



F0 statistics for speakers of Urdu

Sarmad Hussain Univ. of Engineering and Technology, Lahore, Pakistan
Saad Irtza Univ. of Engineering and Technology, Lahore, Pakistan
Toby Hudson Univ. of Cambridge, United Kingdom
Gea de Jong Philipps-Universität Marburg, Germany


IAFPA 2012
 Santander
 Spain



Fundamental frequency: context

- F0 is robust
 (e.g. impervious to background noise and telephone transmission)
 - Useful parameter for speaker identification
 (Nolan 1983, French 1990a, Hollien 1990, Künzel 1987, Rose 2002)
- Individual differences in the length of the vocal folds (Titze 1994, Nolan 1983)
- 

Fundamental frequency: context


- Average F0 – Gaussian population statistics
 ~ probability
 (Künzel 1987, Hudson et al. 2007)
 - Variability measures (SD-F0, VarCo-F0)
 ~ Speaking melody / monotony
 (Künzel 1987, Jessen et al. 2005)
- 

Fundamental frequency: context

Many factors that can cause a change in F0.
 Caution required! (Braun 1995)

Factors influencing F0

External:	Tape speed speaking style (spontaneous vs read) PRAAT settings (pitch thresholds)
Internal:	age smoking language/ethnological background



Fundamental frequency: context

Factors influencing F0


Internal: **language/ethnological background**

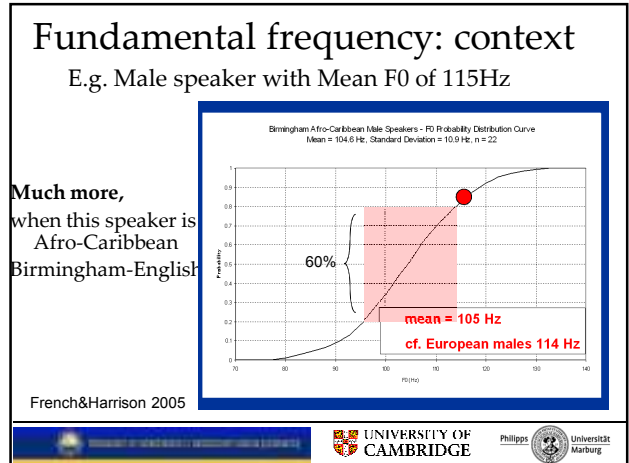
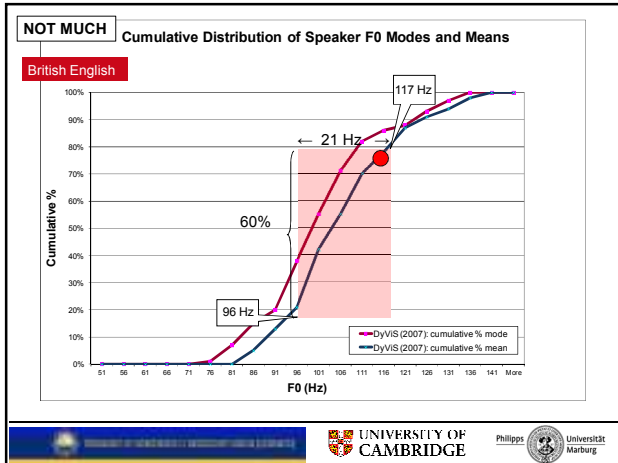


Fundamental frequency: context

E.g. Male speaker with Mean F0 of 115Hz

This means?





Fundamental frequency: context
Useful: F0 data for language/ethnological background!

Research so far:

Poles vs US Americans	(Majewski et al 1972)
Germans vs US Americans	(Scherer 1979)
Turks vs Germans	(Braun 1992b)
Portuguese	(Guimarães & Abberton, 2005a)
Finnish	(Leino 1998)
Swedish	(Pegoraro Krook, 1988, Lindh 2006)

Fewer probability curves:

- German: Künzel 1989, Jessen 2005
- British English (SBE): Hudson et al. 2007
- British English (Birmingham: Afro-Caribbean): French, unpubl.

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Urdu

4th in the list of languages by number of speakers: 100 - 240 Million

Urdu –	Pakistan:	10 million (Urdu)
	India:	48 million (Hindi-Urdu)

(Lewis 2009)

Other regions with significant populations of Urdu speakers:

Bangladesh/Nepal/Afghanistan	
United States	150,000
Saudi Arabia	120,000
Canada	80,000
United Arab Emirates	50,000
United Kingdom	25,000

http://en.wikipedia.org/wiki/Urdu_people
Census: 2001

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Urdu vs Hindi

Urdu-Hindi: from same source language: the Khari Boli dialect of Delhi

Urdu	Hindi
Borrow-words from Arabic and Persian	Borrow-words from Sanskrit
Written in Perso-Arabic Script	Written in devanagari script
تمام انسان آزاد اور حق و عزت کے اعتبار سے برابر ہونے میں اسٹیجی مشمول اور عقل و حکمت علی سے اسٹیجی نہیں ایک دوسرے کے ساتھ جہلی پائے کا ملکا کرنا چاہیے۔	सभी मनुष्यों को गौरव और अधिकारों के मामले में जन्मजात स्वतन्त्रता और समानता प्राप्त है। उन्हें बुद्धि और अन्तरात्मा की देन प्राप्त है और परस्पर उन्हें भाईचारे के भाव से बर्ताव करना चाहिए।

<http://www.omniglot.com/writing/urdu.htm>
<http://www.omniglot.com/writing/hindi.htm>

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Urdu vs Hindi

Urdu	Hindi
Borrow-words from Arabic and Persian	Borrow-words from Sanskrit
Written in Perso-Arabic Script	Written in devanagari script
Urdu speech corpus 2010	सभी मनुष्यों को गौरव और अधिकारों के मामले में जन्मजात स्वतन्त्रता और समानता प्राप्त है। उन्हें बुद्धि और अन्तरात्मा की देन प्राप्त है और परस्पर उन्हें भाईचारे के भाव से बर्ताव करना चाहिए।

<http://www.omniglot.com/writing/urdu.htm>
<http://www.omniglot.com/writing/hindi.htm>

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Urdu speech corpora

The Enabling Minority Language Engineering (EMILLE) corpus, 2004.

The U.S. Army Research Laboratory (ARL) Urdu Speech Database, 2007.

Urdu speech corpus 2010,
The National University of Computer and Emerging Sciences
Lahore, Pakistan (Sarfranz et al. 2010).



Urdu speech corpora

The Enabling Minority Language Engineering (EMILLE) corpus

Source: The Evaluations and Language Resources Distribution Agency (ELDA)
EPSRC project at Lancaster University and Sheffield University

Languages: Urdu, plus Bengali, Gujarati, Hindi, Punjabi, Singhalese and Tamil

Content: 2,627,000 words of transcribed spoken data for per language
Smaller for Assamese, Kannada, Kashmiri, Malayalam, Marathi, Oriya, Sinhala, Telegu

Release date: September 2004

Costs: Free for academic research purposes



Urdu speech corpora

The U.S. Army Research Laboratory (ARL) Urdu Speech Database:

Content: 200 adult native speakers of Urdu
Origin: Pakistan and Northern India
Dialects: South Sindh, North Sindh, South Punjab, North Punjab, North West regions and Baluchistan. (26-30 speakers per dialect)
Source: Linguistic Data Consortium (LDC), Philadelphia
Release date: February 2007.
Costs: Free for members



Urdu speech corpora

The U.S. Army Research Laboratory (ARL) Urdu Speech Database:

Content: 200 adult native speakers of Urdu
Origin: Pakistan and Northern India
Dialects: South Sindh, North Sindh, South Punjab, North Punjab, North West regions and Baluchistan. (26-30 speakers per dialect)
Source: Linguistic Data Consortium (LDC), Philadelphia
Release date: February 2007.
Costs: Free for members




Linguistic Data Consortium (LDC)



	<p><i>Item Name:</i> ARL Urdu Speech Database, Training Data <i>Authors:</i> Appen Pty Ltd, Sydney, Australia <i>LDC Catalog No.:</i> LDC2007S03 <i>ISBN:</i> 1-58563-412-3 <i>Release Date:</i> Feb 20, 2007 <i>Data Type:</i> speech <i>Sample Rate:</i> 22050 Hz <i>Sampling Format:</i> pcm <i>Data Source(s):</i> microphone speech <i>Language(s):</i> Urdu <i>Language ID(s):</i> urd <i>Distribution:</i> 8 DVD, Web Download <i>Member fee:</i> \$0 for 2007 members <i>Non-member Fee:</i> US \$4000.00 <i>Reduced-License Fee:</i> US \$2000.00 <i>Extra-Copy Fee:</i> US \$1600.00</p>
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Urdu speech corpus 2010




The National University of Computer and Emerging Sciences

Details: Sarfraz et al. 2010



Contact: Sarmad Hussain
sarmad.hussain@kics.edu.pk

Designed for

- Speech recognition using CMU Sphinx Toolkit
- General research





geology.com

Methodology: materials

Speakers:

- 82 participants (41 male and 41 female)
- Recruited from a university campus and nearby residences (including students, faculty and staff)
- Accent: Native speakers of the Lahore suburban Urdu accent
- Age: ranging from 20 to 55 years without speech impediments.

Methodology: materials



Type of speech sample:

Spontaneous speech from an interview (+ read speech)

Subjects were asked to talk about topics like their daily routine, hobbies and interests or past experiences.

Environment: quiet offices and homes.

Microphone (+ Telephone recording)







Methodology: materials

Tech specs:

- Dell Latitude E5400 laptop + Logitech USB Desktop microphone
- Linksys SPA400 telephone gateway (telephone recording)
- Format: 16 kHz

Audio preparation:



- 2-3 mins 
- from end of interview
- intrusive sounds eliminated





Methodology: PRAAT – F0 script

Analysis settings: Thank you P. Harrison!

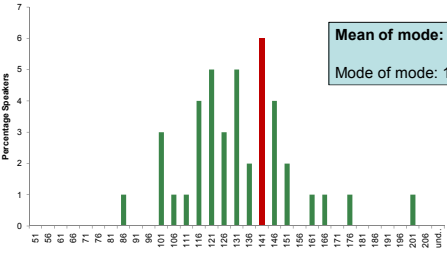
- 50 - 300 Hz male speakers (same as DYVIS)
- 50 - 500 Hz female speakers
- 5 Hz bins (smooths out perturbations)
- Other settings:
 - Time step = 0.01 s
 - Max. number of candidates = 15
 - Very accurate = yes
 - Silence threshold = 0.03
 - Voicing threshold = 0.45
 - Octave jump cost = 0.5
 - Voiced/unvoiced cost = 0.14
 - Acoustic periodicity detection – autocorrelation





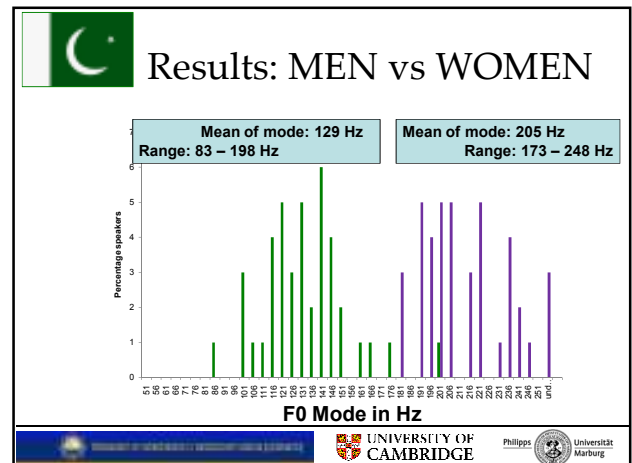
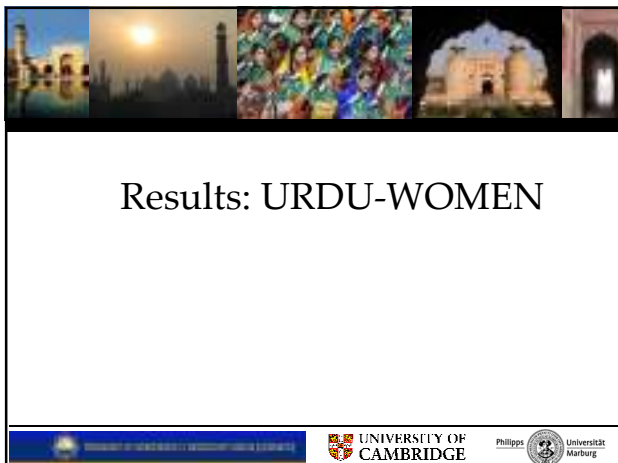
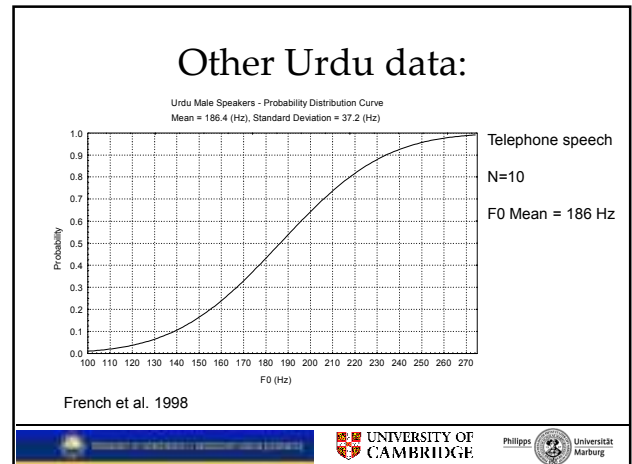
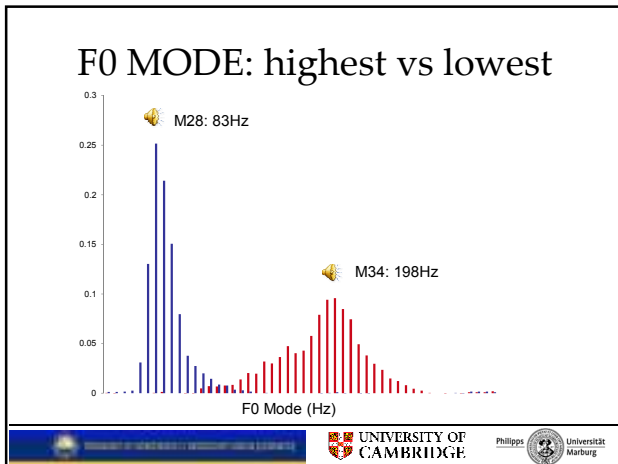
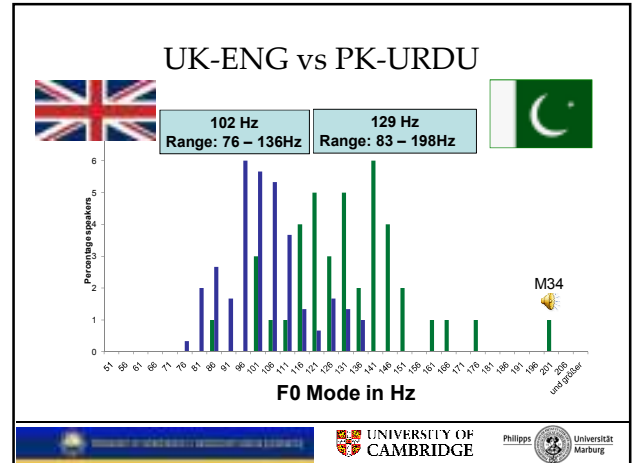
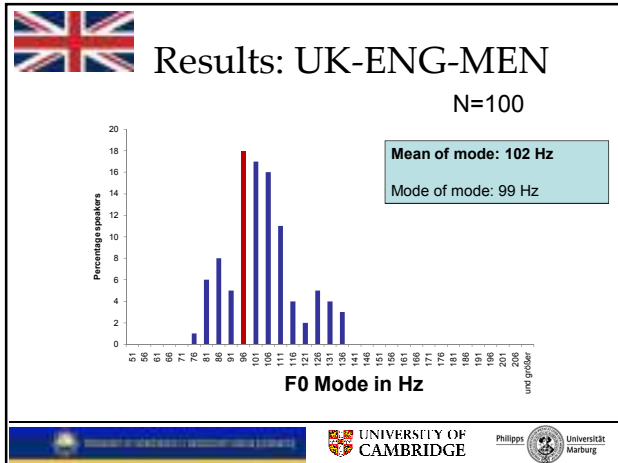
Results: PK-URDU-MEN

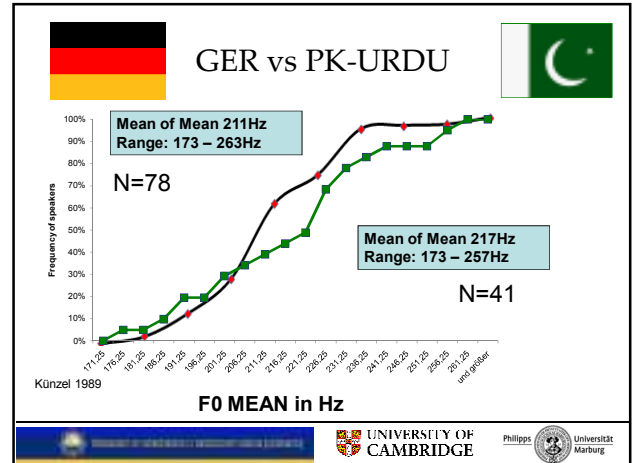
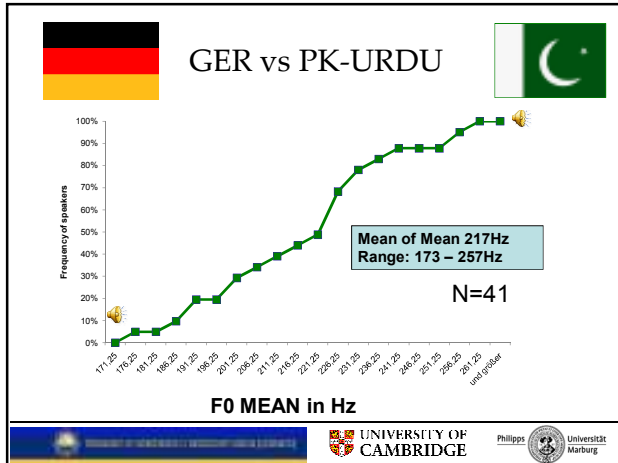
N=41



F0 Mode in Hz





Conclusions

Male Urdu speakers (from the Lahore area) exhibit a much higher F0 compared to known data for (West-European) speakers of English and German.

Female speakers of Urdu exhibit similar F0 values (to Eng and GER).

F0 probability data for one group of speakers cannot be directly extrapolated to speakers from a different language/ethnological background.

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Forensic implications

When working on cases that involve speakers with a different language/ethnological background, it is important to consult F0 probability data for that particular speaker community.

When the case involves Urdu speakers from the Lahore area, one can expect higher F0 values compared to speaker data for German and English (from speakers with a West-European ethnological background).

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

References:

Braun, A. (1995). Fundamental frequency - how speaker-specific is it? In: Braun, A., Köster, J.-P. (eds.), *Studies in Forensic Phonetics*: BEIPHOL 64, 9-23.

Census 1999 report of Pakistan. Islamabad: Population Census Organization, Statistics Division, Government of Pakistan, 2001.

Harrison, P. Praat script for the analysis of F0.

Hudson, T., de Jong, G., McDougall, K., Harrison, P. and F. Nolan. (2007). F0 statistics for 100 young male speakers of Standard Southern British English. ICPHS conference proceeding. In: J. Trouvain and W. Barry.

Hollien, H. (1990). *The acoustics of crime*. New York: Plenum Press.

French, P., Saxena, M., Harrison, P. and Kunezel, H. Normative f0 Data for Panjabi and Urdu: Preliminary Report on Male Speakers over Telephone Lines. Paper delivered to the 7th IAPF(A) Conference, Voorburg, Netherlands, 1998.

French, P., and Harrison, P. 2005 Lay-witness voice description and identification: consideration of shouting, pitch and accent. IAFFA, Marrakesh, Morocco, 2005.

Jessen, M., Köster, O. and Gfroerer, S. 2005 Influence of vocal effort on average and variability of fundamental frequency. *Speech, Language and the Law* 12(2).

Künzel, Hermann J. (1987): *Sprechererkennung: Grundzüge forensischer Sprachverarbeitung*. Heidelberg: Kriminalistik Verlag.

Künzel, Hermann J. (2000): Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics* 7: 149-79.

Lewis, M. Paul (ed.). *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex: SIL International, (2009). Online version: <http://www.ethnologue.com/>

Masica, Colin (1991). *The Indo-Aryan Languages*, Cambridge: Cambridge University Press

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

References:

Nolan, F. (1983). *The phonetic bases of speaker recognition*, Cambridge University Press.

Nolan, F., McDougall, K., de Jong, G. and Hudson, T. (2006a). 'A forensic phonetic study of 'dynamic' sources of variability in speech: the DyVIS project.' In: P. Warren and C.I. Watson (eds.), *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*, 6-8 December 2006, Auckland: Australasian Speech Science and Technology Association, 13-18.

Reubold, U., Harrington, J., Kleber, F. 2010 Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52 (2010) 638-651.

Sarfraz, H., Hussain, S., Bokhair, R., Raza, A.A., Sarfraz, Z., Pervez, S., Mustafa, A., Javed, I. and Parveen, R. (2010). Speech corpus development for a speaker independent spontaneous Urdu speech recognition system, *Proceedings of O-COCOSDA*, Kathmandu, Nepal.

Titze, I. 1994 *Principles of Voice production*, Englewood Cliffs, NJ: Prentice Hall.

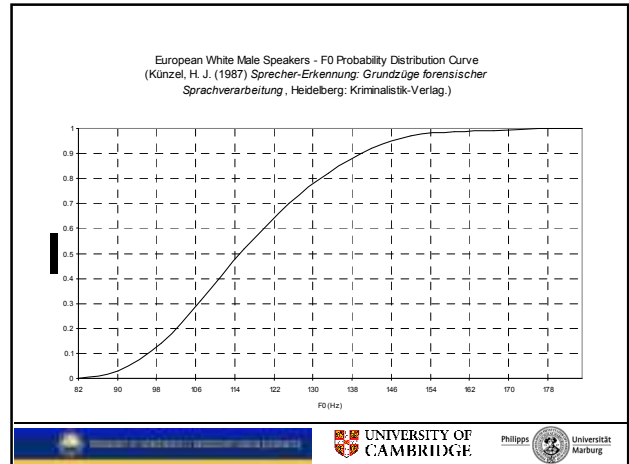
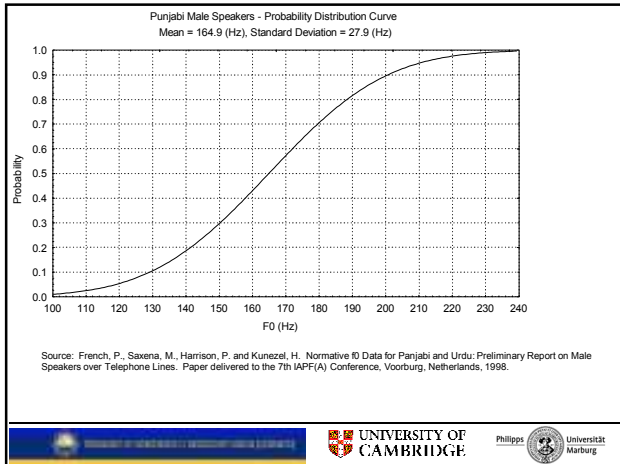
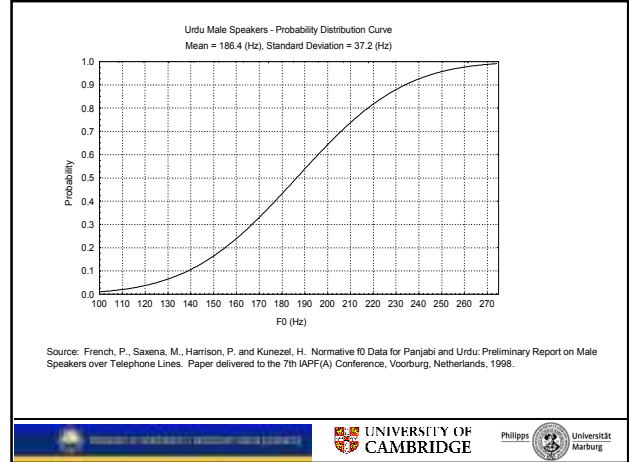
UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg



Thank you

<p>Sarmad Hussain Saad Irtza Toby Hudson Gea de Jong</p>	<p>sarmad.hussain@kics.edu.pk saad.irtza@kics.edu.pk toh22@cam.ac.uk gea.dejong@staff.uni-marburg.de</p>
-----------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------





Urdu language, member of the [Indo-Aryan](#) group within the [Indo-European](#) family of languages. Urdu is spoken by more than 100 million people, predominantly in [Pakistan](#) and [India](#). It is the official state language of [Pakistan](#) and is also officially recognized, or "scheduled," in the constitution of [India](#). Significant speech communities exist in the [United Arab Emirates](#), the [United Kingdom](#), and the [United States](#), as well. Notably, Urdu and [Hindi](#) are mutually intelligible.

Urdu developed in the 12th century ce from the regional [Apabhramsha](#) of northwestern India, serving as a linguistic *modus vivendi* after the Muslim conquest. Its first major poet was [Amir Khusrow](#) (1253–1325), who composed *dohas* (couplets), folksongs, and riddles in the newly formed speech, then called [Hindi](#). This mixed speech was variously called [Hindi](#), [Zaban-e-Hind](#), [Hindi](#), [Zaban-e-Delhi](#), [Rekhta](#), [Gujari](#), [Dakkhani](#), [Zaban-e-Urdu-e-Mualla](#), [Zaban-e-Urdu](#), or just Urdu, literally 'the language of the camp'. Major Urdu writers continued to refer to it as [Hindi](#) or [Hindi](#) until the beginning of the 19th century, although there is evidence that it was called [Hindustani](#) in the late 17th century ([Hindustani](#) now refers to a simplified speech form that is India's largest [lingua franca](#)).

Urdu is closely related to [Hindi](#), a language that originated and developed in the Indian subcontinent. They share the same Indic base and are so similar in [phonology](#) and [grammar](#) that they appear to be one language. In terms of lexicon, however, they have borrowed extensively from different sources—Urdu from [Arabic](#) and [Persian](#), Hindi from [Sanskrit](#)—so they are usually treated as independent languages. Their distinction is most marked in terms of writing systems: Urdu uses a modified form of [Perso-Arabic](#) script, while Hindi uses [Devanagari](#).



Phonologically, the Urdu sounds are the same as those of Hindi except for slight variations in short [vowel allophones](#). Urdu also retains a complete set of aspirated stops (sounds pronounced with a sudden release with an audible breath), a characteristic of Indo-Aryan, as well as [retroflex](#) stops. Urdu does not retain the complete range of [Perso-Arabic consonants](#), despite its heavy borrowing from that tradition. The largest number of sounds retained is among the spirants, a group of sounds uttered with a friction of breath against some part of the oral passage, in this case /l/, /z/, /zh/, /x/, and /g/. One sound in the stops category, the glottal /q/, has also been retained from [Perso-Arabic](#).

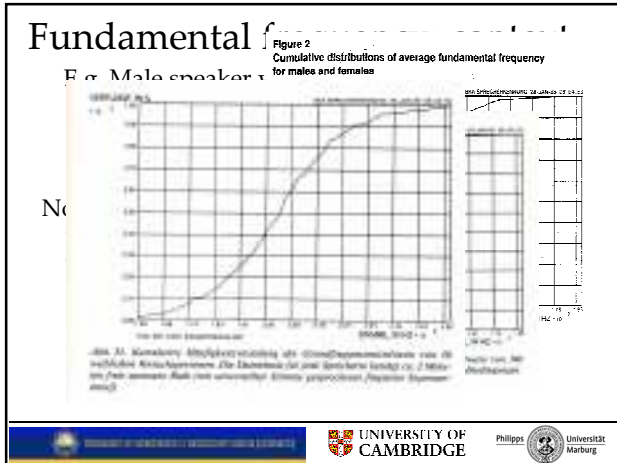
From the grammatical point of view, there is not much difference between Hindi and Urdu. One distinction is that Urdu uses more [Perso-Arabic](#) prefixes and suffixes than Hindi; examples include the prefixes *dar-* 'in', *ba-* 'with', *be-bila-* 'without', and *bad-* 'ill, miss' and the suffixes *-dar* 'holder', *-saz* 'maker' (as in *zinsaz* 'harness maker'), *-khor* 'eater' (as in *mutkhor* 'free eater'), and *-posh* 'cover' (as in *mez posh* 'table cover'). Although both Urdu and Hindi typically mark the plural by changing the singular suffix *-aa* to *-ee*, Urdu uses *-aat* in some cases, such as *kaagazaat* 'papers', *jawaharaat* 'jewels', and *makaanaat* 'houses'. In addition, where Hindi and Urdu both use the suffix *-ka* 'of' in many constructions, Urdu marks the genitive 'of' with *-e* (as in *subhe-azadi* 'the morning of freedom' and *khoon-e-ligar* 'the blood of heart').

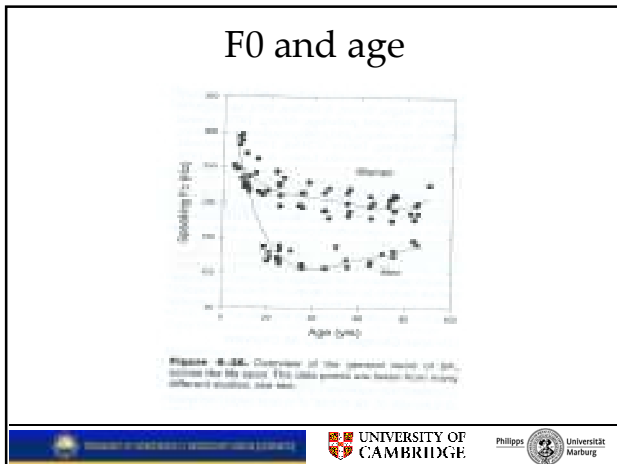
How to do a cumulative graph

- Excel- options
- Add-ins
- Select Analysis funtions > install
- Go back to the excel file of interest
- Daten
- Select Daten analyse > histogram



- ### Parameters affecting F0:
- Technical
 - Tape speed (Dejong&Hones 2002)
 - Sample size (Horii 1975: 14 Sec., French 1990 and Barry et al. 1991: 120sec.)
 - Physiological
 - Race (Hudson&Holbrook 1981)
 - Age (Hollien, Hollien & de Jong 1997, Reubold et al. 2010)
 - Language (Majewski et al. 1972, Scherer 1979, Braun 1992b, Hudson et al. 2009, Loakes 2006)
 - Smoking (Gilbert&Weismer 1974, Sorensen&Horii 1982, Braun 1994)
 - Intoxication Sobell et al. 1982, Pisoni/Martin 1989, Klingholz et al. 1988, Künzel et al. 1992, Hollien et al. 2001.
- Braun, A 1995
- UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg



Parameters affecting F0:

Technical factors:

Parameter	direction	quantity	reference
Tape speed:	+/-	any	
Register change (disguise)	+	app. 2x modal app. 2.5x model	Künzel (1987) Hollien/Michel (1968)
Electronic voice changers	+/-	any, often doubling or halving	Own observation A. Braun
Sample size	+/-	any	Steffen-Batalog et al. 1970: 50 sec. Mead 1974: 75 sec. Horii 1975: 14 sec. French 1990: 120 sec.

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

Parameters affecting F0:

Physiological factors:

Parameter	direction	quantity	reference
Race	Black < white	7Hz male 24Hz female	Hudson/Holbrook (1981)
Age			
Smoking	-	19 Hz (f) 11Hz (m), 4Hz (f)	Gilbert/Weismer 1974 Sorensen/Horii 1982
Alcohol	+	SD only 20% SD 100% SD Mean and SD	Sobell et al. 1982 Pisoni/Martin 1989 Klingholz et al. 1988 Künzel et al. 1992
Drugs-Testosterone	-	< 6 ST	Berendes 1962
Anabolic steroids	-		Damste 1964
Operations:			
Removal of cysts	+	9% (f)	Bouchayer/Cornut 1992
Removal of nodules	+	11%(f)	
Removal of polyps	+	14%(f) 4%(m)	
Shortening of vocal	+		Oates/Dacakis 1983

UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg

- ### Fundamental Frequency
- **Mean** = the sum of the frequency measurements divided by the number of waves measured.
 - **Median** = the fundamental frequency value that marks the 50th percentile of the distribution. Half of all values in the set are greater than the median, and half are smaller.
 - **Mode** = the value that occurs most often.
- UNIVERSITY OF CAMBRIDGE
Philipps Universität Marburg