Parallel adjustment of phonetic targets in L2 English voice onset time

Eleanor Chodroff[1] and Melissa Baese-Berk[2]
[1]Northwestern University, Department of Linguistics
[2]University of Oregon, Department of Linguistics

Northwestern

UNIVERSITY OF OREGON

Introduction

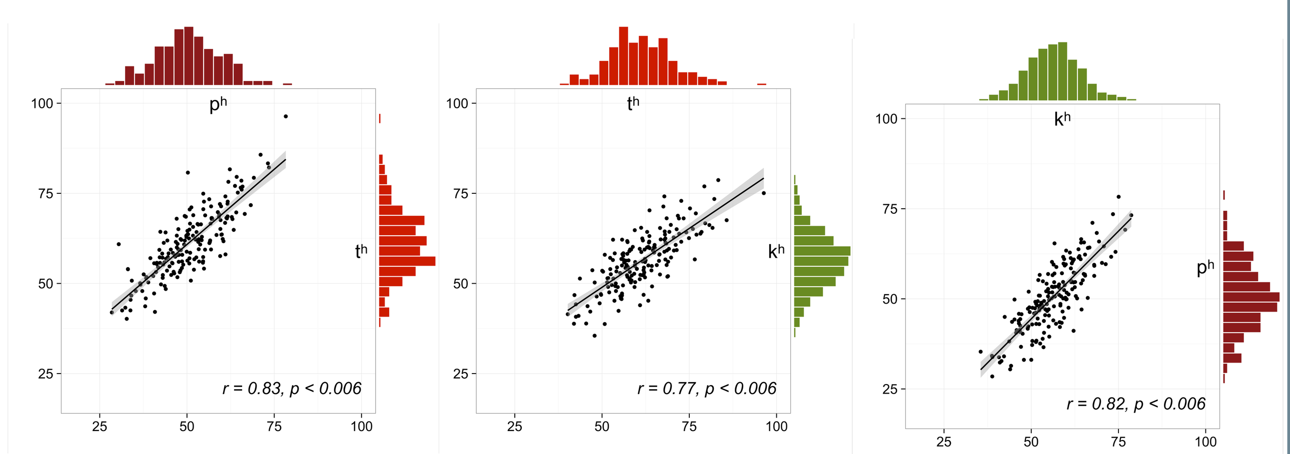
L2 speech has typically been considered more variable in its phonetic realization than L1 speech (e.g., Flege, Takagi, & Mann 1995; cf. Vaughn, Baese-Berk, & Idemaru 2018).

In learning a new phonetic category, a speaker may have:

- Uncertainty in the targets
- Uncertainty in the implementation

Previous research has found constraints on permissible variation *between* speech sounds within a natural class in L1 speech.

Talker mean VOTs of [pʰ tʰ kʰ] strongly covary with one another in L1 American English, indicating systematic relationships of VOT within the natural class (Chodroff & Wilson 2017).



Data from 180 American English speakers in the Mixer 6 Corpus, Chodroff & Wilson 2017

VOT covariation has also been observed across over 100 languages, but in L1 speech only (Chodroff, Golden, & Wilson under review).

Given the increased uncertainty in L2 representations, it seems plausible that these structured relations in VOT may break down in L2 speech.

Do L2 English speakers maintain structured relations in VOT among the voiceless stop consonants?

Does VOT covariation arise from the use of L1 phonetic targets or from a parallel shift in phonetic targets?

ALLSSTAR Corpus

Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings

Connected speech tasks in L1 and L2
Declaration of Human Rights: 20 sentences
HINT 1: 60 sentences
HINT 2: 60 sentences
Le Petit Prince: 30 sentences
The North Wind and the Sun passage

140 speakers from Northwestern University (86 M, 54 F)
114 bilingual speakers
26 monolingual English speakers

L2 English VOT Covariation

Do L2 English speakers maintain structured relations in VOT among the voiceless stop consonants?

Methods

Employed **English** read speech from L1 and L2 speakers
Forced phonetic alignment using FAVE
Automatic VOT alignment using AutoVOT

Represented L1 | Number of speakers | Presence of aspiration in L1

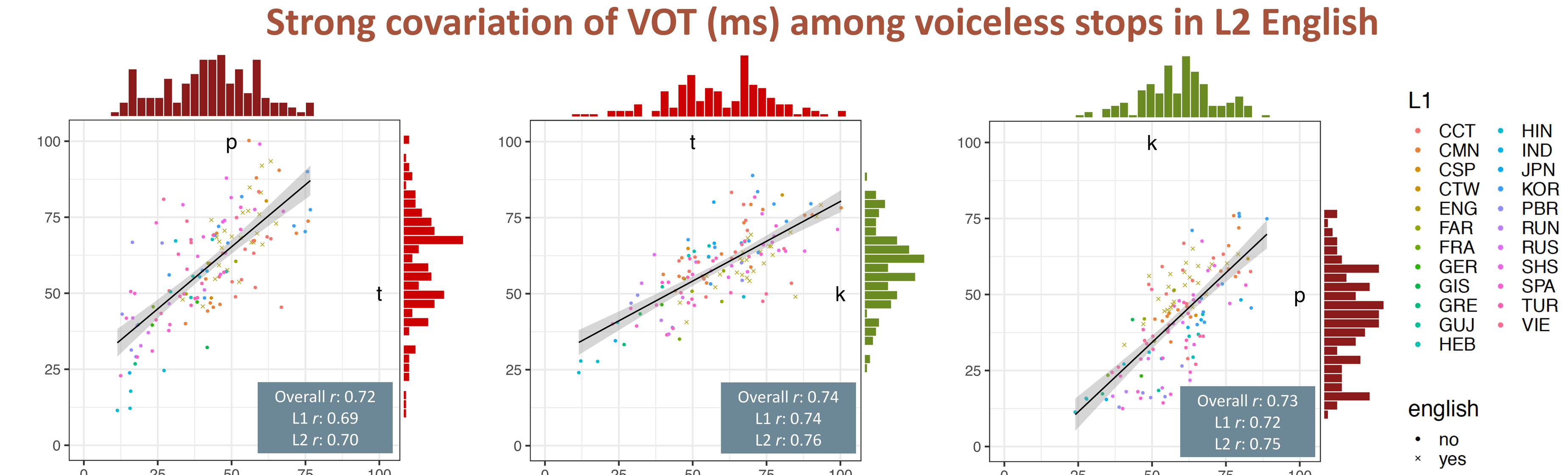
Cantonese	14	✓	Hebrew	4	~	Portuguese (Brazilian)	5	
English (American)	26	✓	Hindi	5		Runyankore	1	
Farsi	3		Indonesian	1		Russian	5	
French	1		Japanese	3	~	Spanish	11	
German	2	✓	Korean	11	✓	Spanish (US Heritage)	12	
Gishu	1		Mandarin (China)	14	✓	Turkish	13	✓
Greek	1		Mandarin (Singapore)	1	✓	Vietnamese	4	[tʰ]
Gujarati	1	✓	Mandarin (Taiwan)	1	✓			

Each L1 has voiceless /ptk/ in native phonological inventory

Results

Total # stops: ~3800 /p/, 3500 /t/, 6200 /k/ = ~13,500 | **Per speaker:** ~27 /p/, 25 /t/, 45 /k/ = ~97

Strong covariation of VOT (ms) among voiceless stops in L2 English



Range and median of speaker means (ms)	/p/	/t/	/k/
L1 English	33 – 65; 53	46 – 93; 69	41 – 79; 60
L2 English	11 – 77; 41	11 – 100; 56	24 – 89; 60
Overall	11 – 77; 44	11 – 100; 60	24 – 89; 60

Range and median of speaker SDs (ms)	/p/	/t/	/k/
L1 English	14 – 34; 22	7 – 35; 19	14 – 35; 20
L2 English	2 – 40; 21	3 – 35; 20	6 – 40; 20
Overall	2 – 40; 22	3 – 35; 20	6 – 40; 20

Speaker mean and SD fairly correlated (*rs* between 0.50 and 0.80)

L1 to L2 VOT Covariation

Does VOT covariation arise from the use of L1 phonetic targets or from a parallel shift in phonetic targets?

Methods

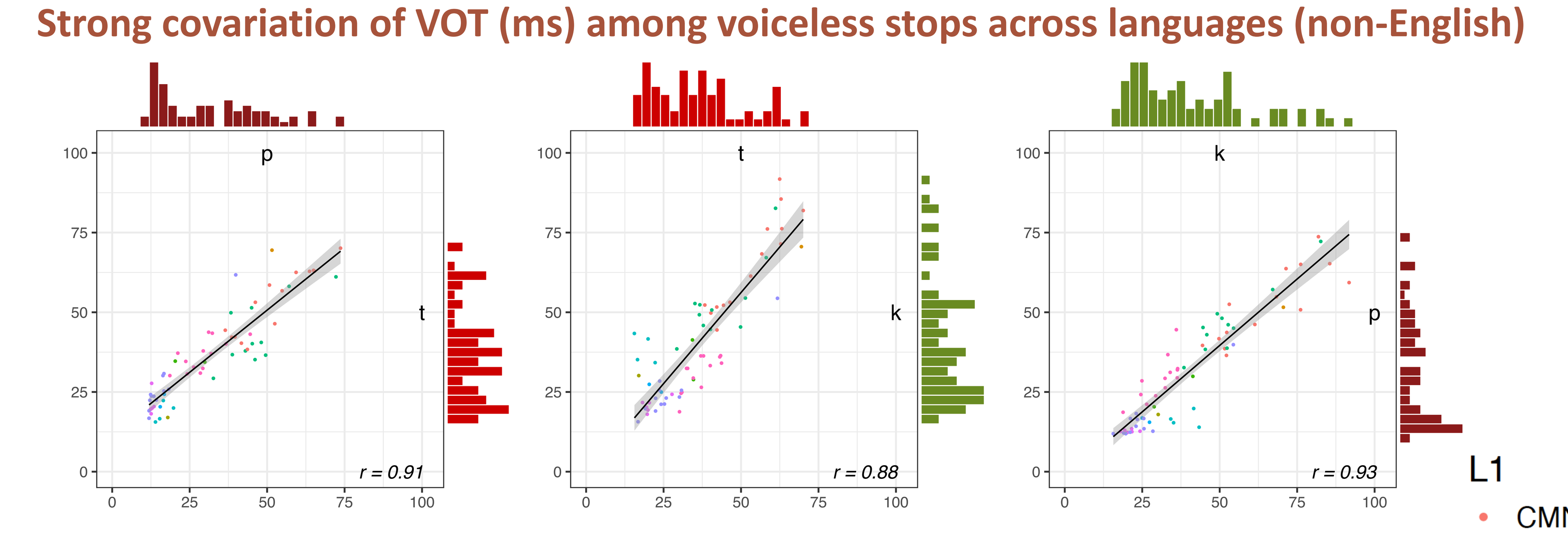
Employed **non-English L1** read speech from languages that could be force aligned
Forced phonetic alignment using Montreal Forced Aligner with pre-trained acoustic and grapheme-to-phoneme models

French | German | Korean | Mandarin | Portuguese (Brazilian) | Russian
Spanish | Spanish (US Heritage) | Turkish

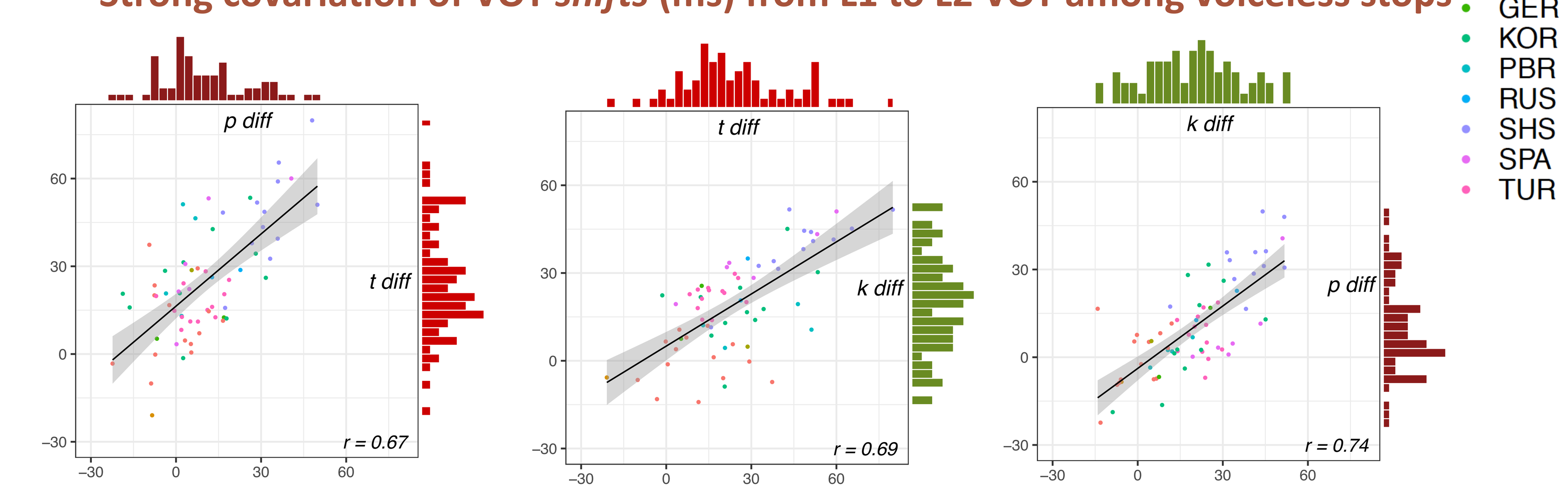
Automatic VOT alignment using AutoVOT

Results

Strong covariation of VOT (ms) among voiceless stops across languages (non-English)



Strong covariation of VOT *shifts* (ms) from L1 to L2 VOT among voiceless stops



The phonetic targets underlying VOT for /p/, /t/, and /k/ may differ from the L1 to L2 grammar, but the relationship *among* those segments is approximately the same

Discussion

Summary and implications

Strong linear relationships between VOT means of /p/, /t/, and /k/ in L2 English

Some representation of natural class in L2 grammar: phonetic targets underlying VOT for /p t k/ shift in parallel (rarely the case that an individual acquires a more English-like VOT for /k/, but not for /p/ and /t/

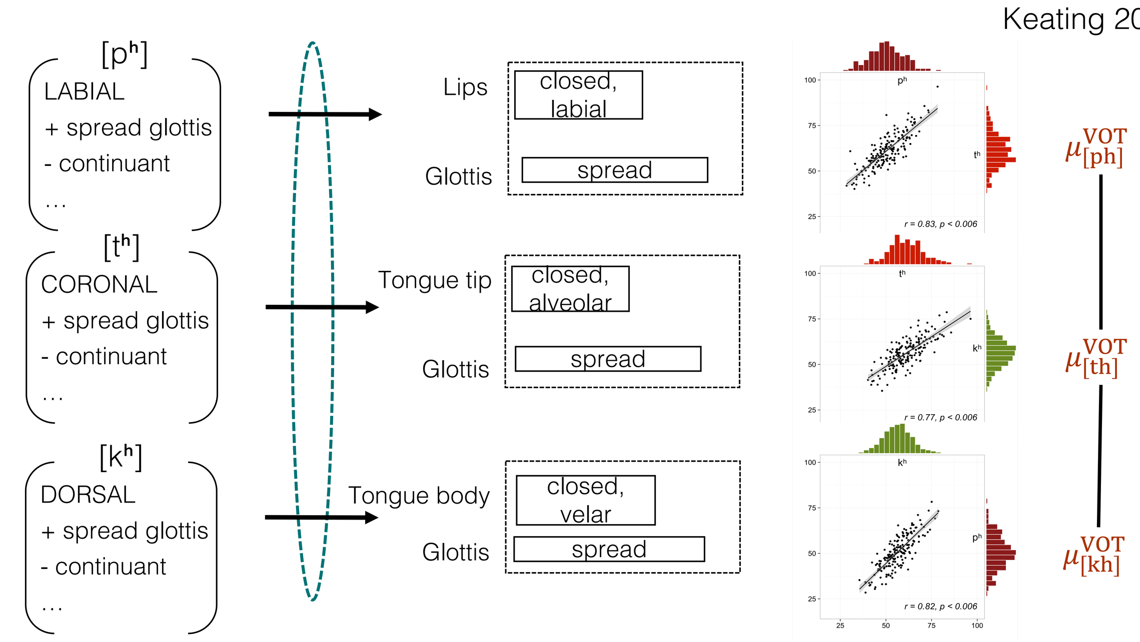
Need to further investigate cases when shifts are not entirely parallel (e.g., Spanish /p/)

Acknowledgments
We would like to thank Huanyang Xu and Kayla Walker for their help in data processing. We also thank Ann Bradlow, Chun-Liang Chen, Yasmin Elmasri, Jordan Hosier, and Linda Nwumeh for their support in data processing and management. Additional thanks go to the ASA Strategic Plan Goal 2 Task Force for supporting the first author's travel to this conference through an ASA Early Career Travel Subsidy.

What gives rise to covariation?

Structure in the output indicates structure in the input

Keating 2003



Same vocal tract but physically and theoretically possible for speakers to produce the following patterns:

$\mu^{VOT}_{[ph]}$ = 80 ms and $\mu^{VOT}_{[kh]}$ = 40 ms
$\mu^{VOT}_{[ph]}$ = 40 ms and $\mu^{VOT}_{[kh]}$ = 80 ms

A speaker can have two distinct phonetic targets for /p t k/ (see above data)

Principle of uniformity

- Mapping from phonological feature value to corresponding set of phonetic targets must be uniform for all segments with that feature value
- Phonetic targets underlying VOT may be articulatory in nature
- Covariation arises from underlying (near-)identity in targets for /p t k/
- Applies to L1 and L2 grammars