

F0 of Adolescent Speakers

First Results for the German Ph@ttSessionz Database

Chr. Draxler, F. Schiel, T. Ellbogen
BAS Bavarian Archive of Speech Signals
University of Munich, Germany

Introduction

- previous f_0 studies of adolescents
 - small numbers of speakers
 - limited and artificial speech material, e.g. sustained vowels
 - no speech data available
- forensic databases
 - not available for German

Ph@ttSessionz: Goals

- 1000 speakers
- 50% male, 50% female ($\pm 5\%$)
- 13-19 years
- good dialect coverage
- recorded via Internet in secondary schools
- 22.05 kHz, 16 bit linear PCM, stereo

Session Contents

item	#	item	#
isolated digit	10	date	3
numbers 11-100	19	time	3
PC command phrases	12	directory assistance	9
telephone numbers	13	spelling	10
mobile phone keys	3	phonetically rich	30
credit card	3	spontaneous	5
PIN	3	narrative	2

Session Contents

item	#	item	#
isolated digit	10	date	3
numbers 11-100	19	time	3
PC command phrases	12	directory assistance	9
telephone numbers	13	spelling	10
mobile phone keys	3	phonetically rich	30
credit card	3	spontaneous	5
PIN	3	narrative	2

- SpeechDat and RVG-I compatible

Speaker Data

- date of birth, sex, weight, height
- dialect region (federal state at age 6)
- mother tongue of speaker and family
- smoking habits, dental braces, piercings

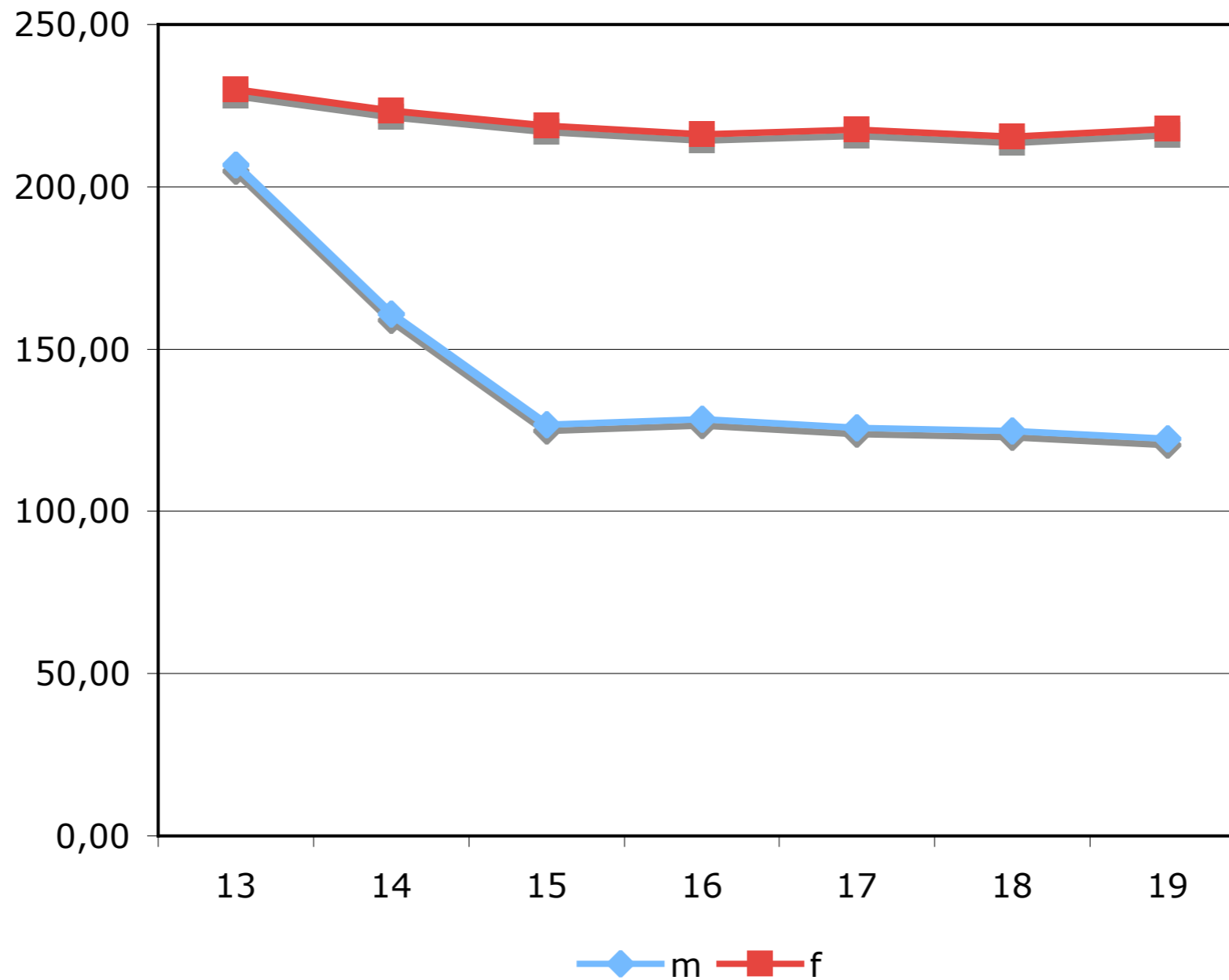
F0 Analysis

- pre-release version of the database
 - 762 speakers
 - ~ 49% f, 51% m
 - good age distribution
 - biased dialect region distribution
- 90829 utterances

F0 Calculation

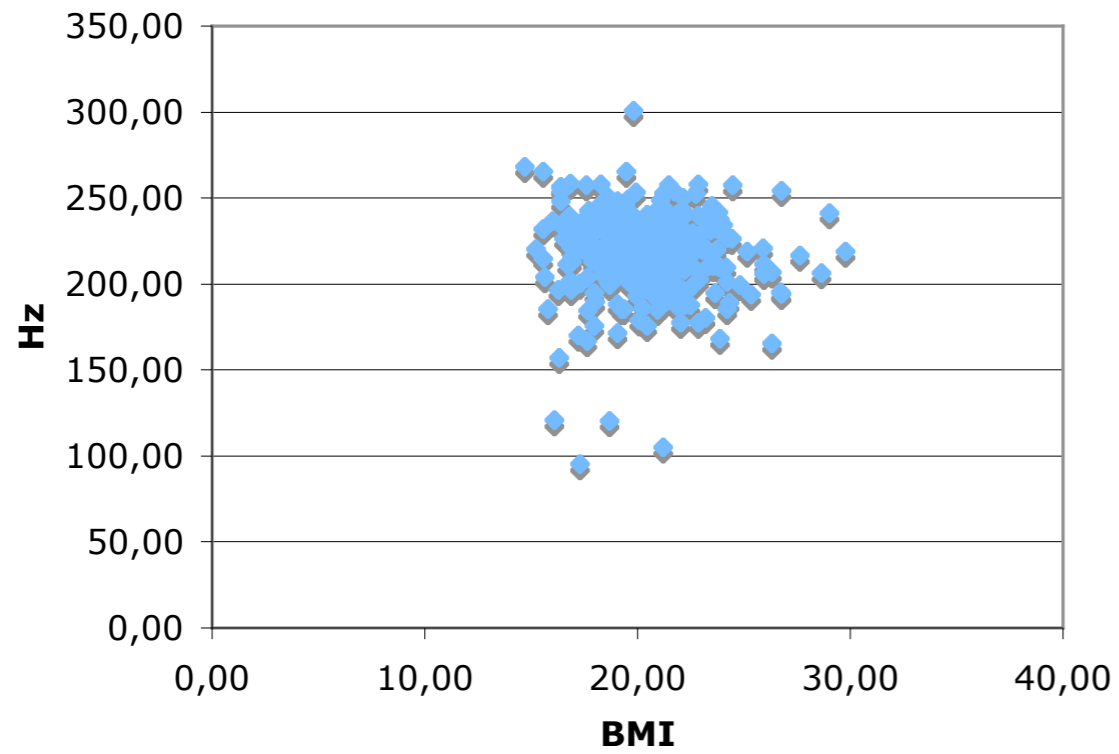
- Praat built-in algorithm
 - frequency 75-400 Hz
 - max candidates 15
 - silence/voicing threshold 0.03/0.45
 - octave/jump/voiced cost 0.01/0.35/0.14
- f0 mean, min, max (in Hz and mel)

F0_{mean} vs. Age

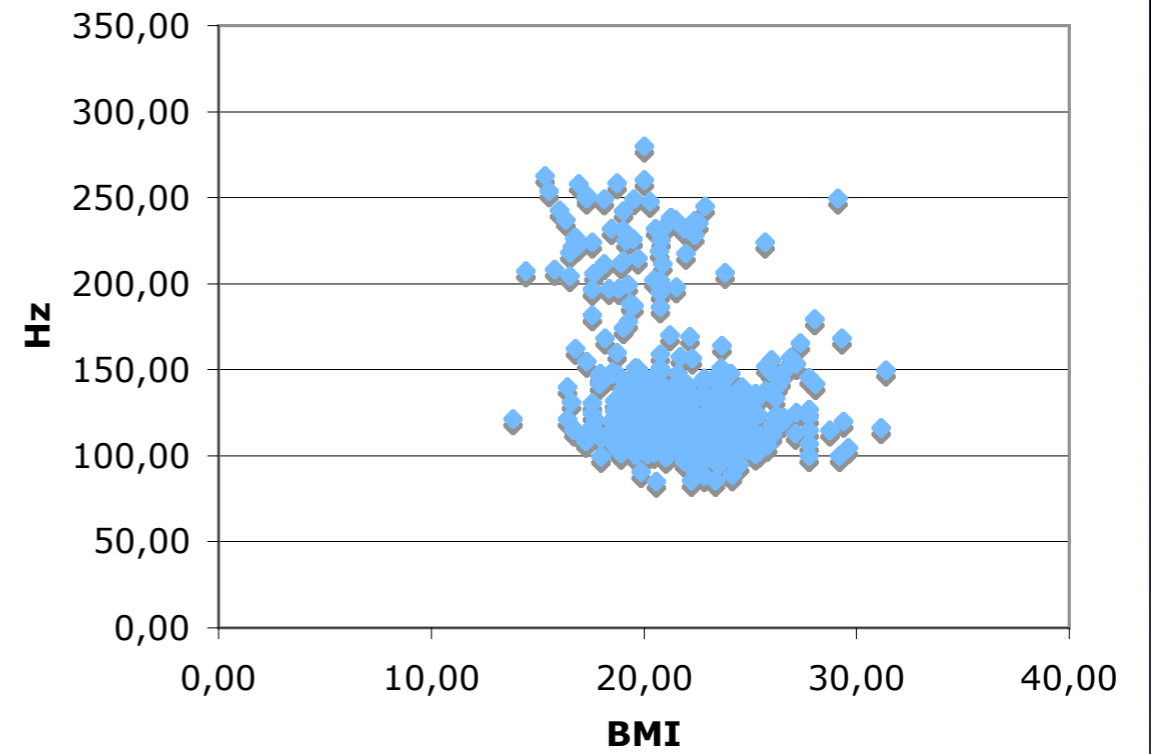


F0 vs. BMI

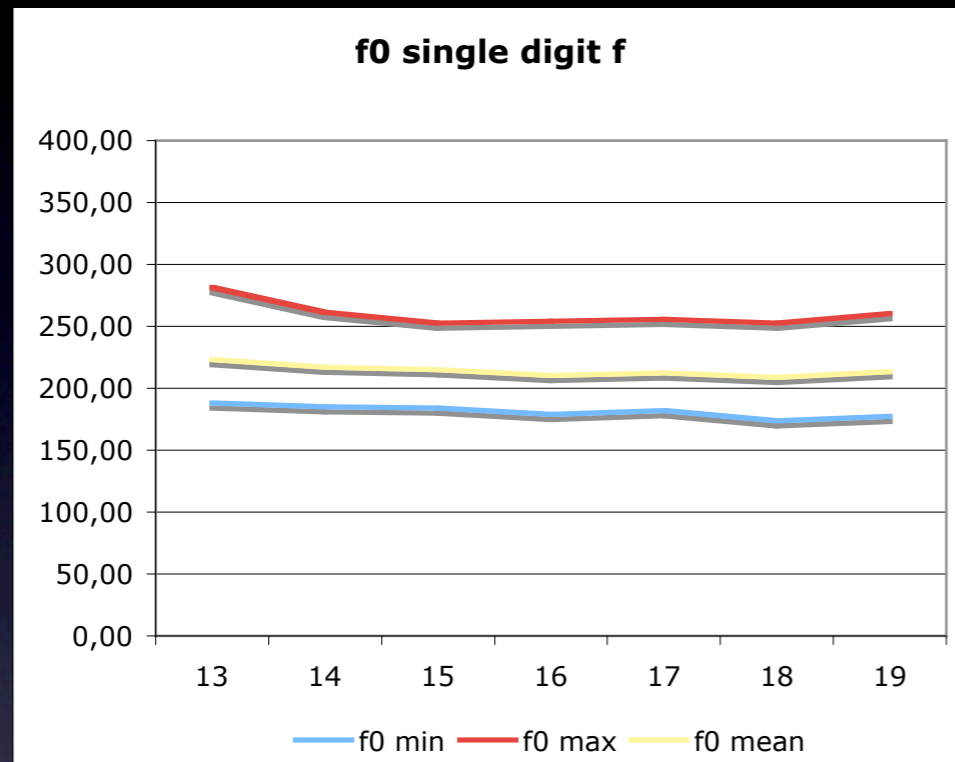
mean f0 vs. BMI (female)



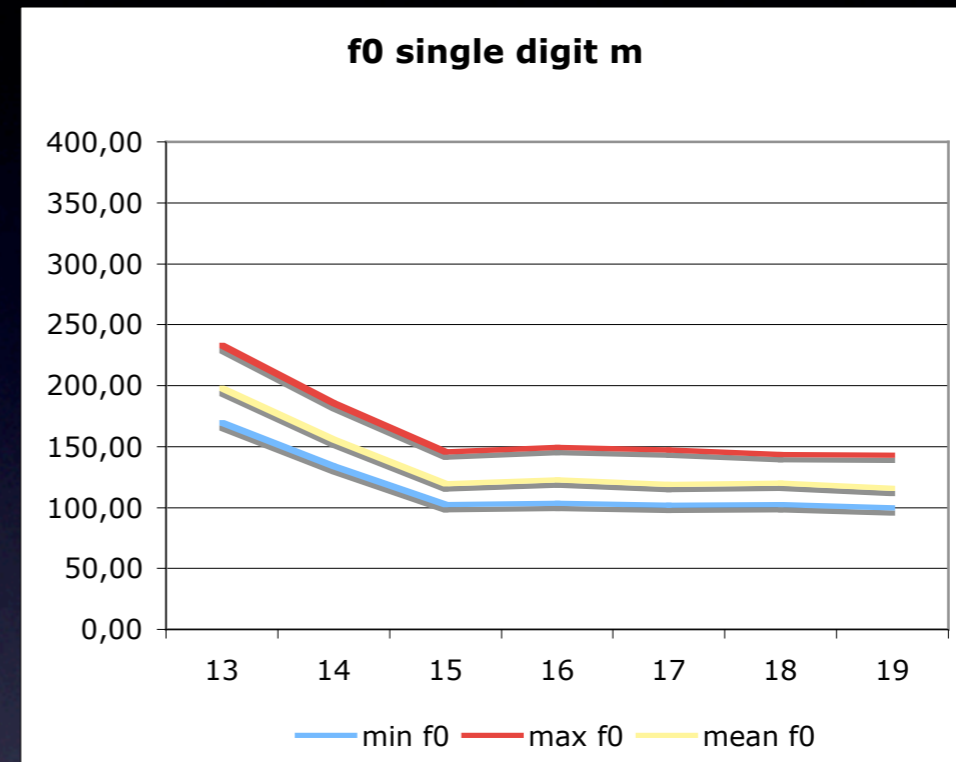
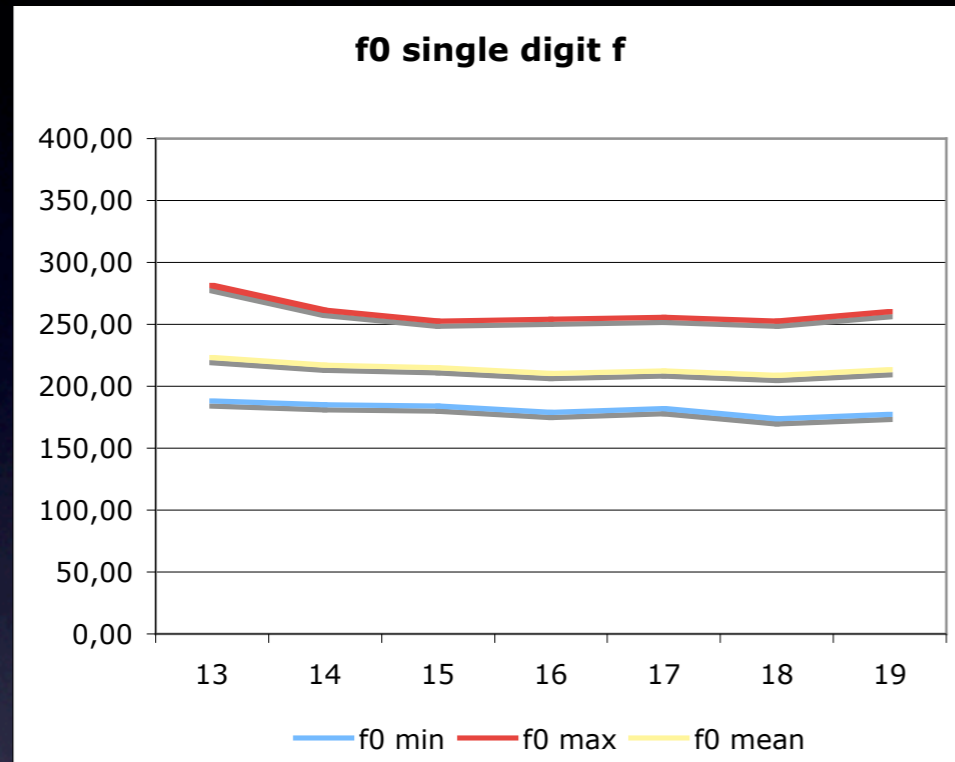
mean f0 vs. BMI (male)



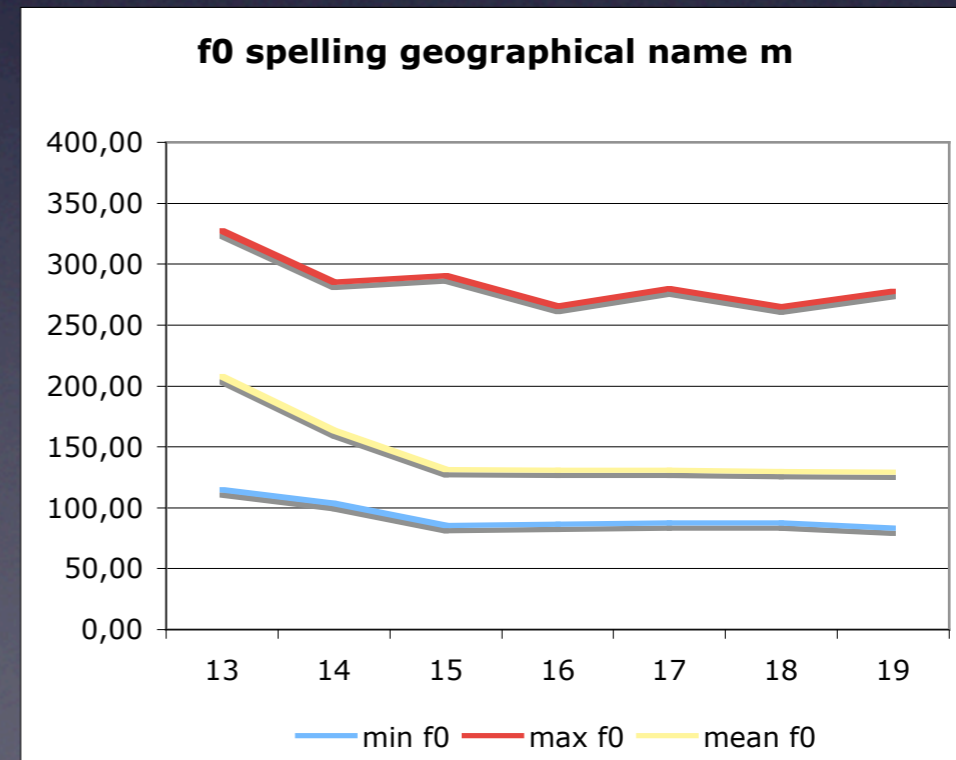
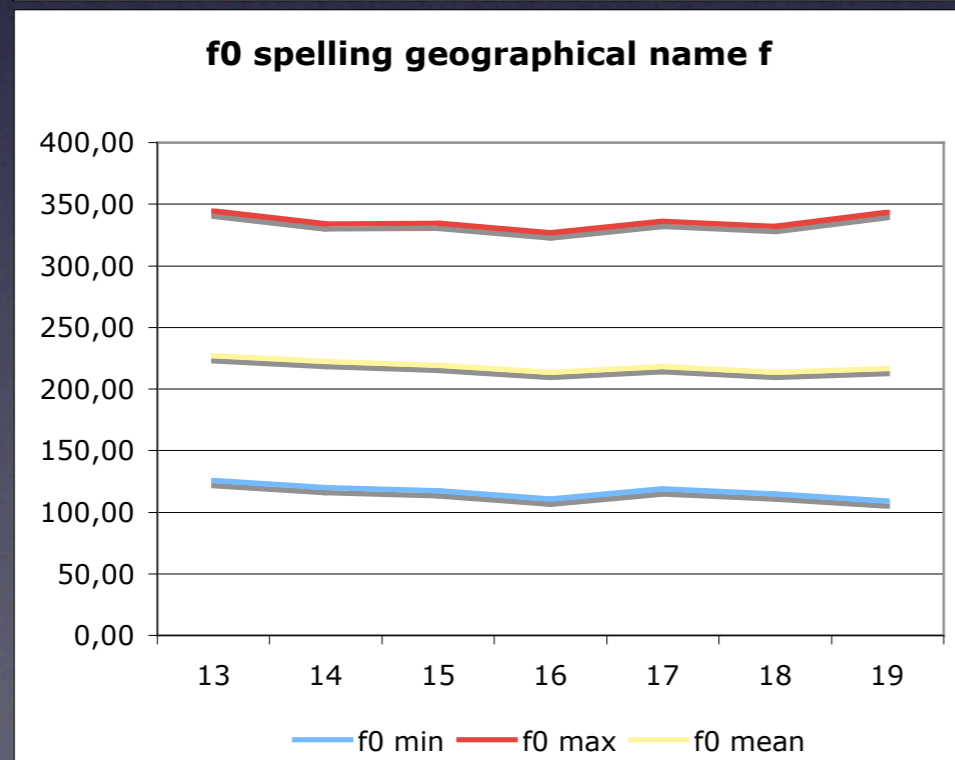
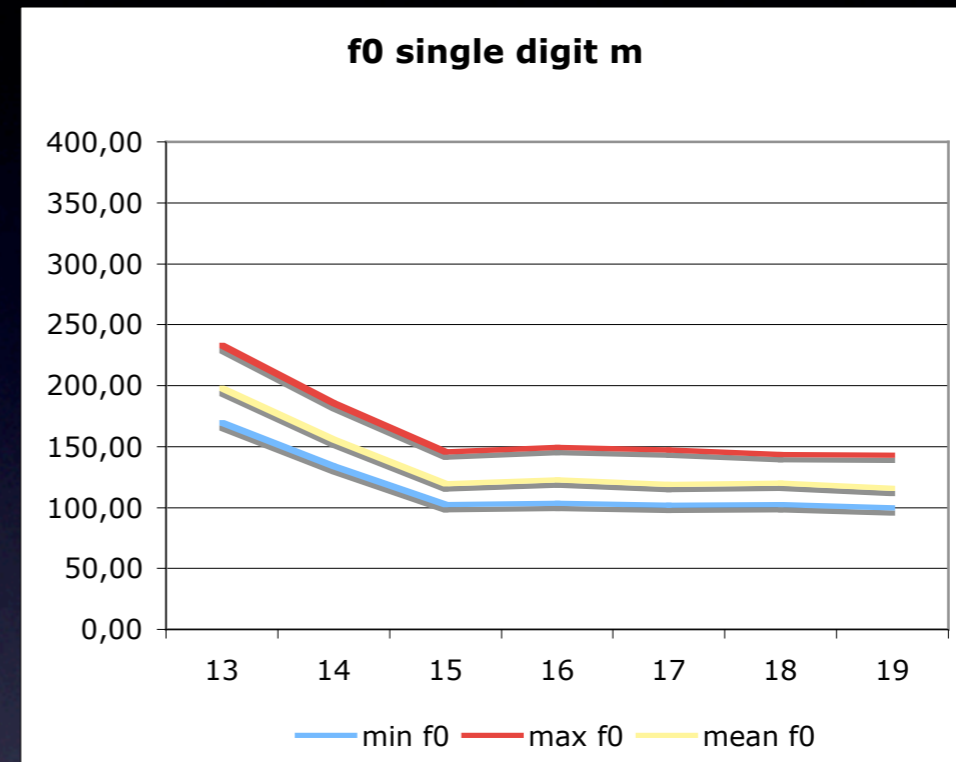
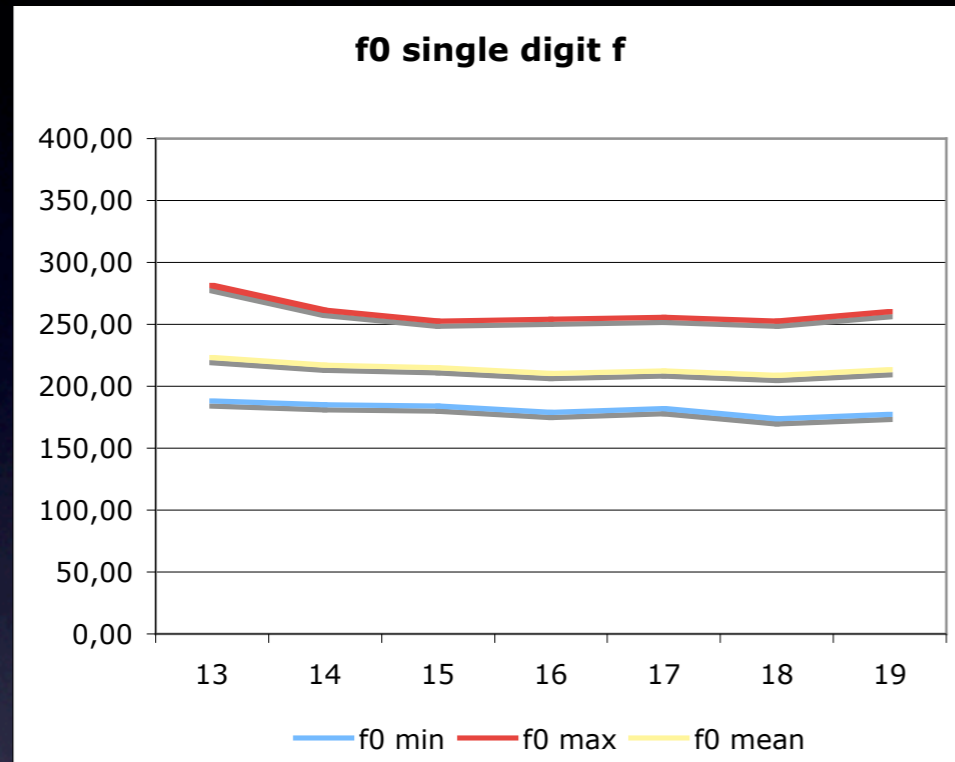
F0 Data



F0 Data



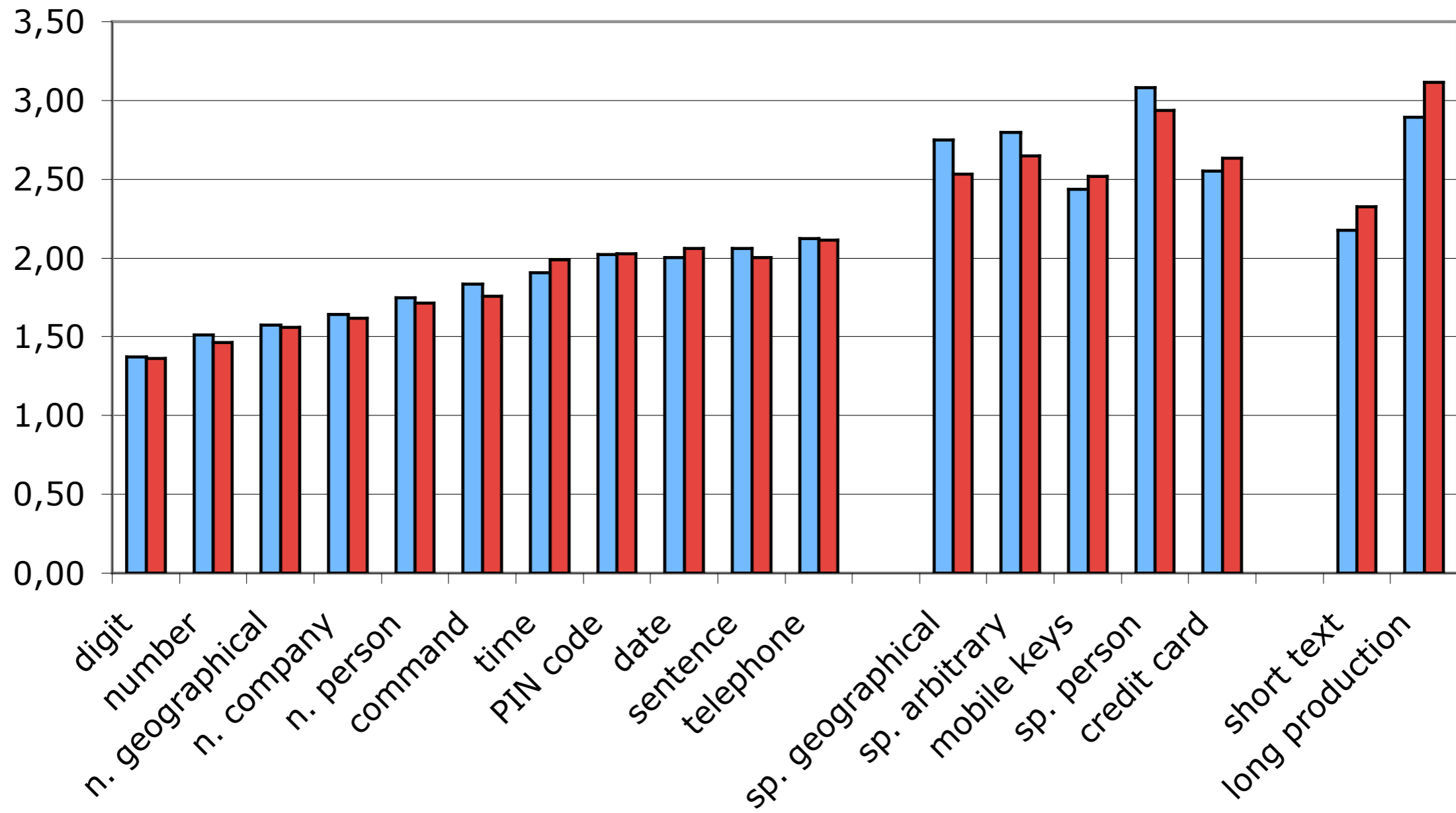
F0 Data



F0 Range

- $F0_{\text{abs}} = F0_{\text{max}} - F0_{\text{min}}$
- $F0_{\text{rel}} = F0_{\text{max}} / F0_{\text{min}}$
- scale
 - absolute Hz scale
 - perception-based mel scale

F0_{rel} mel



Outlook

- use final release of the database
 - 864 speakers
- refine analysis
 - re-compute F0 for phrases

Summary

- Ph@ttSessionz database
 - largest database for adolescent speakers
 - technology development and research
- statistically reliable voice data for German
- F0 variation dependent on utterance class

Summary

- Ph@ttSessionz database
 - largest database for adolescent speakers
 - technology development and research
- statistically reliable voice data for German
- F0 variation dependent on utterance class?