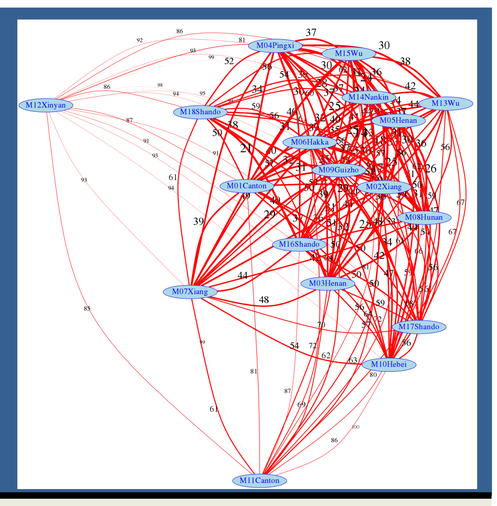
Variability in Mandarin Tone Perception

a multidialectal approach

Dafydd Gibbon[1], Huangmei Liu[2]

[1]Bielefeld University, Bielefeld, Germany & Jinan University, Guangzhou, China

[2]Tongji University, Shanghai, China & Fryske Academy, Leeuwarden, NL



ABSTRACT

As a preliminary to a larger scale dialect study, variability in the perception of the four Mandarin lexical tones by native speakers with different regional dialect backgrounds was examined.

In a novel sociophonetic survey of the ascription of pitch descriptors to tones, respondents rated the applicability of descriptors of pitch contour and height to recordings of tones on a 5-point Likert scale.

Each response contained metadata, with self-reported experience with regional varieties of Chinese.

Results showed differences in variability between pitch contour and pitch height descriptors, as well as some dependence between descriptor scores and regional dialect background, due to categorial tone perception.

A number of statistical and visualisation techniques were applied, including a set of hierarchical classifiers with dendrogram visualisation for comparison with dialect areas.

The significant results (analysis of variance, classification) indicate that the sociophonetic survey method tentatively fulfils its purpose and yields new results, but needs more data in the later more extensive study.

INDEX TERMS

Mandarin Chinese
multidialectal survey
categorial tone perception
pitch descriptor
regional variability
dialect
sociophonetics

OSCAR

wwwhomes.uni-bielefeld.de/gibbon/OSCAR/OSCAR_cmn01/

CONTACT

Dafydd Gibbon
<gibbon@uni-bielefeld.de>

Huangmei Liu
<lauraliu4321@163.com>

GOALS

Test new method
- multidialectal comparison
- contrast with previous bidialectal and bilingual comparisons
- exploratory rather than confirmatory

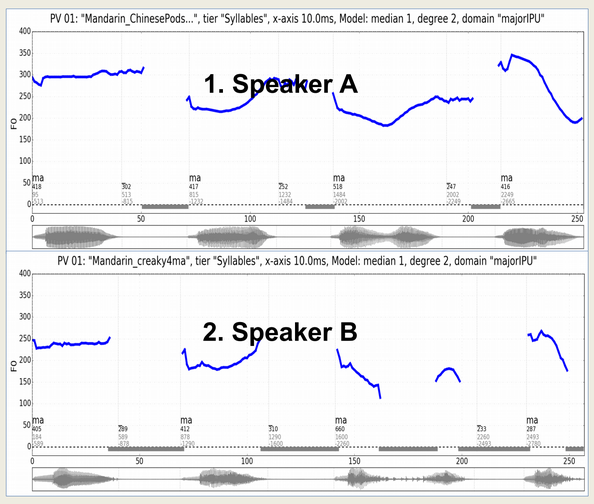
Sociophonetic focus
- assignment of descriptors of perceived pitch:
  - to standard Mandarin (Pǔtōnghuà) tones
  - by native speaker responders from different regions
- comparison of height and contour descriptors
- focus on inter-rater variability (contrast with inter-rater reliability)
- preparation for a large-scale multidialectal study
- longer-term goal of relating pitch descriptor assignments
  - to self-ascribed regional dialects
  - to linguistic dialect classification

TASK

Assign pitch descriptors to tones
- metalinguistic, cf. judgment paradigm of auditory phonetics and phonology
- sociophonetics, 'folk linguistic' opinions

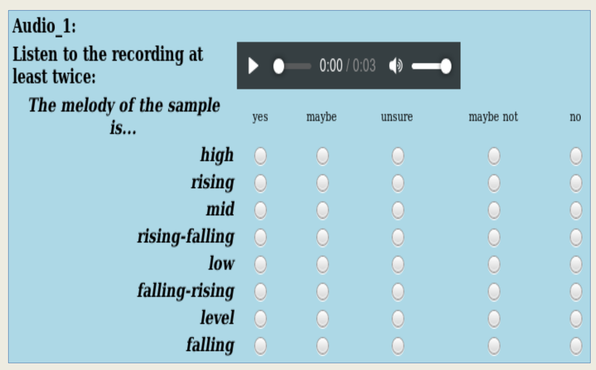
Custom online tool OSCAR
- input:
  - responder metadata:
    - age group, sex, L1, regional variety
  - single-page Likert format survey form
  - list of tones + pitch descriptor choices
- output:
  - for responders: confirmation
  - for experimenter: automatic evaluation

Stimuli
- 16 tone items (4 tones x 2 tokens each x 2 female speakers, standard Beijing Mandarin
- for all responders: same randomised token order, no adjacent repetitions

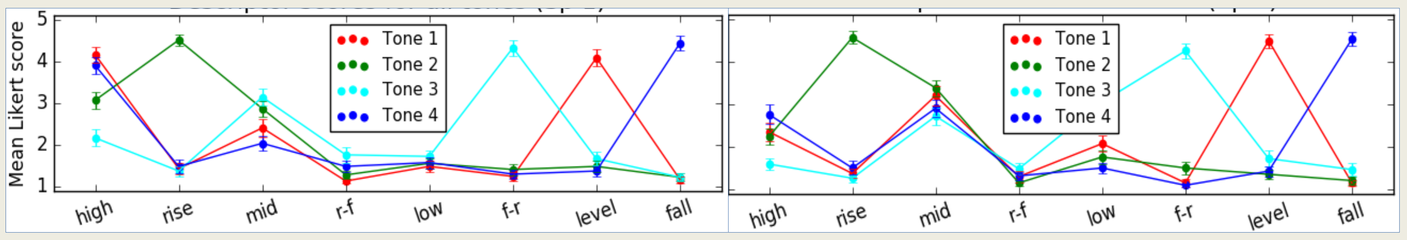


8 pitch descriptors
- contours: *level*, *rise*, *fall-rise*, *rise-fall* (distractor), *fall*
- heights: *high*, *mid*, *low*

5-point input scale
- *yes*, *maybe, not sure*, *maybe not*, *no*
- coded 5, 4, 3, 2, 1



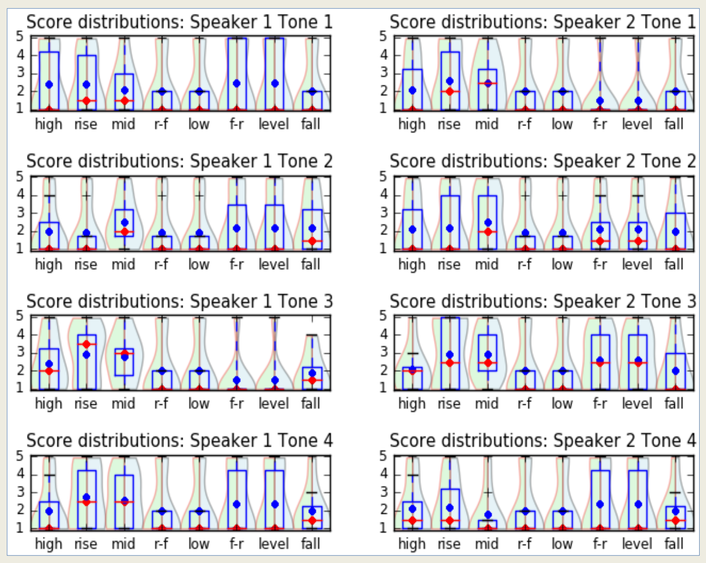
1. DISTRIBUTION



Descriptor scores per tone: Speaker A (left), Speaker B (right)

SpA	high	rise	mid	r-f	low	f-r	level	fall
T1	4.12	1.41	2.39	1.17	1.47	1.27	3.92	1.20
T2	2.95	4.52	2.74	1.27	1.52	1.41	1.45	1.24
T3	2.11	1.36	3.03	1.85	1.70	4.27	1.59	1.23
T4	3.79	1.45	1.97	1.53	1.64	1.32	1.39	4.52
SpB	**high**	**rise**	**mid**	**r-f**	**low**	**f-r**	**level**	**fall**
T1	2.32	1.36	3.29	1.36	1.97	1.21	4.24	1.18
T2	2.11	4.58	3.32	1.18	1.70	1.47	1.33	1.21
T3	1.56	1.27	2.71	1.59	3.00	4.14	1.70	1.44
T4	2.77	1.44	2.68	1.30	1.47	1.11	1.45	4.58

Descriptor distribution
- contour descriptors:
  - high density for high scores
  - Mandarin lexical tone system:
    Tone 1: *level;* Tone 2: *rise*
    Tone 3: *fall-rise*; Tone 4: *fall*
  - distractor tone: low score
    - *rise-fall*
- height descriptors
  - high density for low scores, but inconsistent
    - bimodal and/or very broad densities
    - *high*, *mid*, *low*



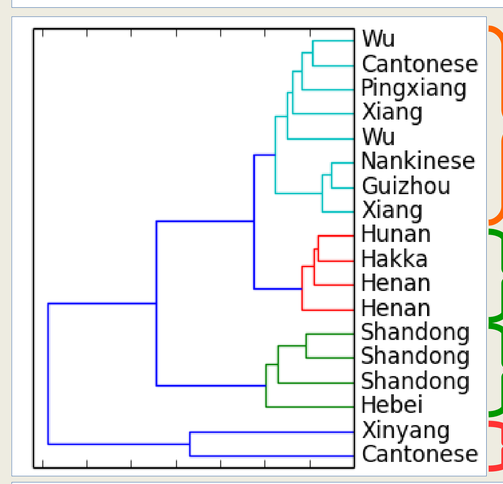
Inter-speaker variability
- some inter-speaker variability
  pitch height descriptor *mid* varies for tones but not so much for speakers
- *low* varies strongly for speaker B but not for A
- higher score of *high* for Speaker A reflects her overall higher pitch

MANOVA
- fixed factors:
  - *tone type*, *pitch descriptor*, *dialect background*, *speaker*, with interactions
- significant effects
  - *dialect region, pitch descriptor*
- strong interactions
  - *tone + descriptor, speaker + descriptor*
  - *dialect + tone + shape* (multiinteraction)

Factors	Df	Sum Sq	Mean Sq	F	p
dial	16	480	30	12.966	<0.001
descr.	1	123	123.19	53.252	<0.001
tone:descr.	1	130	130.39	56.365	<0.001
sp.:descr	1	38	38.26	16.54	<.0001
dial:tone:descr	16	89	5.58	2.413	<0.01

2. RESPONSE CLUSTERS

Pearson Distance classification
- 7 clustering algorithms compared
  - comparison with geographical areas
  → **Ward incremental variance minimisation**
    - Shandong+Hebei: geogr. close, N. Mandarin
    - Hunan, Hakka, Henan: geog. close, hist. rel.
    - others: prosodic typology partly plausible, geographically and historically less plausible
- noise due to
  - small data set with large number of classes
  - inaccuracies in self-ascription
  - normative element in self-ascription:
    - responders are language graduates
    - strong influence of standard Mandarin



Geographically and/or historically distant pairs:
WU-Cantonese, Nankinese-Guizhou, Henan-Hebei-Wu, Henan-Hebei, Xiang-Wu, Xinyang-Cantonese

CONCLUSIONS

Main descriptive outcomes
- expected canonical status of contour vs. height descriptors confirmed: canonical descriptors assigned more consistently than non-canonical descriptors
- significant differences for factors *dialect* and *descriptor*
- significant interactions for *tone* + *descriptor*, *speaker* + *descriptor*, *dialect* + *tone* + *descriptor*
- partly plausible classification results
- despite small dataset – but more data needed

Main strategic outcome

The novel method is fit for purpose as a baseline for current planning of a larger dialect survey using more complex contextual data such as tone sandhi, accentuation, intonation, with larger numbers of speakers of each dialect.