

– Unless you do it shortly or nasally: How the phonetics of filled pauses determine their subjective frequency and perceived speaker performance

Oliver Niebuhr & Kerstin Fischer

- Hesitations in the form of filled pauses (FPs) such as “err”, “uh”, “um”, and “mmh” generally **have a bad reputation**.
- Rhetorical coaches and manuals strictly recommend speakers to ban them from their speech:
- *The Speaker's Handbook*: “Do not be afraid to pause between sentences or thoughts when you speak. But avoid filling those pauses with distracting and meaningless sounds and phrases [...]” Sprague et al. (2013:336).
- *Here's the Pitch*: Silent pauses are an effective way to “eliminate distracting nonwords such as *ums* and *uhs*” from a speaker's speech.” (Soorjoo 2012:122)
- Bell (2011) provides the reader with “3 tips to eliminate filled pauses from your professional presentation”.

- **Scientifically, this bad reputation of FPs is not justified.**
- For example:
 - FPs facilitate the listeners' cognitive processing of upcoming information (in that they occur before less frequent words or new information) → Corley and Hartsuiker (2003): the “um advantage”.
 - FPs indicate to listeners through phonetic cues how long they will have to wait (Fox Tree 2001) and whether the speaker continues with the same or a different message (Fischer 2000).
 - FPs mitigate potentially impolite utterances (Levinson 1983; Schegloff 2010) and showcase a speaker's affiliation to a specific cultural or social group.
 - FPs convey spontaneity and listener-orientation. That is, they are critical “**contact signals**” (cf. Fischer 2006).

- Based on anecdotal evidence and own experience: Is it the dose that makes the poison?
 - And how is the salience of FPs in a speaker's presentation related to this dose?
- Do listeners systematically over/underestimate the frequency of FPs depending on the FPs' phonetic characteristics?



- 68 experienced business speakers from our own charisma coaching
- All gave an “investor pitch” of about 3-5 minute, L2 English
- 32 females, 36 males; 27-58 years old
- 68 1-minute excerpts (from the middle of the presentation)

- **Acoustic analysis (N=430)**: frequency count (FPs/min), FP duration, FP vowel quality (F1-f0, F2-F1), %Nasal (um, umm, ummm, mmm)

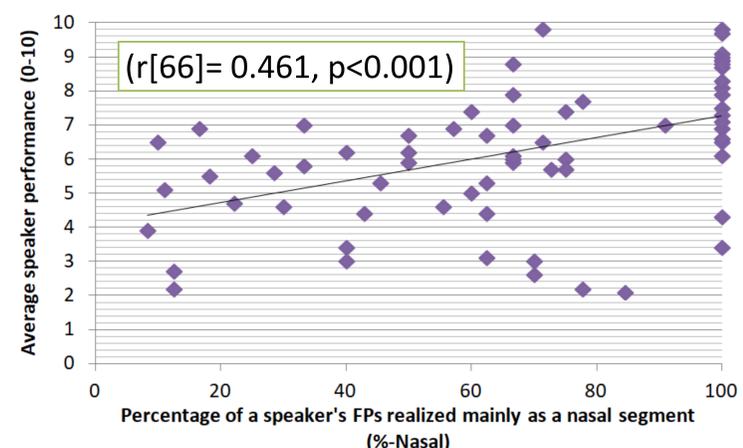
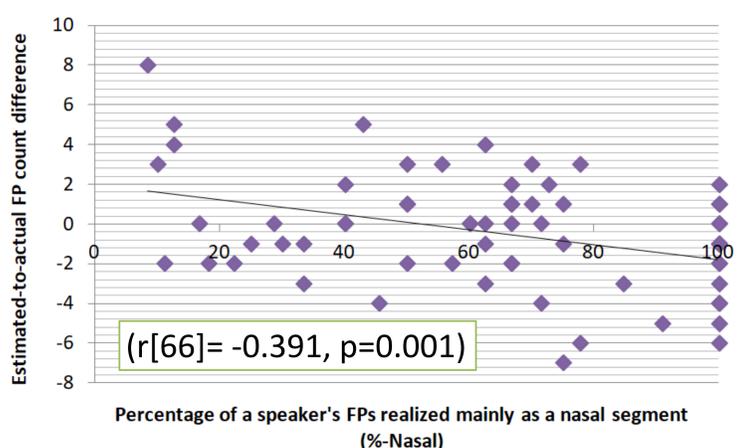
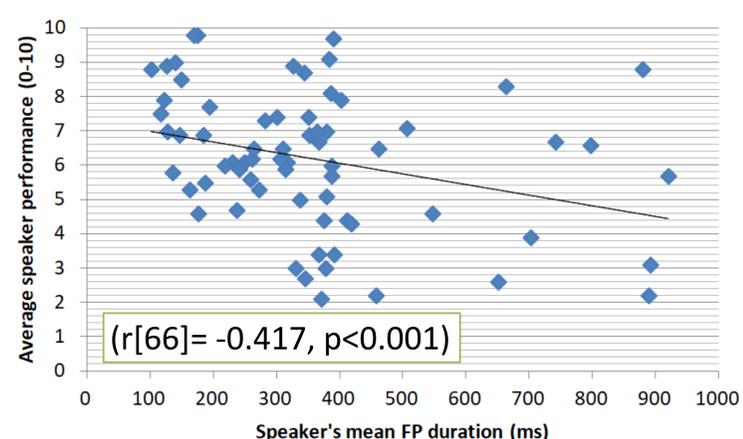
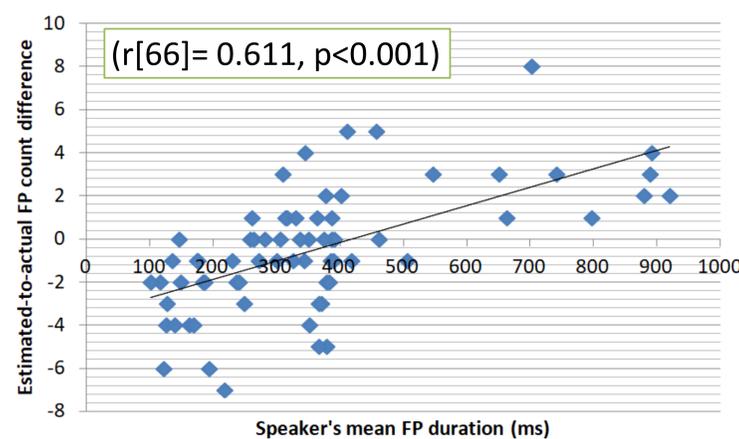
- **Perception**: 29 listeners rated the performance and counted the FPs
- Group A (N=16): Estimate (Do not consciously count!) the total number of FPs in each stimulus → Extra task: shuffle a set of playing cards – only number cards – while listening to the stimulus
- Group B (N=13): After having listened to each stimulus, rate how skilled you perceived the speaker's presentation performance on a 10-point scale from 0='extremely bad' to 10='absolutely excellent'.

- **Yes, the dose makes the poison → The more FPs were produced the higher their estimated frequency, the lower the rated performance**
 - Listeners are good at estimating number of FPs – However, the phonetic characteristics of FPs (not vowel quality) matter as well!!
 - Listeners underestimate the actual FP frequency when FPs are <400ms and >50% nasalized → Speakers are rated better

In accordance with previous studies, we found strong inter-individual variation along all three acoustic dimensions, while, at the same time, speaker-specific FP profiles emerged as well.

Listeners were most precise at estimating a speaker's actual FP count, when the FPs were “normal” realizations (i.e. with all parameters close to the German average)

We assume the effect of over-/underestimation to be due to the perceptual salience of FPs; shorter and a higher nasal-segm. percentage (lower energy) = less salient = more FPs are missed



The strict FP ban of rhetoric is unnecessary and premature. FPs fulfill important communicative functions, and trying to reduce a speaker's FP frequency is only useful if their number is exceptionally high (> 8/min). Working on the **quality** of filled pauses is more effective in terms of improving a speaker's performance rating.