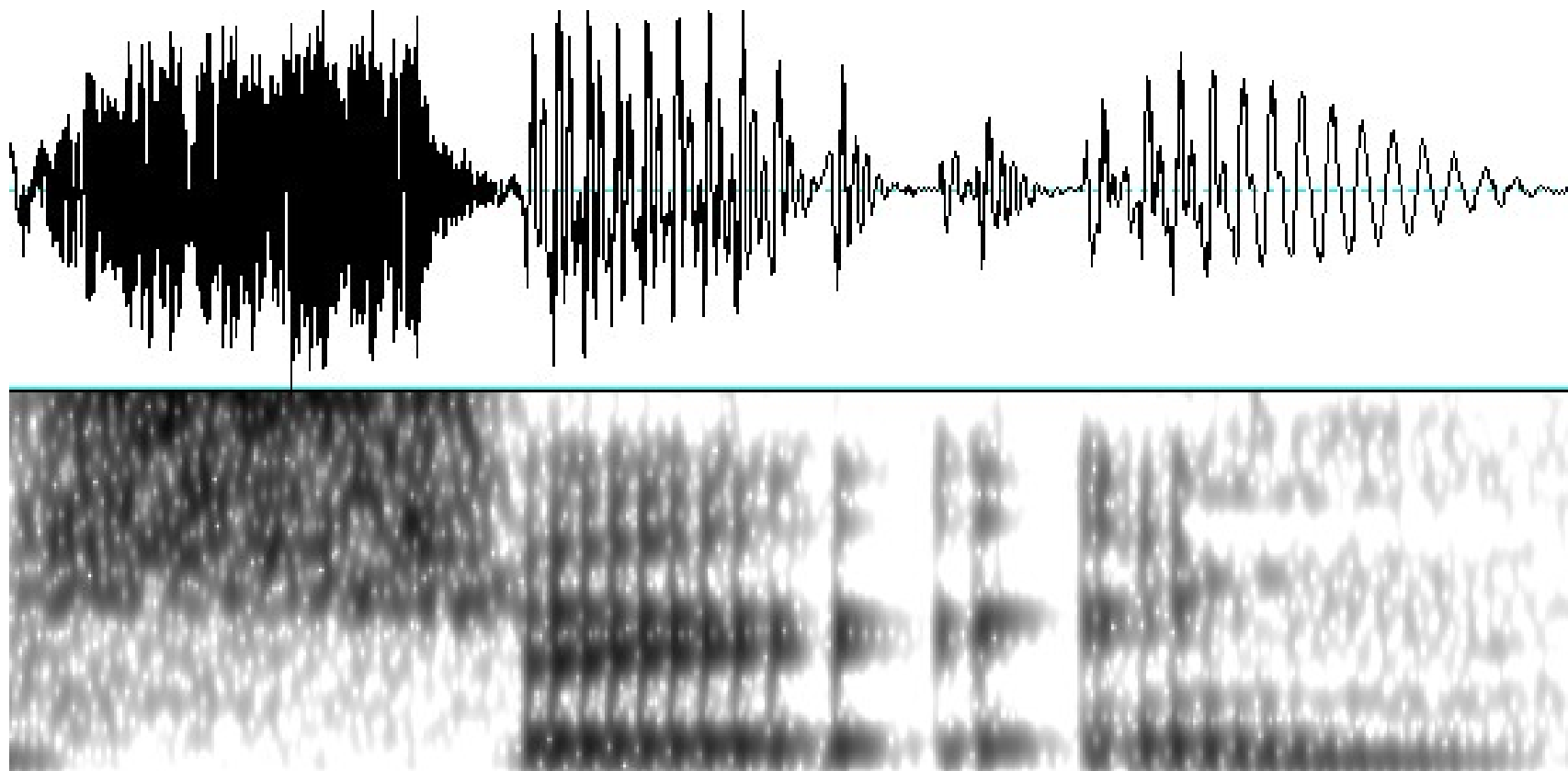
'uden'
syllabisk δ : [u: δ ŋ]

Cues: Grundtoneforløb



'gade der'
to stavelser: ['gæ:ð də]

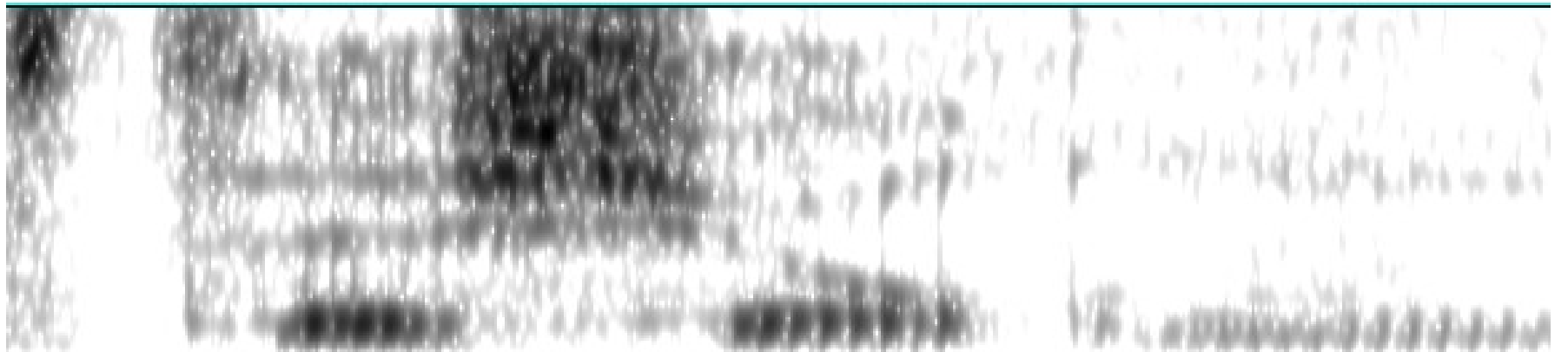
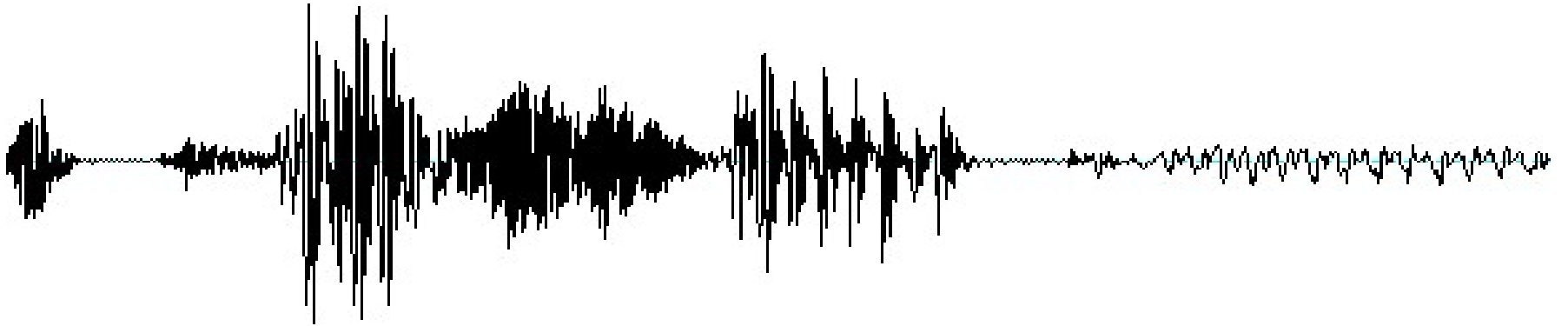
Cues: Stød



'søen'

['sø:ʔøn] eller ['sø:ʔn]?

Cues: Stød



'stationen'
[sdə'ɕo:ʔnn]

Hvilke stavelser falder bort?

Ca. $\frac{1}{4}$ af alle schwa-stavelser. Især:

- ordfinalt schwa i verber
(fortsætte, lægge, passe)
- i funktionsord
(ikke, lige, nogen)
- i proparoxytone ord
(væltede, længere, stenene)
- i sidste stamme i sammensatte ord
(børnehave, jernbane, stengærde)

Hvilke stavelser falder bort?

Der er visse fonotaktiske begrænsninger

Vi ser ikke bortfald i ord som

straks, grønt, spring, frisk

Dvs. den resulterende stavelsesstruktur efter stavelsestab skal være rimelig

Hvilke stavelser falder bort?

Bortfald sker på fraseniveau

Isoleret kan schwa ikke falde bort i *cykle* *[syg!]

Men det kan falde bort i en frase som

Du skal cykle igennem parken

Hvilke ord taber stavelser?

Top 20 (segmenterbare ord)

216/295	ikke	36/45	jernbaneoverskæring
195/452	lige	35/62	bare
163/415	sådan	30/58	gået
93/179	sige	29/40	siger
77/86	nogle	29/62	længere
68/130	noget	29/71	jamen
55/78	skulle	29/33	væltede
49/58	have	29/64	stykke
40/56	også	27/33	jernbaneoverskæringen
39/81	være	26/55	oppe

Hvilken Part-of-speech taber stavelser?

#	POS	%	DanPASS
1998	2+	44 %	9 %
853	ADV	19 %	16 %
704	V	15 %	14 %
463	N	10 %	12 %
228	ADJ	5 %	6 %
112	PRON	2 %	14 %
91	ART	2 %	3 %
58	KONJ	1 %	7 %
23	PRÆP	1 %	9 %
21	NUM	0	1 %
8	INTERJ	0	9 %
8	EGEN	0	1 %
1	UNIK	0	1 %

2+ ord

(= segmenterede intervaller som indeholder 2 eller flere ord)

4322 intervaller i alt

hvoraf 1998 intervaller taber en eller flere stavelser

Stavelsestab i ikke-segmenterbare intervaller a 2+ ord

Top 20

167	det er	19	du et
48	du er	16	det er et
41	så er	15	har ikke
40	du ikke	14	jeg ikke
35	der er	14	på et
30	det vil	13	du har
25	skal jeg	13	du en
21	jeg er	13	i en
21	sådan en	12	de er
21	har jeg ikke	11	har et

Typiske emner inden for forskning i udtalevariation

Hvordan påvirkes udtalen af

alder, geografi, køn, ordfrekvens osv.?

Men hvad med intratalervariation? Hvorfor udtaler en og samme person ord og orddele forskelligt i forskellige sammenhænge?

Reduktion og funktion

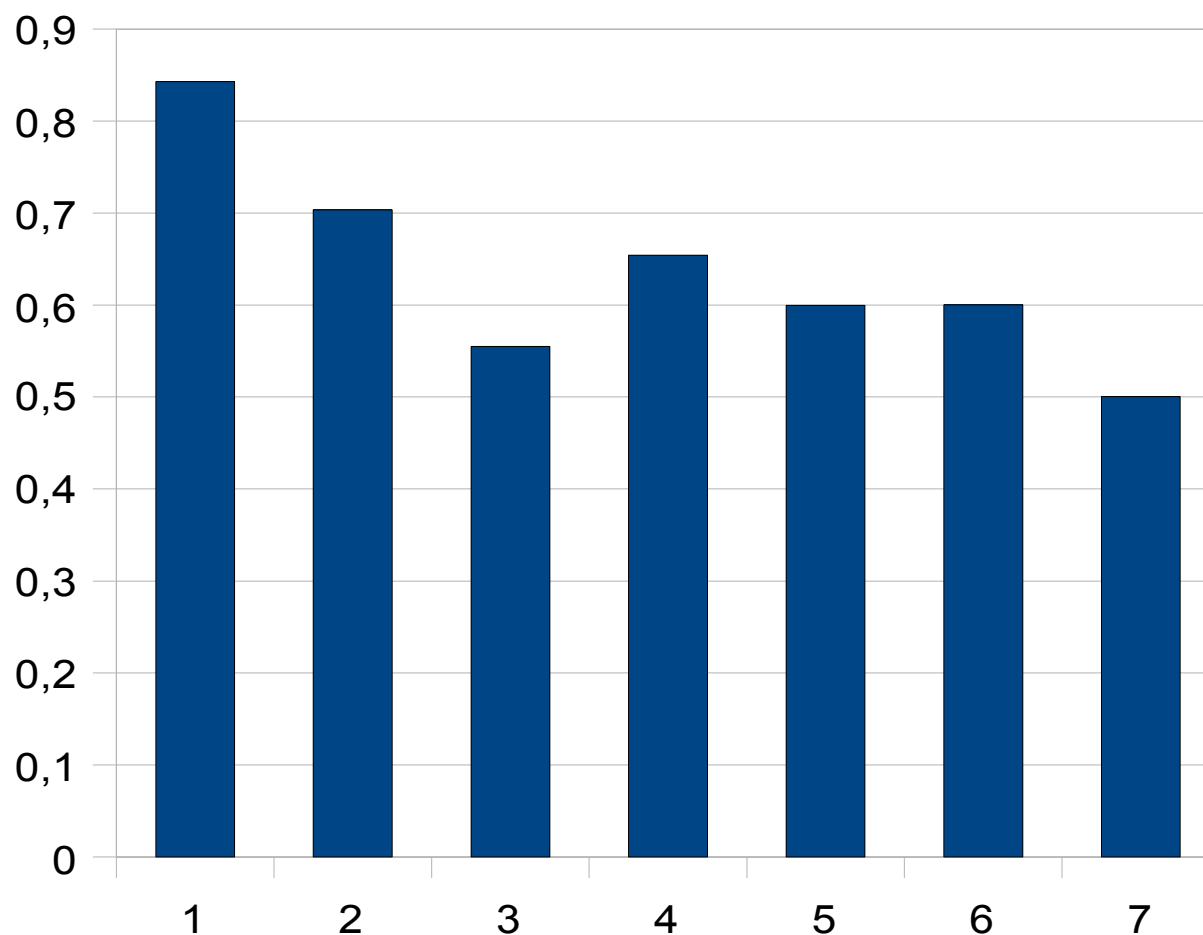
Er reduktioner blot et spørgsmål om artikulatorisk økonomi?

Ville vi opnå bedre kommunikation hvis vi til alle tider udtalte ord tydeligt?

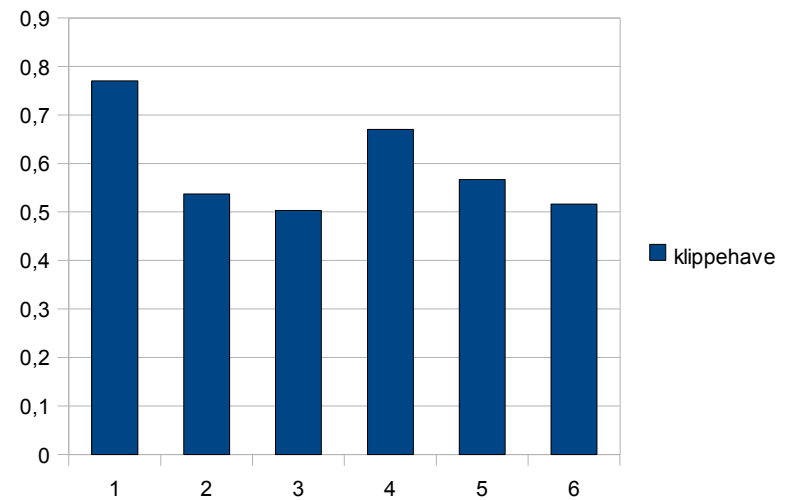
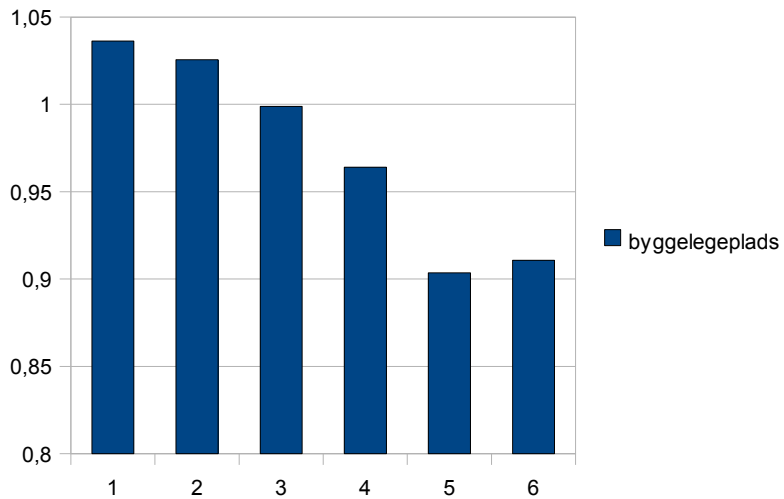
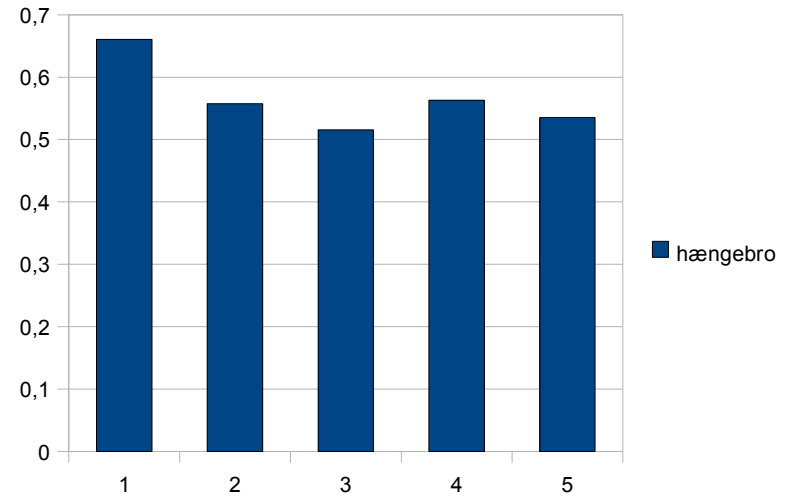
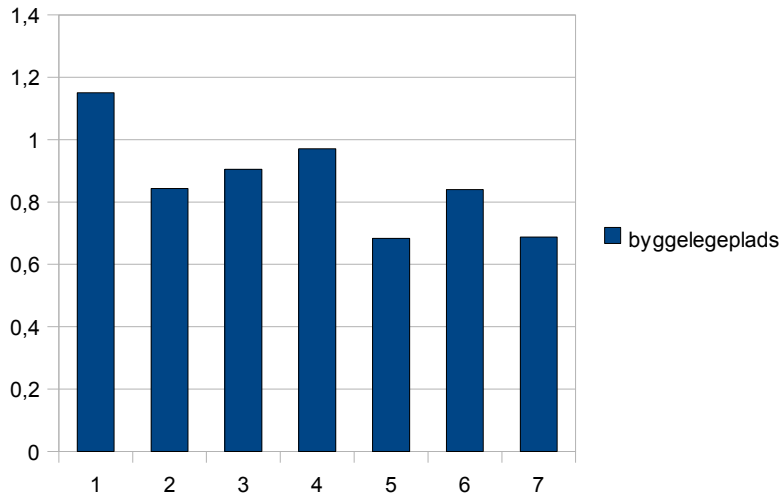
Eller er reduktioner funktionelle?

Ny vs. gammel information

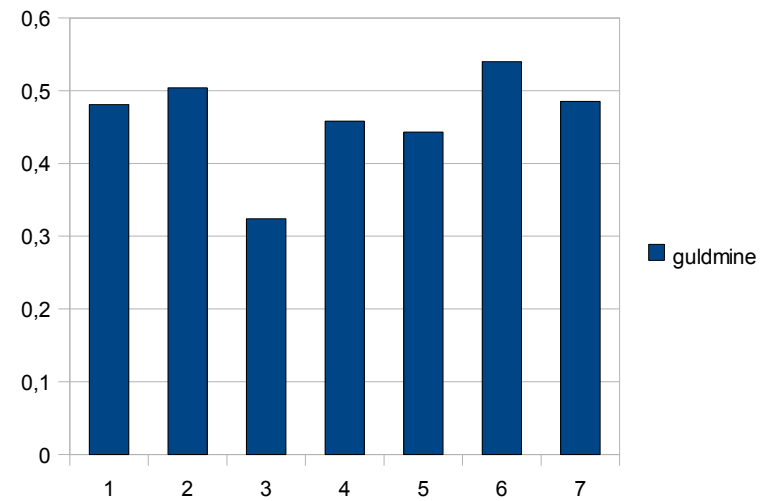
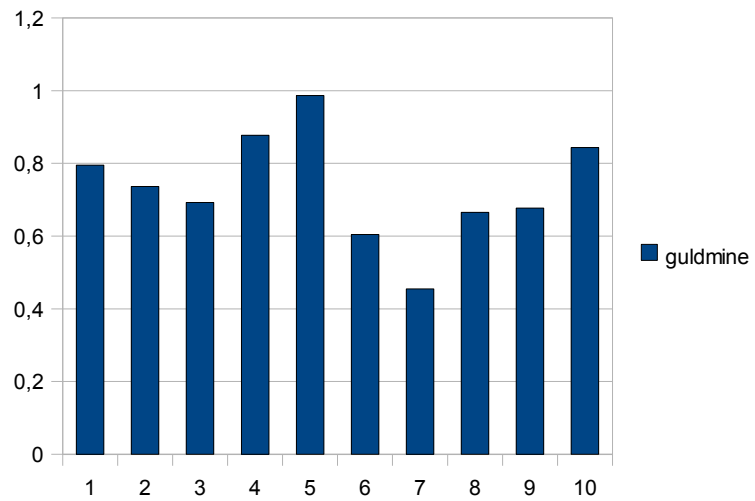
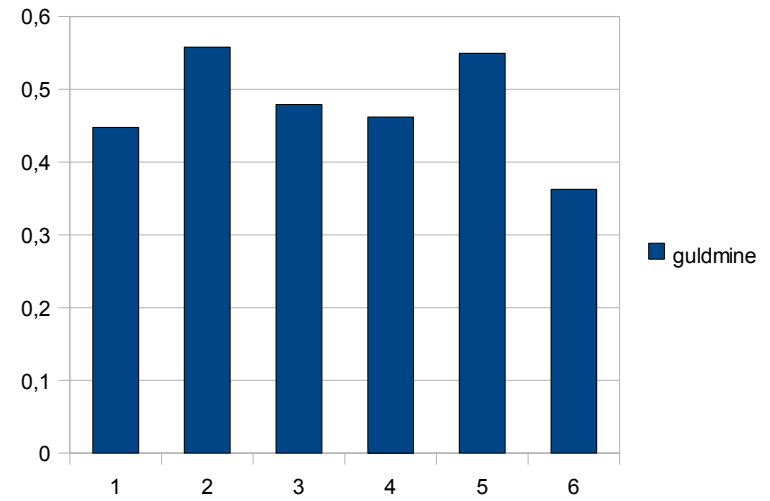
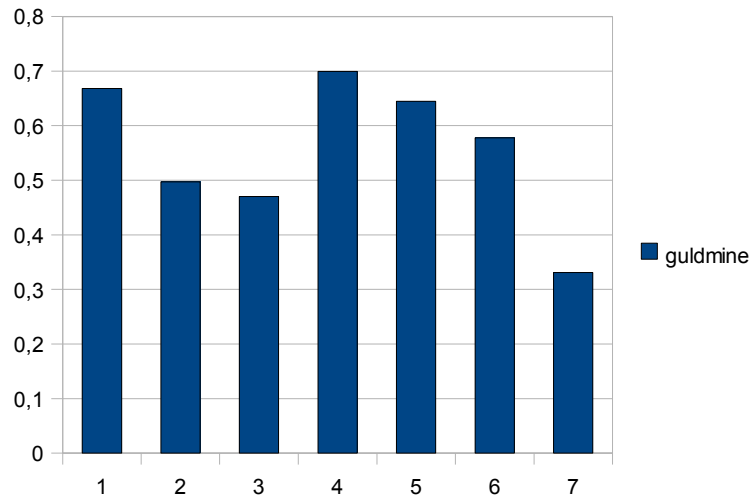
Varighed af 'kirkegården' i løbet af en dialog



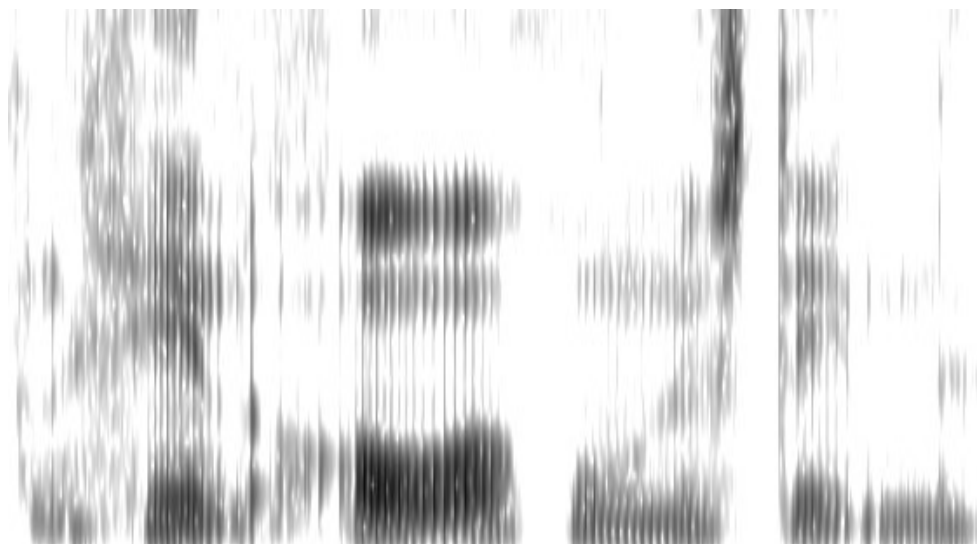
Flere figurer



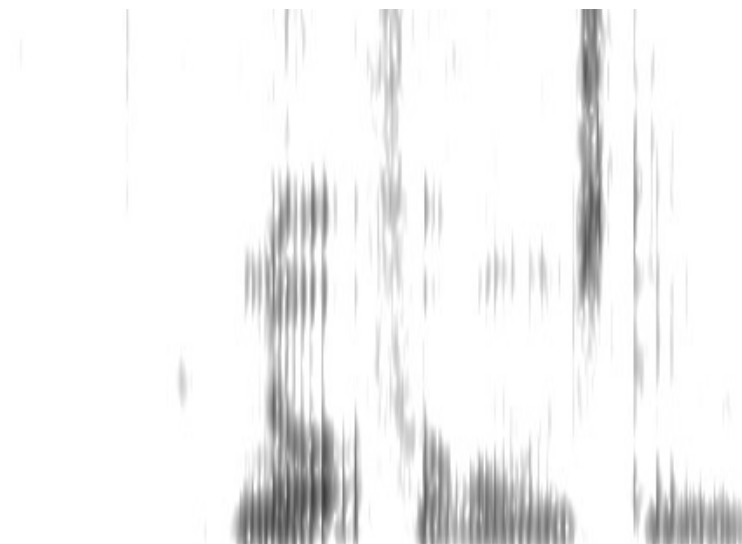
Guldmine-forvirring



Fokus/ikke-fokus



hen over runestenene
[ˈvʉ:nəsdeʔn̩nə]
0.69 sec.



nord for runestenene
[ˈvʉ:nsdeʔn̩]
0.44 sec.

