University of Glasgow

Perceiving phonetic variation in noise: structure, speaker, accent

Rachel Smith
(University of Glasgow)

University of Glasgow

Collaborators

Rachel Baker

Sarah Hawkins

Rachael-Anne Knight

Michèle Pettinato

Sophie Holmes-Elliott

University of Glasgow

Purpose

To discuss what speech-in-noise studies can tell us about the contribution of **fine phonetic detail (FPD)** to word recognition

3 studies, discussing FPD relating to

- linguistic (grammatical) **structure**
- individual **speaker**
- regional **accent**

University of Glasgow

Background

1. Fine phonetic detail (FPD)

"phonetic phenomena that are systematically distributed according to linguistic/communicative function, but not systematically treated in conventional accounts"

"phonetic information that affects people's responses but is not a primary cue to phonological form of lexical items"

Hawkins (2008)

Thus:	not cues distinguishing	/pa/ from /ba/
	but cues distinguishing	/p/ in *potato* from /p/ in *important*
		/p/ in *displease* from /p/ in *displays*
		/p/ in *it's a ta*[p^h] from *it's a ta*[p']

2. FPD in SPiN

- **Natural/coherent** patterns of FPD improve SPiN

  e.g. *r*-resonances in English

  Hawkins & Slater (1994), Tunley (1999), West (1999), Heinrich, Flory & Hawkins (2010)

- **Familiar** patterns of FPD improve SPiN

  e.g. a familiar voice

  Nygaard, Sommers & Pisoni (1994), Nygaard & Pisoni (1998)

How? FPD increases processing efficiency by giving richer and more redundant cues to structure

Expt 1: SPiN of FPD reflecting grammatical **structure**

Baker (2008)

Patterns of FPD differ between the same phoneme strings when they form a function word (F), e.g. *she's,* vs. part of a content word (C)

Manuel (1992, 1995); Lavoie (2002); Local (2003)

e.g.	**F**	**C**
/ʃiːz/	*she's*	*banshees*
/aɪm/	*I'm*	*time*
/jə/	*you're*	*Yosemite*
/ɔlːə/	*all the*	*all Letitia*

University of Glasgow

Expt 1 Method

20 sentence pairs. Within pairs, sentences matched for foot structure (rhythm), and for segments as far as possible.

Recorded in thematic paragraph context and casual style by 1 female SSBE speaker

Critical phoneme strings x-spliced, to *match* or *mismatch* their context

Cafeteria noise added (SNR +3 dB, based on piloting)

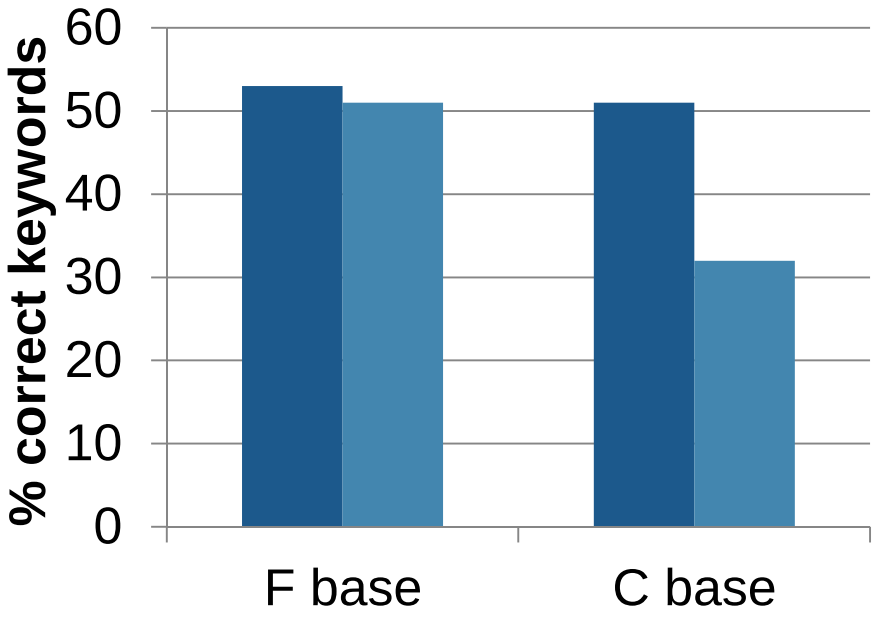
	***Matched* splice**	***Mismatched* splice**
Function word base	**FFF** The girl saw the man /ʃiːz/$_{\text{she's}}$ in love with	**FCF** The girl saw the man /ʃiːz/$_{\text{banshees}}$ in love with
Content word base	**CCC** The girl saw the ban/ʃiːz/$_{\text{banshees}}$ in London	**CFC** The girl saw the ban/ʃiːz/$_{\text{she's}}$ in London

University of Glasgow

Expt 1 Results

Participants typed what they heard; responses scored for key F or C word correct, and non-keywords correct

Analysis: mixed-effects logistic regression



Splicing an F word into a C base disrupts intelligibility in noise; splicing a C word into an F base does not.

Expt 2: SPiN of FPD reflecting **speaker** idiosyncrasy

Patterns of FPD differ between the same phoneme strings depending on the location of word boundaries within them

e.g. *cat size – cat's eyes*

she dyed them – she'd eyed them

sly stroll – sliced roll

la mie (de pain) – l'amie

Lehiste (1960), Hoard (1966)

Expt 2: SPiN of FPD reflecting **speaker** idiosyncrasy

Acoustic cues involved depend on the segmental string, but include

- duration of word-initial vs non-initial consonants
- allophonic cues e.g. aspiration, flapping, /l/-darkness
- intensity
- spectral balance
- vowel quality
- voice quality

Most forced-choice experiments show better than chance identification (~ 60-90% accuracy)

Lehiste (1960), Hoard (1966), Oller (1973), Umeda & Coker (1975), Pierrehumbert & Talkin (1992), Fougeron (2001) and many more

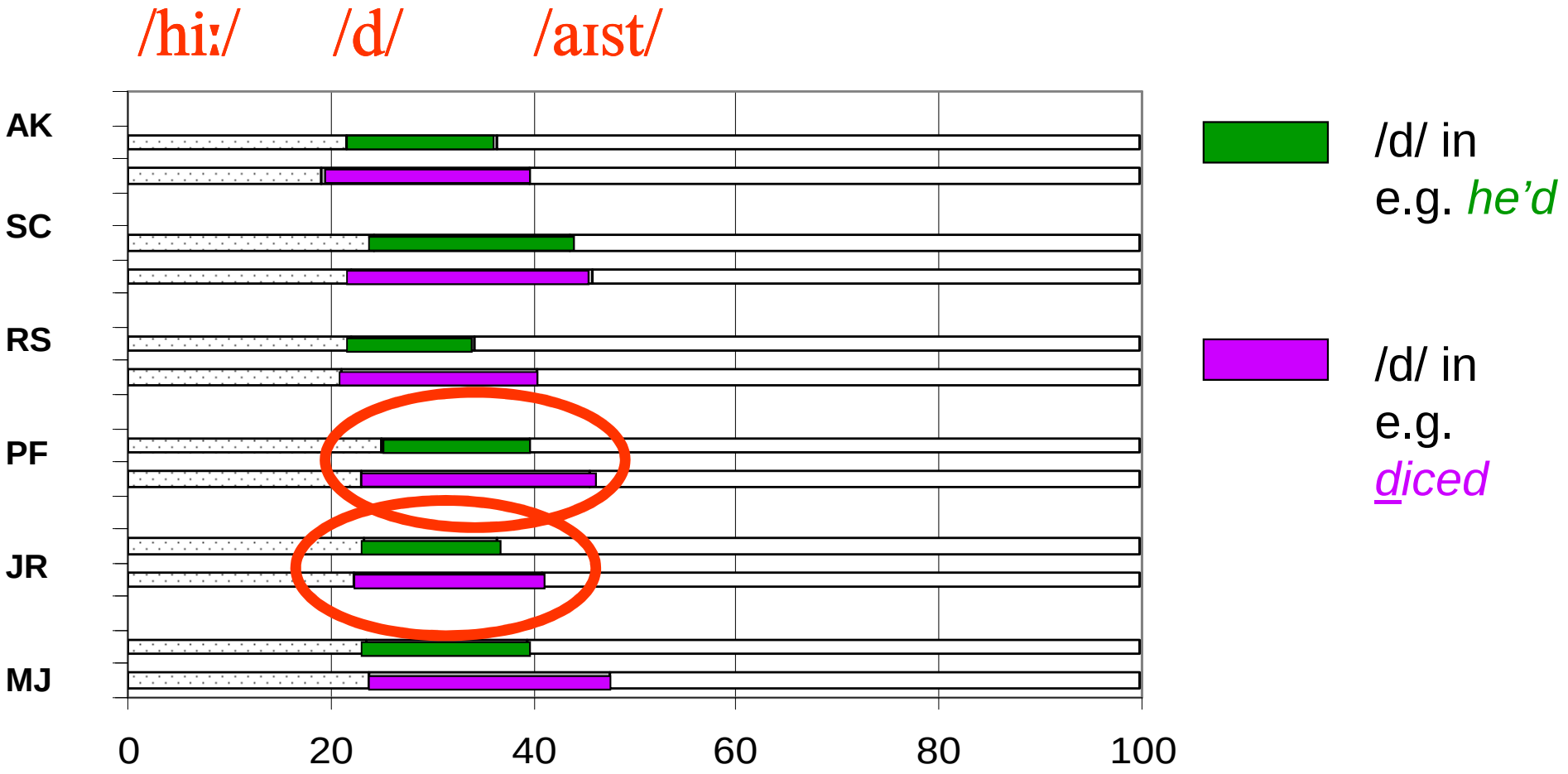
Expt 2: SPiN of FPD reflecting **speaker** idiosyncrasy

Individual speakers vary in their use of word-boundary cues

→ **Listeners' experience with individual voices may affect how they exploit this variation in SPiN tasks.**

University of Glasgow

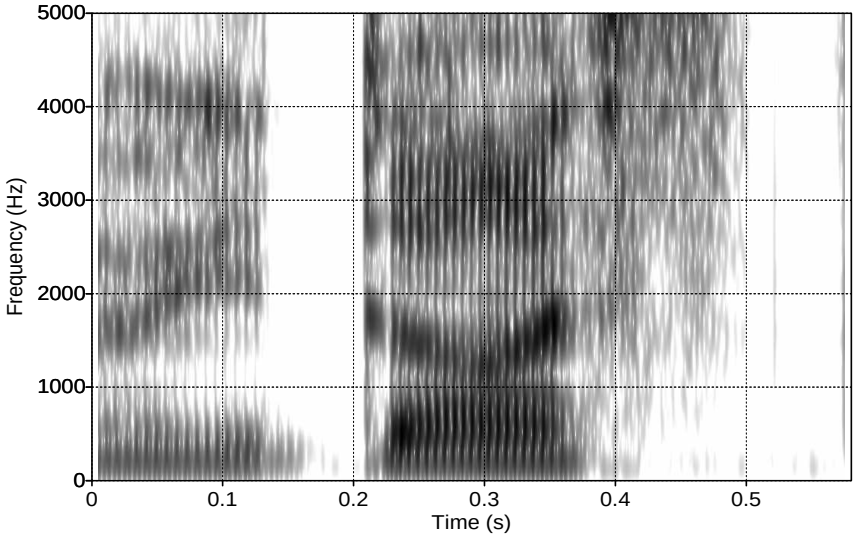
significant inter-spk differences in *durational contrasts*



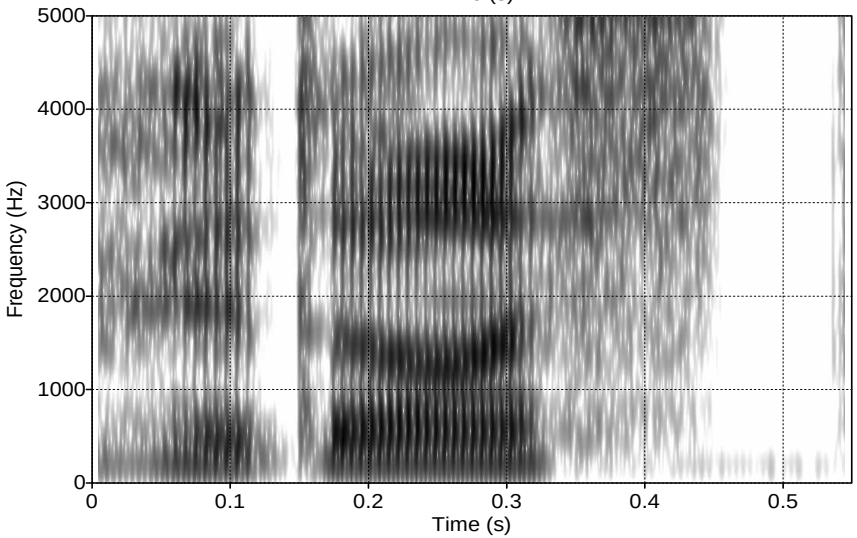
6 speakers of Standard Southern British English (Smith & Hawkins 2012)

significant inter-spk differences in realisation of *word-initial vs –final /d/*

MJ: /d/

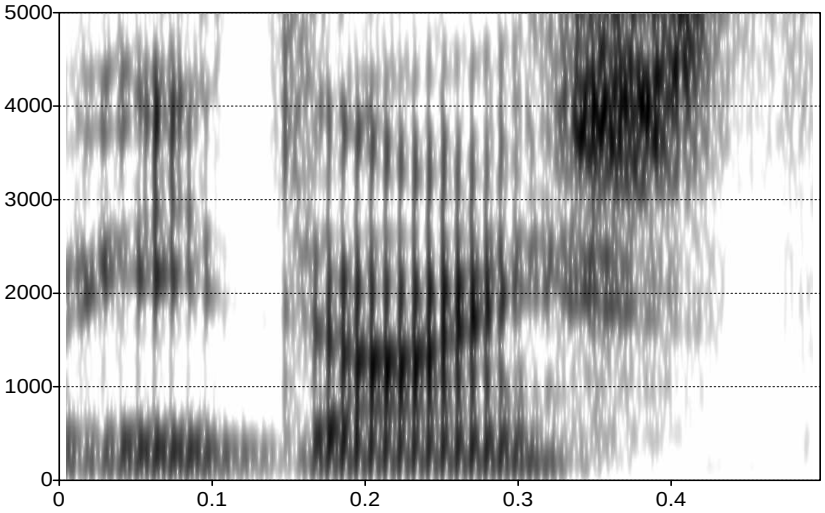


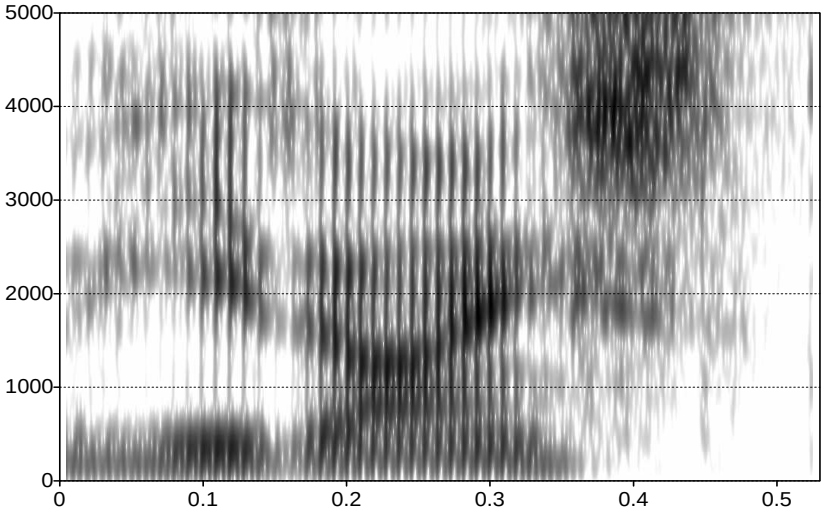
he diced



he'd iced

PF: /d/

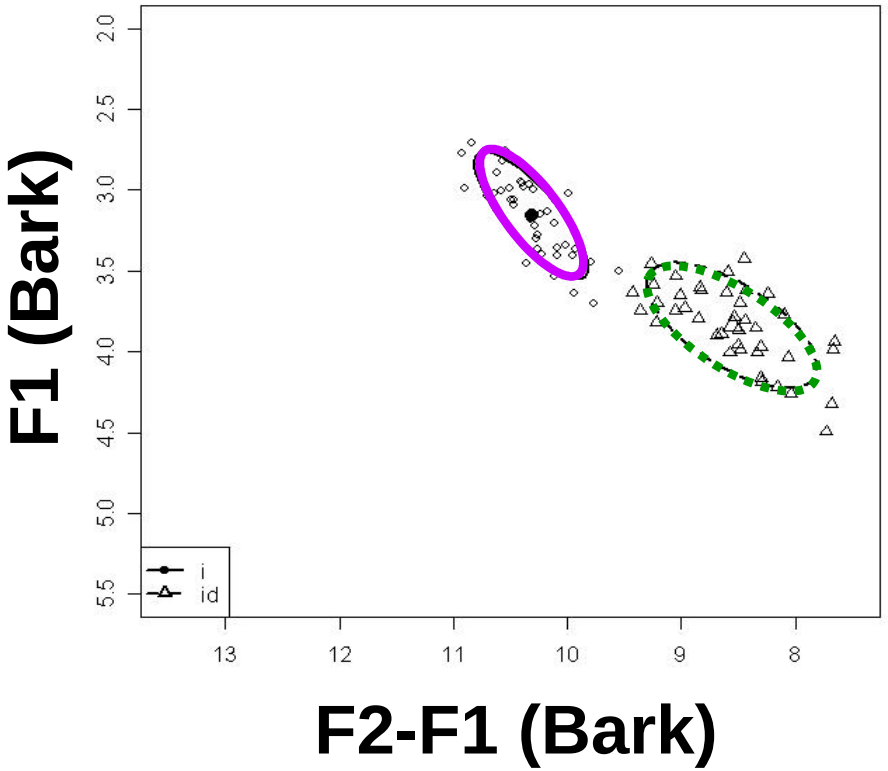




University of Glasgow

significant inter-spk differences in /iː/ F2–F1

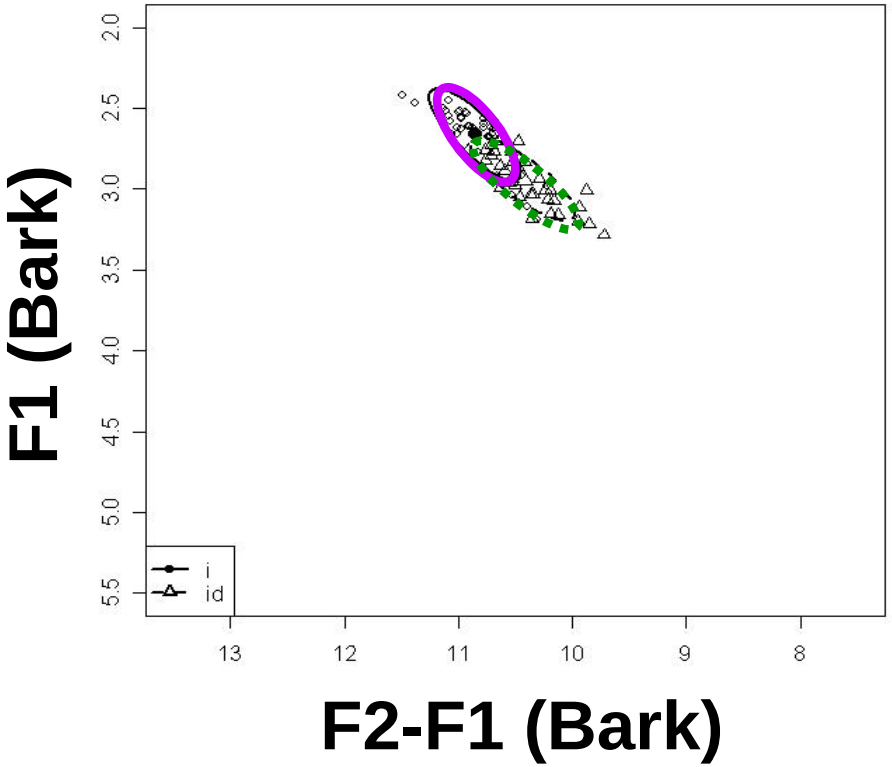
MJ: /iː/



he diced

he'd iced

PF: /iː/



University of Glasgow

Expt 2 Method

Smith & Hawkins (2012)

24 phonemically-identical sentence pairs, e.g.:

So he diced them — So he'd iced them

Other examples:

But Pat sawed them — But Pat's awed them
It's no wonder he didn't recognise that salute — that's a lute
They also offer Mick stability — mixed ability

University of Glasgow

Expt 2 Method

2 M speakers of Standard Southern British English

read sentences 8x in:

casual style

disambiguating contexts

Contexts:

a) He wanted the carrots to cook fast.
So he diced them.

b) The top of the cakes had come out looking uneven. So he'd iced them.

University of Glasgow

Expt 2 Method

Pre- and post-test: SPiN (25 min each)

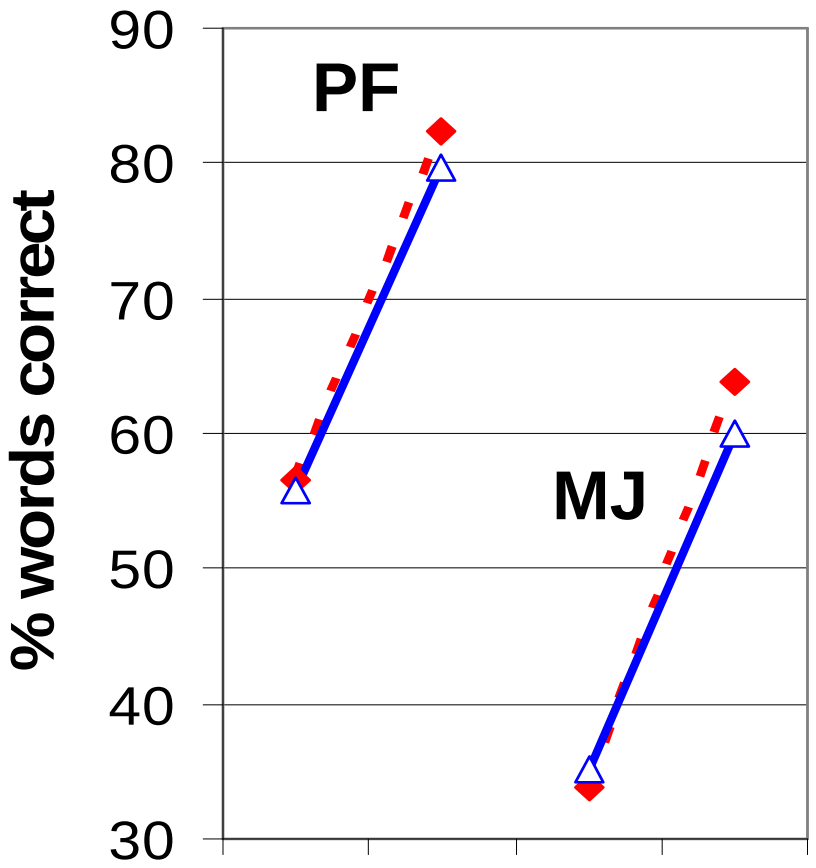
- **sentences presented in cafeteria noise (SNR +2 dB)**
- **task: type in what is heard**

Training: (40 min)

- **sentences presented in disambiguating contexts**
- **no noise**
- **task: answer questions about meaning**
- **training voice Same or Different as test voice**

University of Glasgow

Expt 2 results: % words correct



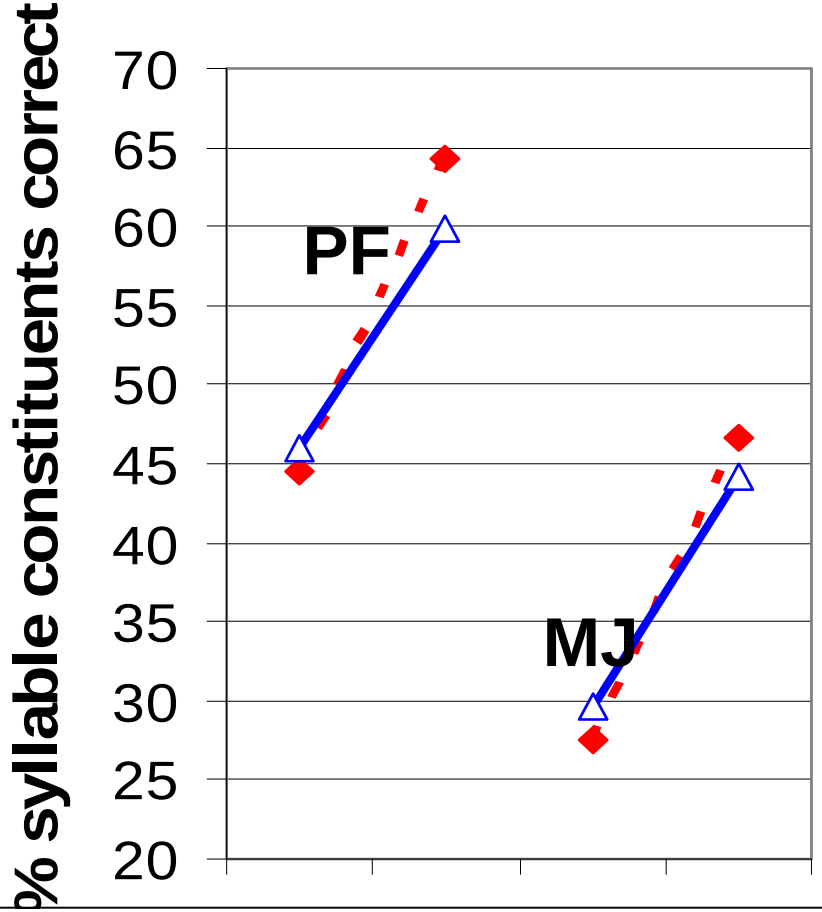
Lines show improvement from pre-test to post-test

Same Voice

Different Voice

Same vs. Different:
$p < 0.0001$, independent of talker

Expt 2 results: syllable constituents at Word 1 End*



Lines show improvement from pre-test to post-test

Same Voice

Different Voice

Same vs. Different:
$p < 0.025$, independent of talker

*e.g. for *he:* word-final /i:/ ✓

he, we, she, tea ✓

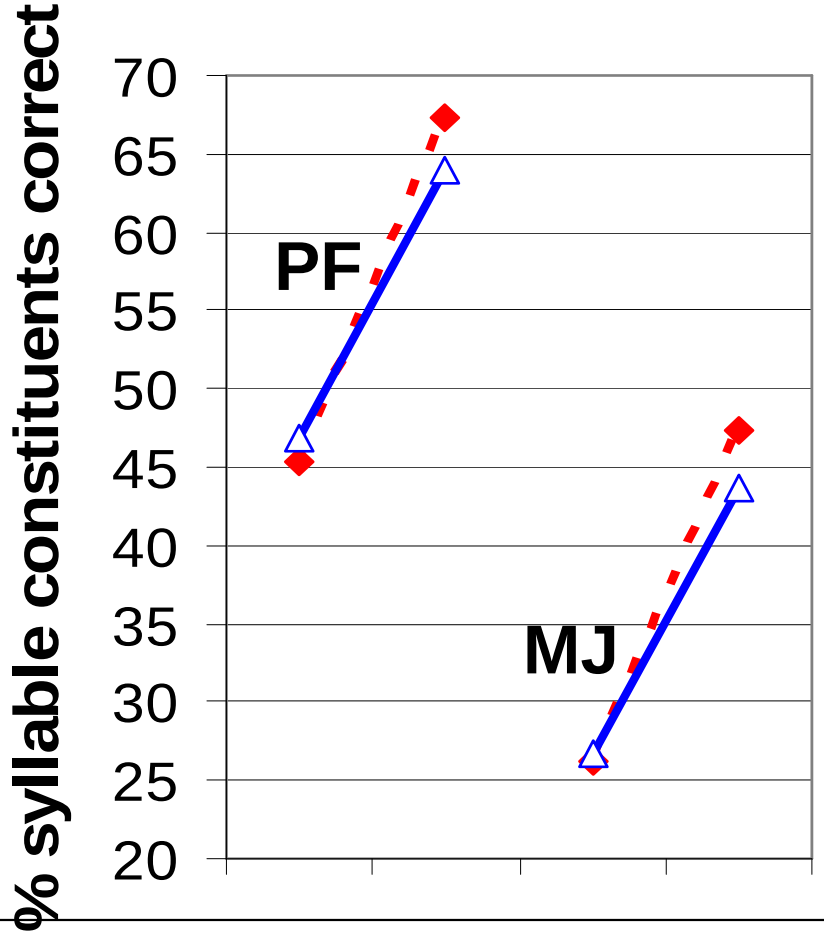
he'd, weed, eat etc ×

for *he'd*: word-final /d/ ✓

he'd, we'd, weed, stampede ✓

he, heat, freeze, heel etc ×

Expt 2 results: syllable constituents at Word 2 Start*



Lines show improvement from pre-test to post-test

Same Voice

Different Voice

Same vs. Different:
$p < 0.05$, independent of talker

*e.g. for *diced:* word-initial /d/ ✓

diced, died, darcy ✓

iced, enticed, guy etc ×

for *iced*: word-initial /aɪ/ ✓

iced, eye, icecream, I ✓

diced, side, asked, etc ×

University of Glasgow

Expt 2 Summary

- Talkers vary in patterns of allophonic detail at word boundaries.
- Familiarity with these patterns helps listeners segment and identify words in non-stationary noise.
- The perceptual benefit is small, but robust, and obtained using natural materials.

Expt 3: SPiN of FPD reflecting regional **accent**

Speech in a regional accent that is not one's own tends to be more difficult to process

Labov & Ash, 1997; Clopper & Bradlow, 2008; Adank et al., 2009; Floccia et al., 2006, 2009

Exception: when the listener has a nonstandard regional accent, and the target speech is in a standard accent, familiar from media.

Glasgow English (GE) is highly unintelligible to listeners of SSBE, but GE listeners show no impairment listening to SSBE.

Adank et al., 2009

Or do they?

NB Trainspotting actually set in Edinburgh! but gives an idea of a similar Scottish vernacular

Expt 3: SPiN of FPD reflecting regional **accent**

Smith, Holmes-Elliott, Pettinato & Knight (2013)

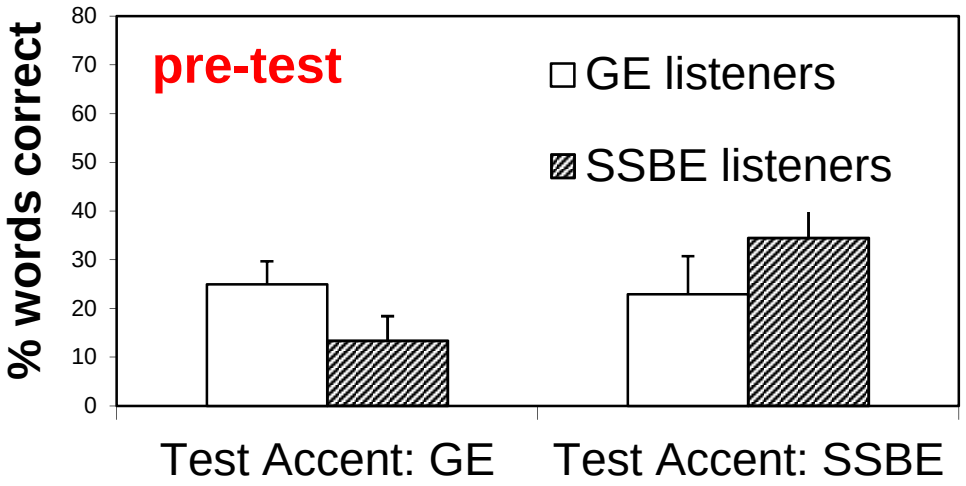
Trained 4 groups x 39 listeners on versions of materials from Smith & Hawkins (2012)

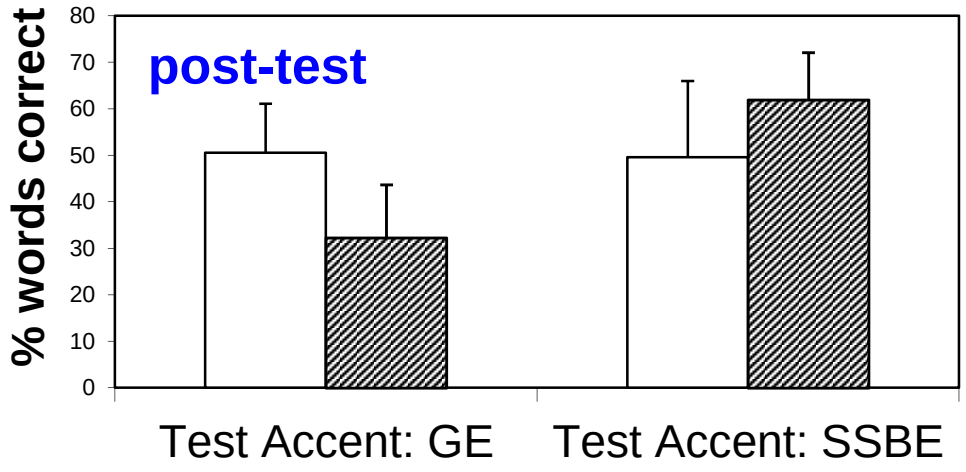
- GG (Glasgow listeners, Glasgow talker)
- GS (Glasgow listeners, SSBE talker)
- SG (SSBE listeners, Glasgow talker)
- SS (SSBE listeners, SSBE talker) (from previous study)

Pre- and post-test: always same dialect as training

identical vs novel talker

Expt 3 Results





word intelligibility

SS > GS and GG > SG; SS>SG, but GG=GS

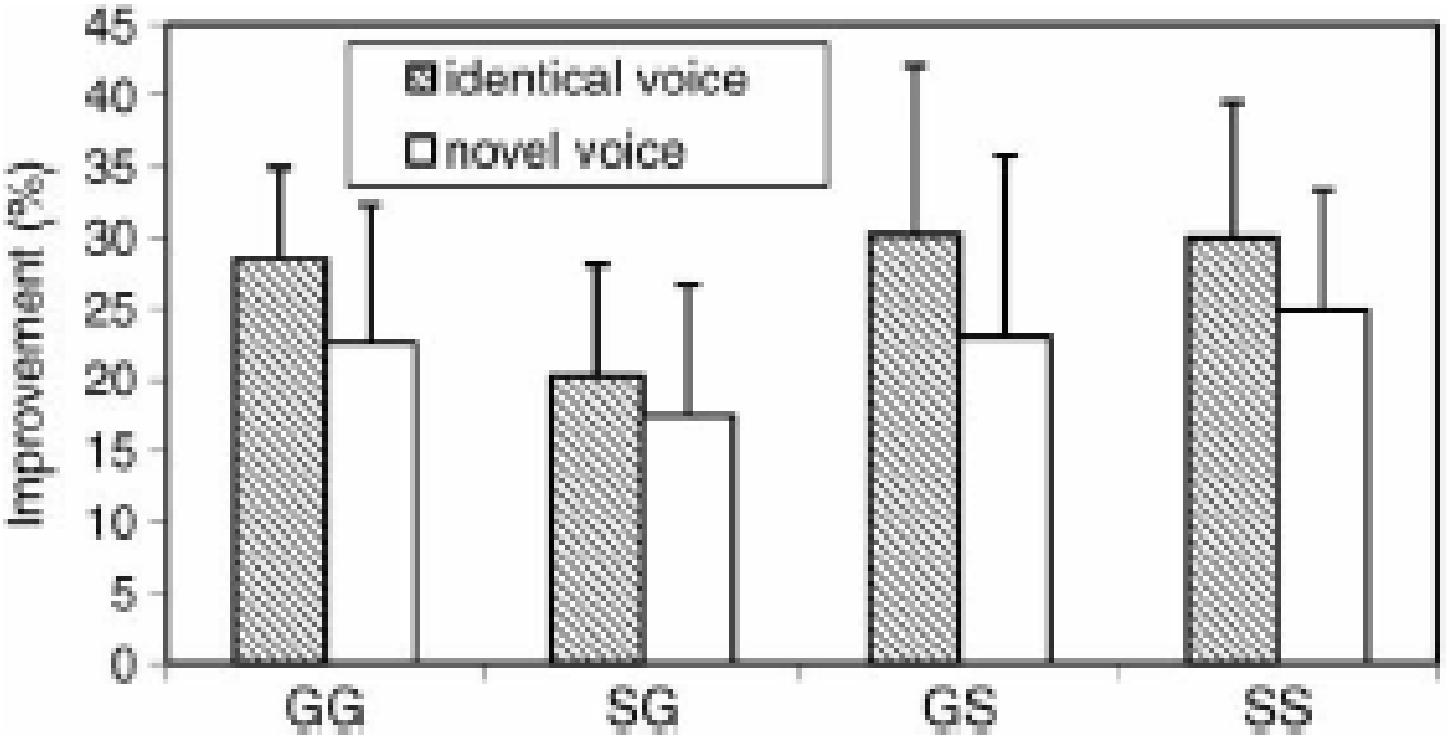
Glaswegian listeners perform less well than SSBE listeners do with SSBE – but better than SSBE listeners hearing GE

Partially supports, partially contradicts previous findings.

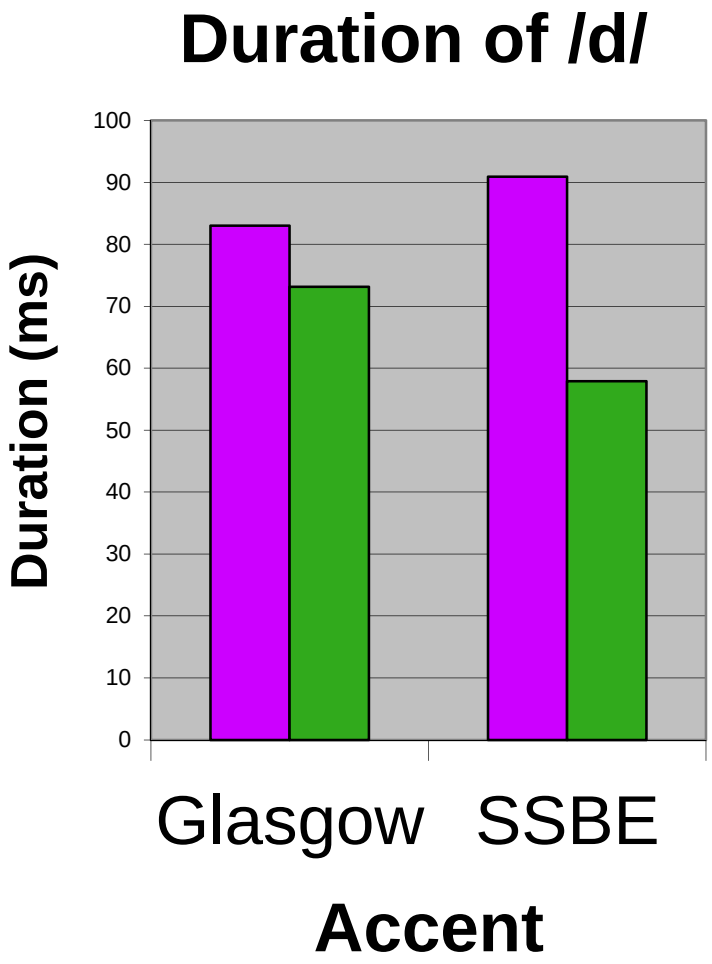
University of Glasgow

effects of speaker familiarity according to dialect familiarity

speaker familiarity benefit found for all groups, regardless of listener's or speaker's dialect



Work in progress: inter-accent differences in durational contrasts



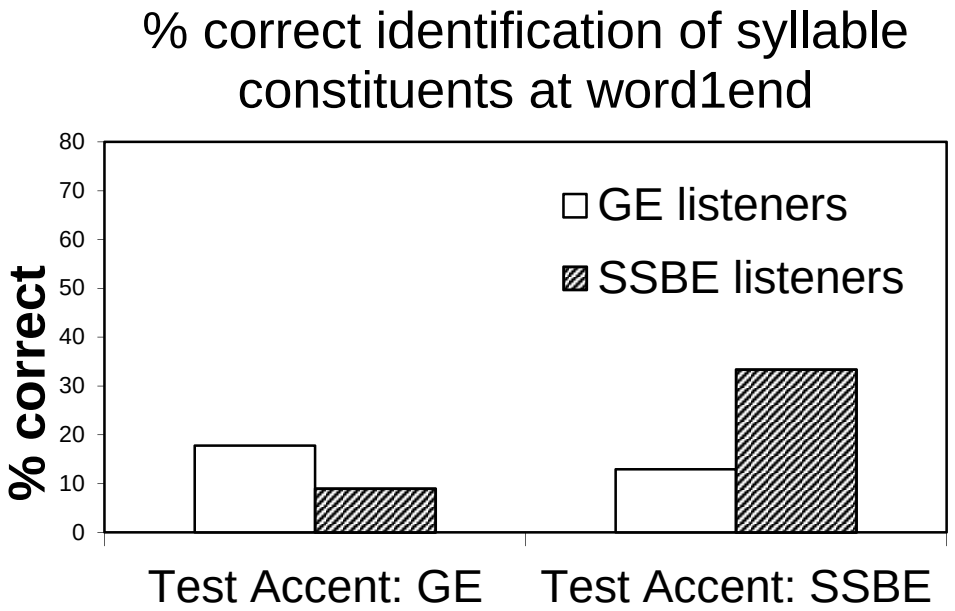
Initial /d/
e.g. *he diced*

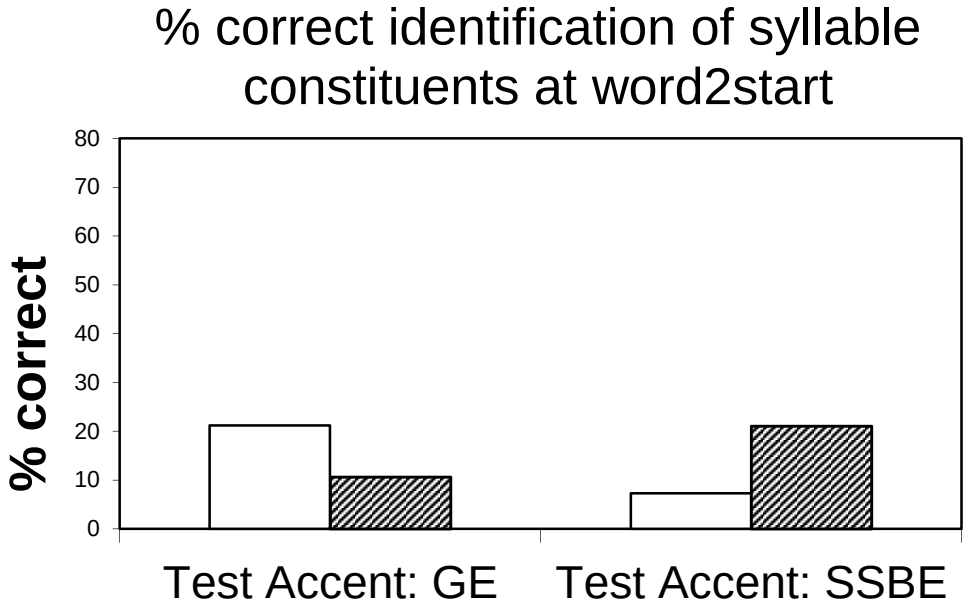
Final /d/
e.g. *he'd iced*

Glasgow speakers use duration less extensively to mark word boundaries, compared to SSBE speakers

4 speakers of Standard Southern British English (Smith & Hawkins 2012) + 4 age-matched speakers of Glasgow English

inter-accent differences in FPD reflected in perception





Before training, **both** groups of listeners are **poorer** at identifying boundary-adjacent segments in the other dialect, relative to speakers of that dialect, and relative to their own dialect.

Glaswegians improve more with SSBE, relative to the other subject groups (19-21%, vs 12-15%): SSBE offers more cues (in the durational domain).

University of Glasgow

Expt 3 Summary

Experience with a regional accent affects its intelligibility in noise.

Contrary to previous studies, experience of a “standard” variety (gained e.g. via media) does not completely protect listeners who have a non-standard accent themselves from processing difficulty, when listening to the standard

Results seem to depend on materials and task used.

University of Glasgow

Conclusions

SPiN conditions can reveal a facilitatory role in word recognition for constellations of weak cues that are coherent with the listener's experience of speech

The combination of SPiN with **casual speech styles** and **specifically-controlled sources of phonetic variability** may reveal areas of processing difficulty that are not apparent with other approaches.

University of Glasgow

Thank you!

ICPhS2015

18TH INTERNATIONAL CONGRESS
OF PHONETIC SCIENCES
10 AUG-14 AUG 2015 SECC GLASGOW UK

www.ICPhS2015.info

University of Glasgow

Queen Margaret University
EDINBURGH

University of Strathclyde

The University of Edinburgh