Time-varying spectral characteristics of Danish stop releases

Universiteit Leiden

Rasmus Puggaard-Rode, Leiden University
Speech Lab Groningen 'winter' school, July 2022

Danish /p t k/

Danish /t/ is strongly affricated, and is sometimes claimed to be an affricate /ts/

- /t/ shows similar behavior to /p k/, which do not appear to be affricates

Otto Jespersen:
Danish is undergoing a sound change whereby [p^h t^h k^h] → [pf ts kx] (1)

Q: What's going on in those stop releases?

Previous research shows:

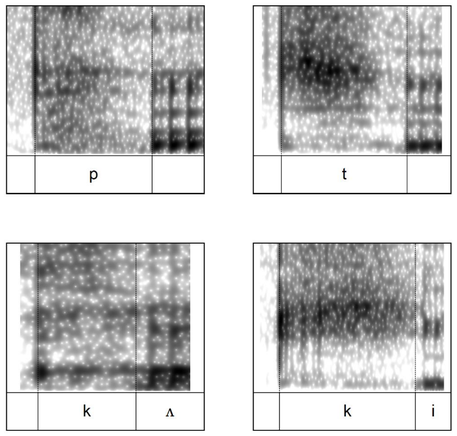
- Closure duration is relatively short
- Peak glottal opening falls relatively late

→ Intraoral air pressure relatively low at the time of release (2, 3)

Aspirates <> Affricates

Composition of aspirates (4)

- Closure (air pressure builds)
- Transient (broad energy distribution)
- Frication (reflects the constriction)



- Aspiration (low frequency noise)

Composition of affricates is similar (5)

- Primary constriction → closure
- Secondary constriction behind the closure → frication

Function-on-scalar regression

- Speech spectra store very complex info
- Many methods in use for boilding down this info to few discrete numbers
  - Spectral moments, (mid-frequency) spectral peak, DCT coefs, …

FOSR models (6) take functional dependent variables, i.e.

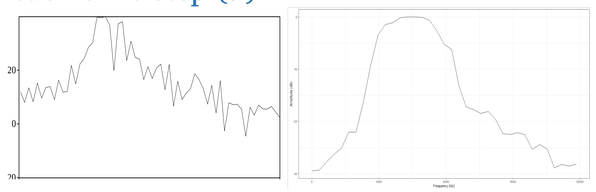
- DV = Amp(Freq) = **the spectrum**
- No need to boil down the info!

Methods

Data come from the DanPASS corpus (7)
Stop releases semi-automatically segmented in Praat (8)

Sound	#	VOT, +str	VOT, -str
/p/	642	57	41
/t/	850	79	68
/k/	842	59	46

Each stop release divided into 20 normalized time steps
Multitaper spectra generated in R for each time step (9)

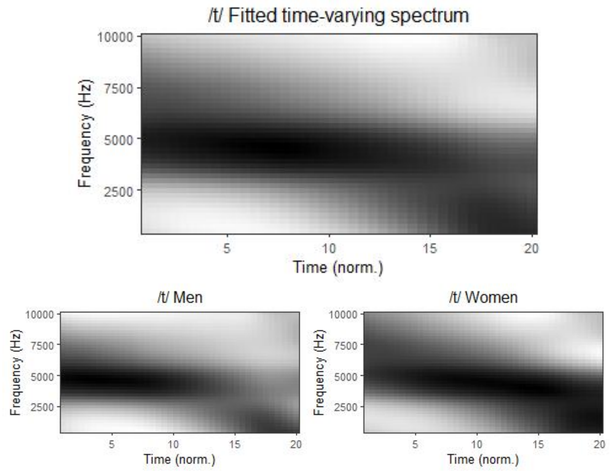


FFT multitaper

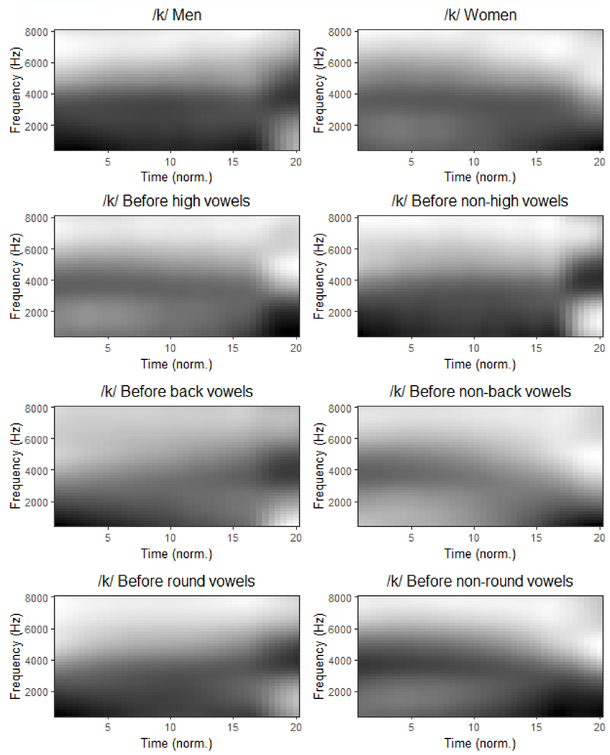
Models fitted in R (10) with this structure:

$$\text{amplitude}_{ij}(F) = \alpha(F) + \gamma(t_{ij}, F) + \text{sex}(t_i, F) + \text{stress}(t_{ij}, F) + \text{height}(t_{ij}, F) + \text{backness}(t_{ij}, F) + \text{roundness}(t_{ij}, F) + \text{speaker}_j\ \gamma(t_{ij}, F) + \text{speaker}_j\ \text{stress}(t_{ij}, F) + \text{speaker}_j\ \text{height}(t_{ij}, F) + \text{speaker}_j\ \text{backness}(t_{ij}, F) + \text{speaker}_j\ \text{roundness}(t_{ij}, F) + \rho e_{i-1} + E_{ij}(F)$$

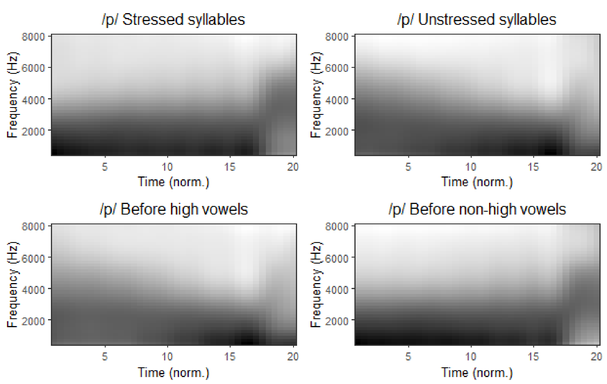
Results /t/



Results /k/



Results /p/



Discussion

- **/t/** is invariably affricated, but latter parts of release are often dominated by affrication
- **/k/** is sometimes affricated, highly affected by CV coarticulation
- **/p/** shows signs of affrication in some contexts, mostly dominated by aspiration

FOSR allows the user to

- Find systematic patterns of variation
- Tease apart influences on results
- Compute the uncertainty associated with variables
  - E.g. using 95% CIs (11)

Useful beyond just stop releases!
Remaining issues include

- Significance testing (perhaps?)
- Fringe effects at right edges

References

(1) Jespersen (1897–1899) *Fonetik* **(2)** Fischer-Jørgensen (1968) *Word* 24, 112–153. **(3)** Frøkjær-Jensen et al. (1973) *ARIPUC* 7, 269–295. **(4)** Stevens (1993) *Speech Comm* 13, 367–375. **(5)** Stevens (1993) *Speech Comm* 13, 33–43. **(6)** Bauer et al. (2018) *Stat Modelling* 18, 346–364. **(7)** Grønnum (2009) *Speech Comm* 51, 594–603. **(8)** Boersma & Weenink, v6.2.04. **(9)** v4.0.3, using packages `tuneR` v1.3.3.1, `multitaper` v1.0-15, and code from Reidy (2013) *OSUWPL* 60, 67–116. **(10)** Uusing package `refund` v0.1-24, visualizations made with `ggplot2` v3.3.5 and `FoSIntro` v1.0.3. **(11)** Marra & Wood (2012) *Scand J Stat* 39, 53–74, implemented in `FoSIntro`.