How does voice familiarity affect speech intelligibility?

Emma Holmes

Division of Psychology and Language Sciences

University College London

Familiar-voice information

Recognition

Intelligibility in noisy environments

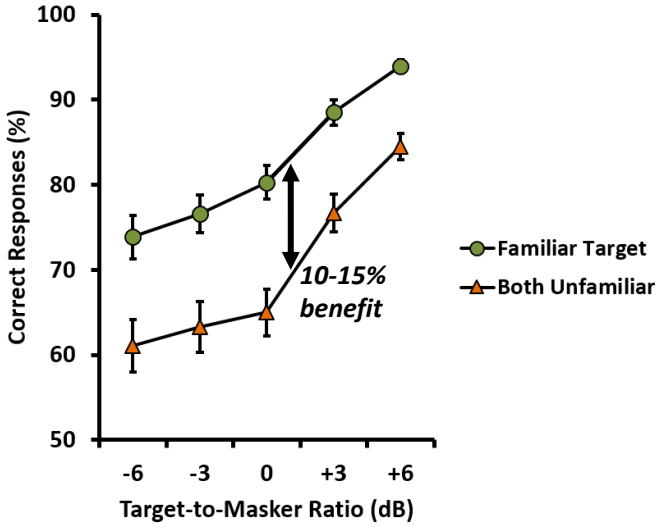
Familiar-voice benefit to intelligibility

Spouses married > 10 years

(Johnsrude *et al.*, 2013, *Psych Science*)

Target
e.g., "Ready Baron, go to *green two* now"

Masker
e.g., "Ready Charlie, go to blue four now"

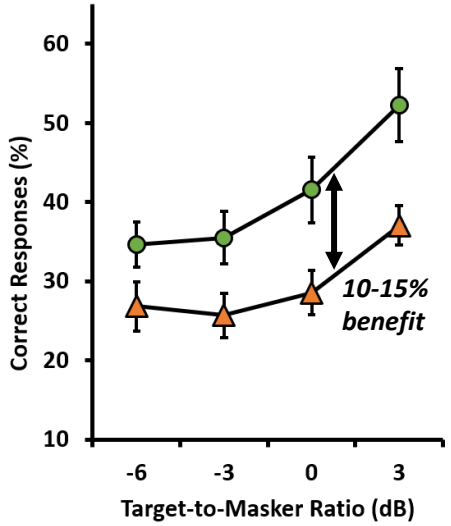


Friends known ≥ 6 months

(Domingo, Holmes, & Johnsrude, 2020, JEP:Applied)

Target
e.g., "Bob *bought three old bags*"

Masker
e.g., "Pat lost three small shoes"



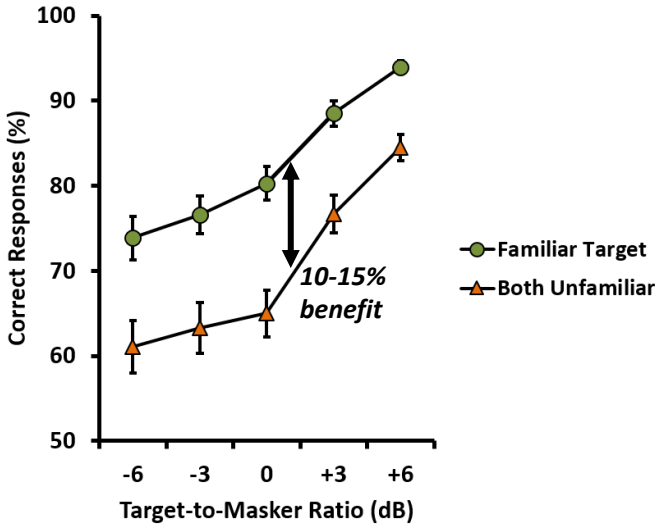
Familiarity with the acoustic characteristics of a voice improves intelligibility

Familiar-voice benefit to intelligibility

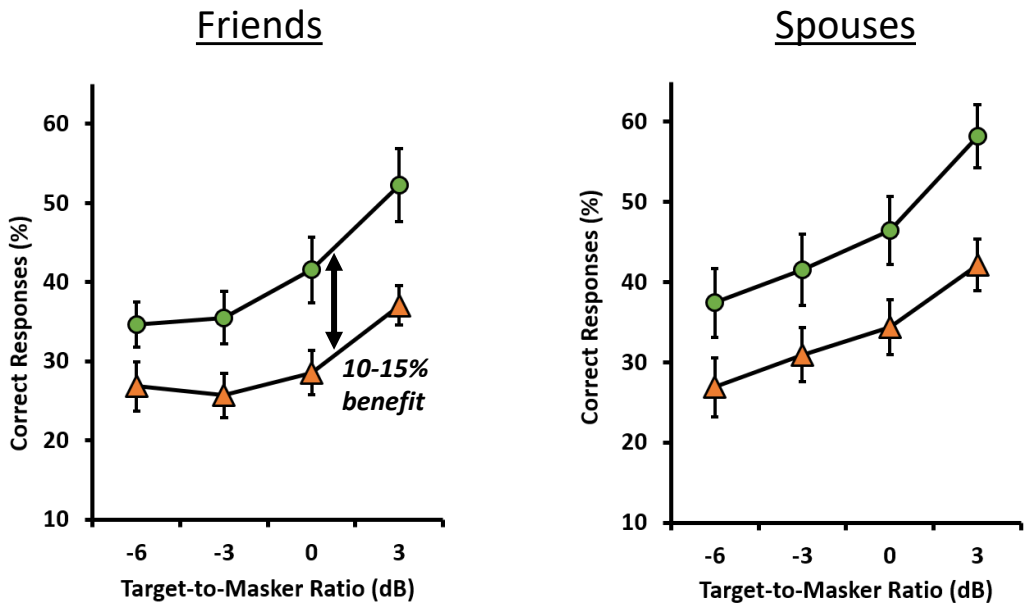
Spouses married > 10 years
(Johnsrude *et al.*, 2013, Psych Science)

Target
e.g., "Ready Baron, go to *green two* now"

Masker
e.g., "Ready Charlie, go to blue four now"



Friends known ≥ 6 months
(Domingo, Holmes, & Johnsrude, 2020, JEP:Applied)



Familiarity benefit develops relatively quickly

What makes a familiar voice more intelligible?

1) What processes underpin the benefit to intelligibility for naturally familiar voices?

2) How do people learn to become familiar with new voices?

What makes a familiar voice more intelligible?

1) What processes underpin the benefit to intelligibility for naturally familiar voices?

2) How do people learn to become familiar with new voices?

General Methods

- Participants: Pairs of friends/couples known each other > 6 months

- Participants hear recordings of their partner (familiar) and the partners of two other participants (unfamiliar)

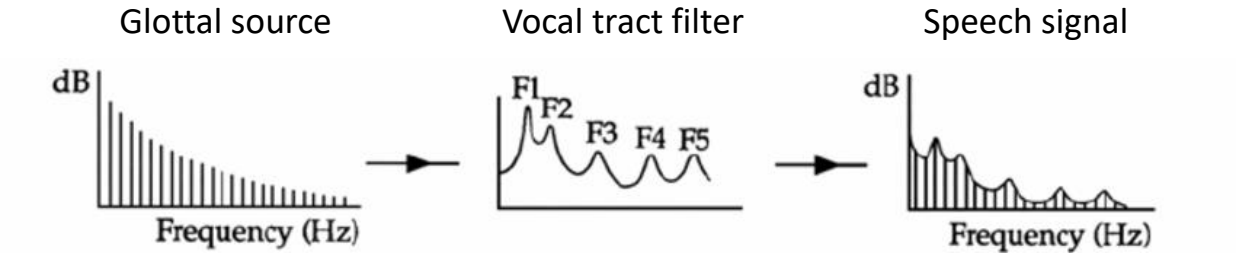
Name	Verb	Number	Adjective	Noun
Pat	bought	two	big	bags
Bob	**found**	three	**blue**	cards
	gave	**four**	cold	**gloves**
	held	five	hot	hats
	lost	six	new	pens
	saw	eight	old	shoes
	sold	nine	red	socks
	took	ten	small	toys

(Kidd, Best, & Mason, 2008)

Which acoustic properties do listeners use?

Voice recognition relies on:

- Fundamental frequency (f_0)
  (e.g., LaRiviere, 1975; Lavner et al., 2000, 2001; van Dommelen, 1987)
- Vocal tract characteristics
  (e.g., Abberton & Fourcin, 1975; van Dommelen, 1990)



(adapted from Keller, 2004)

Do these cues contribute in the same way to the familiar-voice benefit to intelligibility?

Which acoustic properties do listeners use?

Voice Recognition:

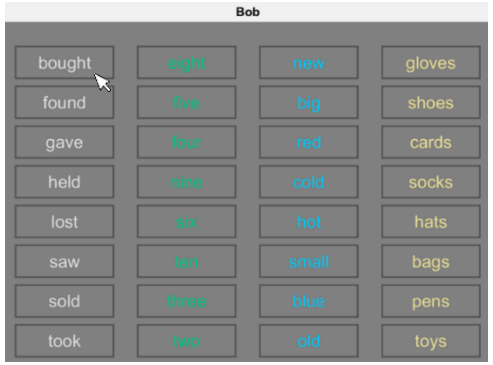
e.g., "Pat sold four big hats"

Was this sentence spoken by your familiar voice?

Yes / No

Speech Intelligibility:

Target
e.g., "**Pat** lost three small shoes"
Masker
e.g., "Bob bought three old bags"



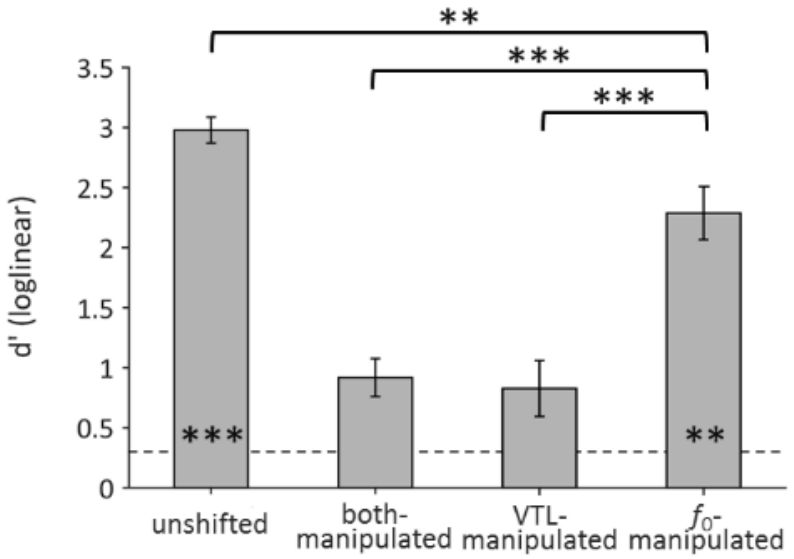
Manipulation conditions:

- **f_0-manipulated**
- **VTL-manipulated**
- **both-manipulated**
- **unshifted**

Holmes, Domingo & Johnsrude, 2018, Psych Science

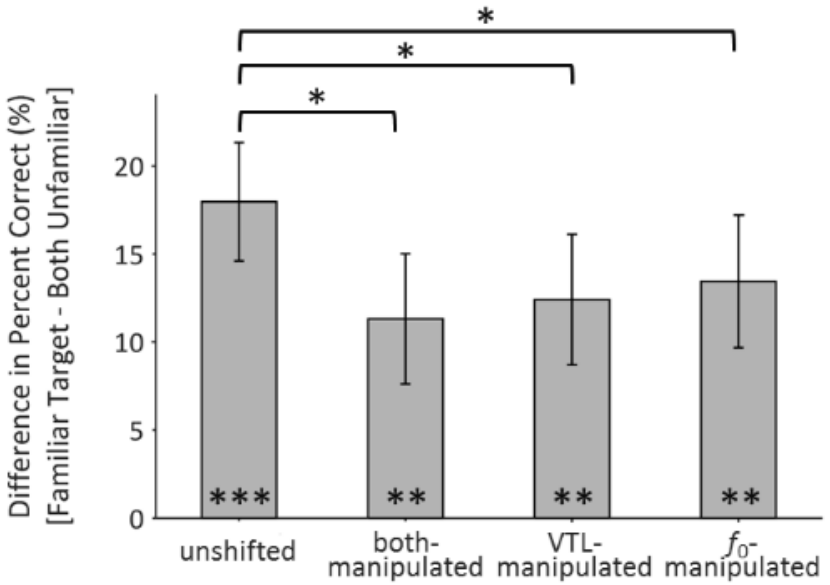
Which acoustic properties do listeners use?

Voice Recognition



Error bars show ± 1 SE

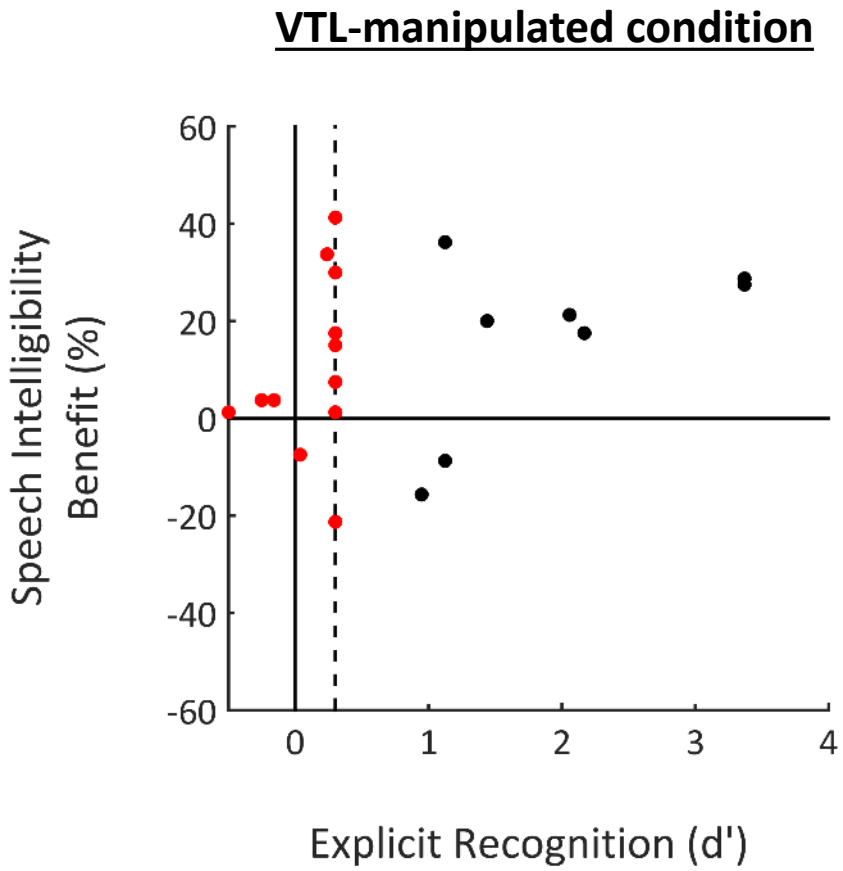
Speech Intelligibility Benefit



Voice manipulations affect recognition and the familiar-voice intelligibility benefit differently

Holmes, Domingo & Johnsrude, 2018, Psych Science

Which acoustic properties do listeners use?

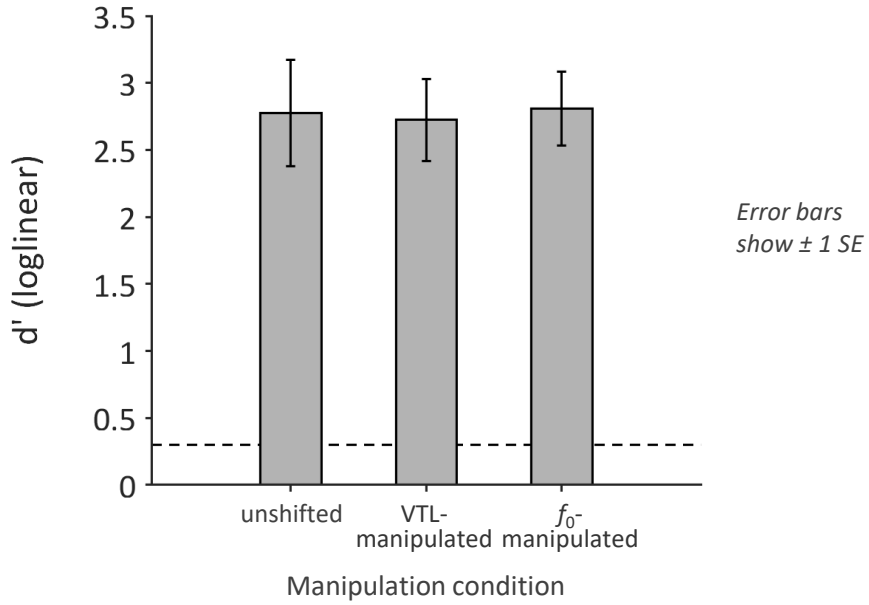


VTL-manipulated familiar voice is more intelligible, even if not recognised

Holmes, Domingo & Johnsrude, 2018, Psych Science

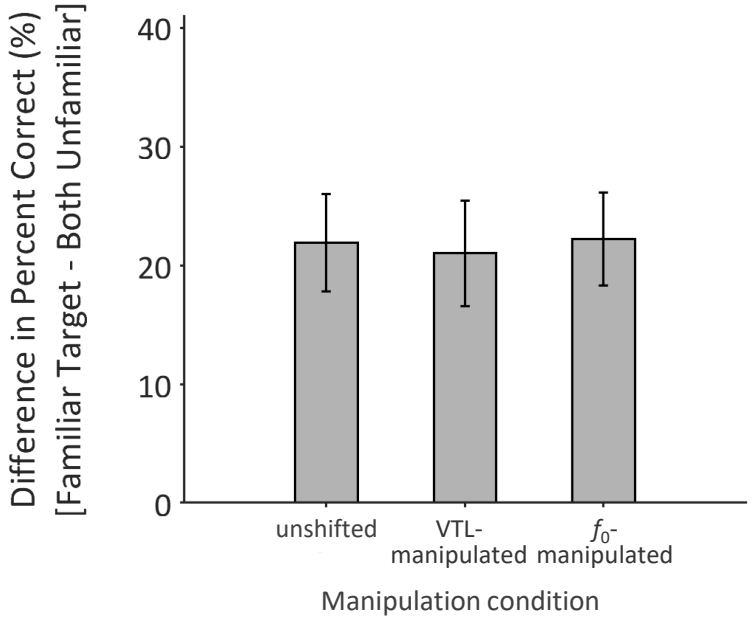
Smaller manipulations to f_0 and VTL

Voice Recognition



Error bars show ± 1 SE

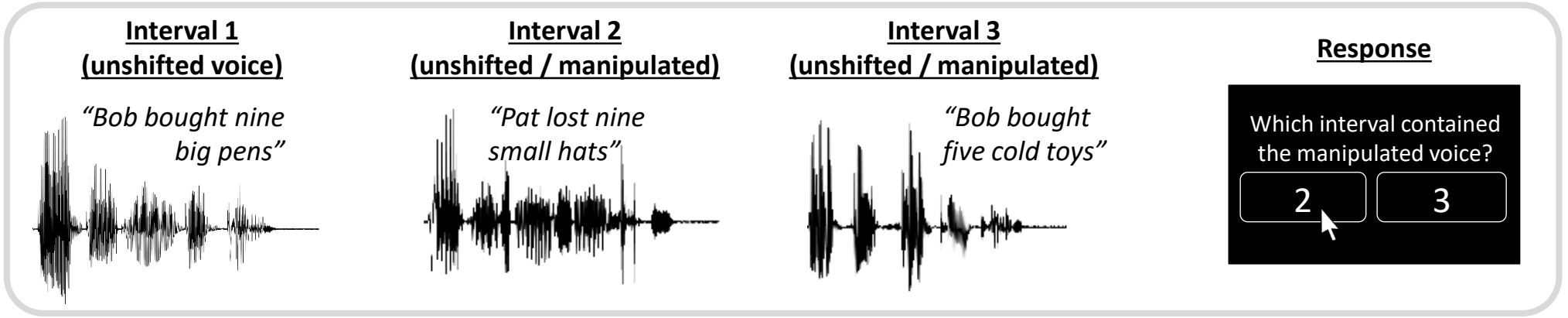
Speech Intelligibility Benefit

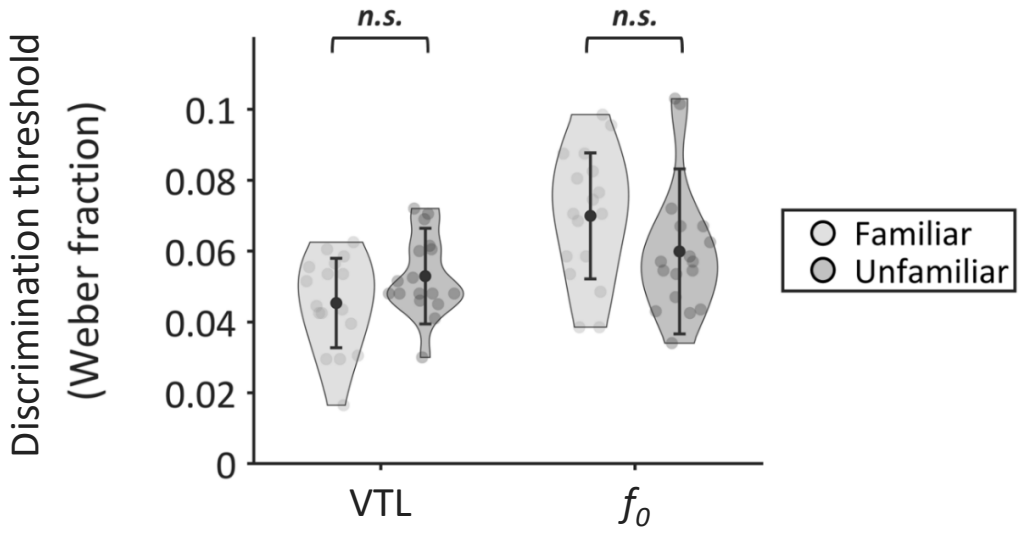


Explicit recognition and speech intelligibility of familiar voices are robust to 'small' manipulations in f_0 and VTL that participants can reliably detect

Holmes & Johnsrude, 2023, Hearing Research

Discrimination of f_0 and VTL





Discrimination thresholds are no better for familiar than unfamiliar voices

Holmes & Johnsrude, 2023, Hearing Research

Does the familiar-voice benefit change with other maskers?

English target sentence
(familiar / unfamiliar voice)

e.g., **Peter got twelve cheap spoons**

Three masker conditions:

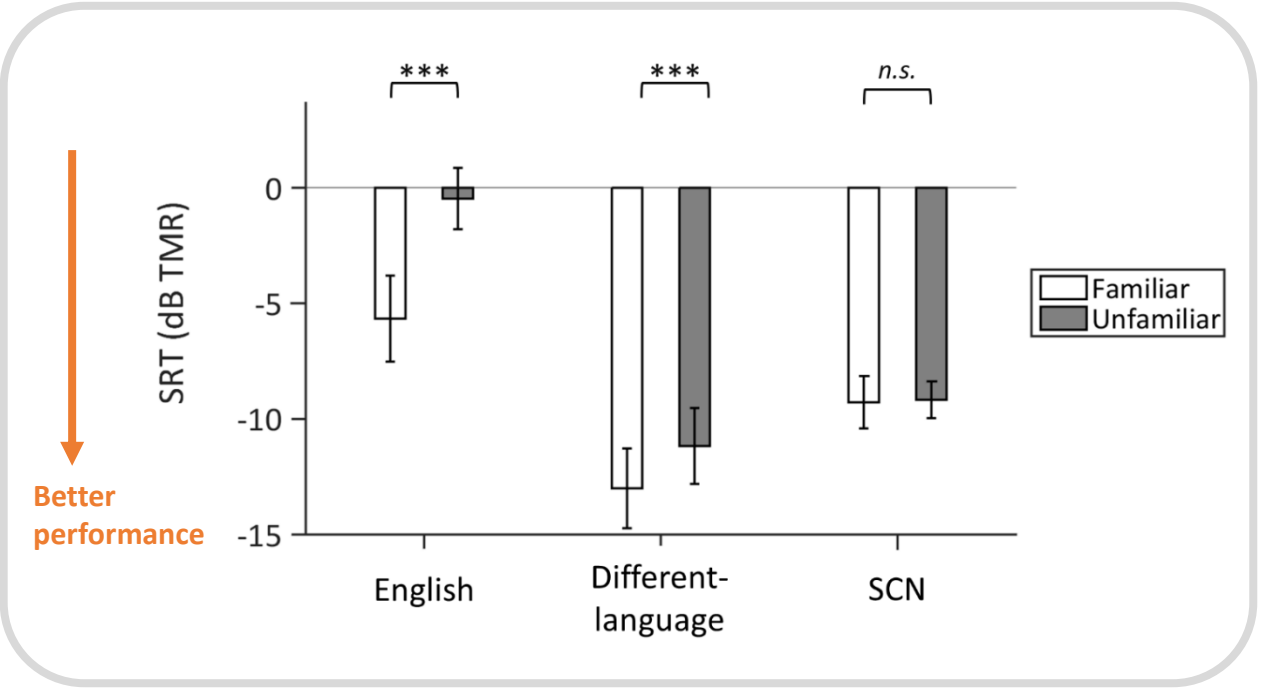
1) English sentence

e.g., **Rachel brought eight cheap toys**

2) Different-language sentence

e.g., **Carlos hace veinte platos baratos**

3) Signal-correlated noise (SCN)



Naturally familiar voices reduce cognitive interference from masker with relevant linguistic information

Holmes & Johnsrude, 2020, JEP:LMC

What makes a familiar voice more intelligible?

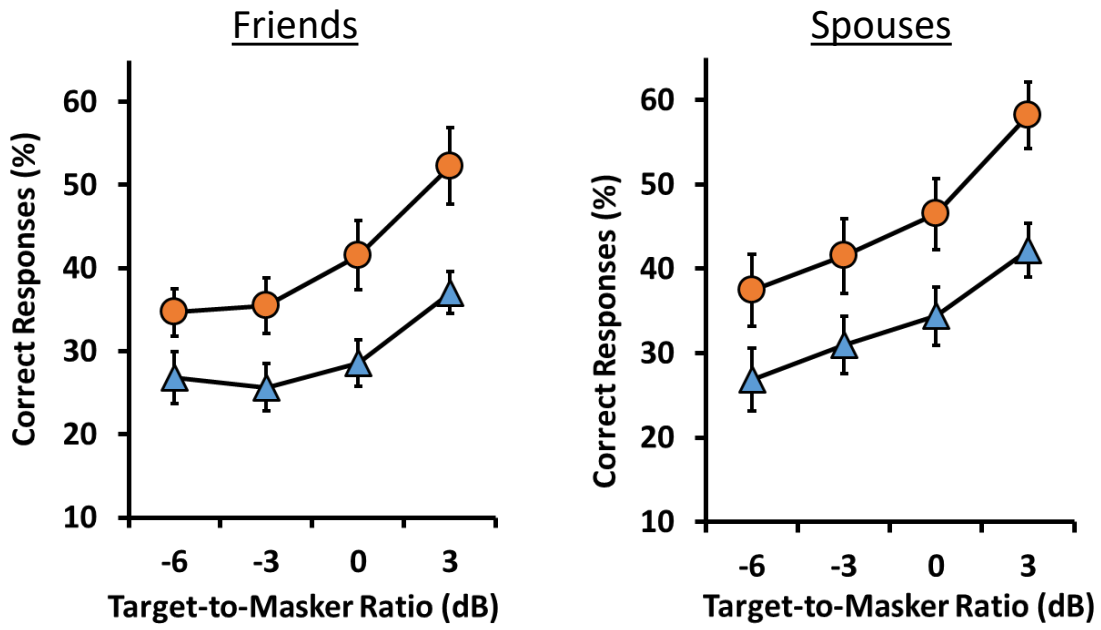
1) What processes underpin the benefit to intelligibility for naturally familiar voices?

2) How do people learn to become familiar with new voices?

Natural familiarity versus voice training

Naturally familiar voices

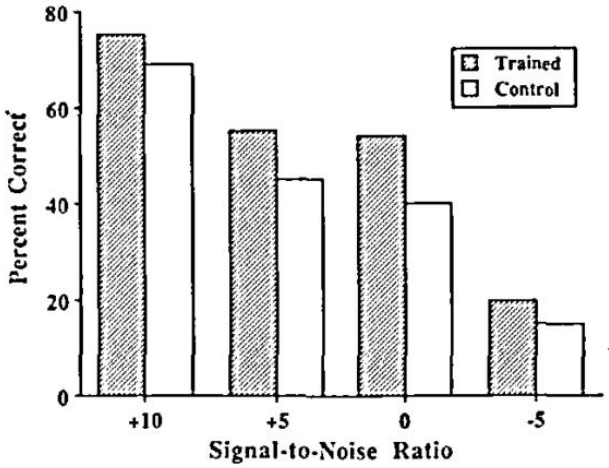
(Domingo, Holmes, & Johnsrude, 2020, *JEP Applied*)



Trained talkers

(e.g., Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998)

Nygaard, Sommers, & Pisoni, 1994



How rapidly does voice training improve intelligibility?

How rapidly does voice training improve intelligibility?

Familiarisation:
(5 mins)

Sentence
e.g., "The actors learned their scripts that afternoon"

Mike

3 talkers

Training:
(approx. 1.5 hours)

Sentence
e.g., "The birds were making a loud racket"

Multi-talker babble
(half of participants)

Who spoke this sentence?

Mike / Jeff / Brad

Correct

Talker was MIKE

Same 3 talkers:

- **Most Familiar: 67%**
- **Moderately Familiar: 22%**
- **Least Familiar: 11%**

Holmes, To, & Johnsrude, 2021, Psych Science

How rapidly does voice training improve intelligibility?

Voice Identification:

Sentence
e.g., "Pat sold four big hats"

Was this sentence spoken by a familiarized voice?

Yes / No

If yes, who?

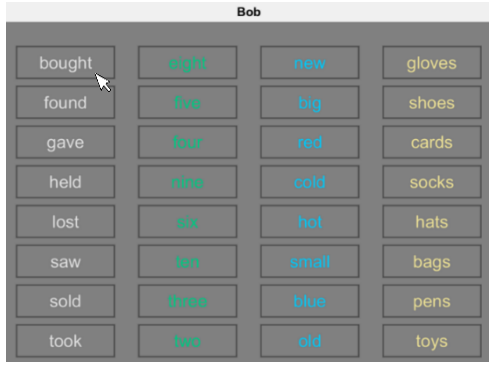
Mike / Jeff / Brad

5 talkers:
- **3 Familiarised**
- **2 Unfamiliar**

Speech Intelligibility:

Target
e.g., "**Pat** lost three small shoes"

Masker
e.g., "Bob bought three old bags"



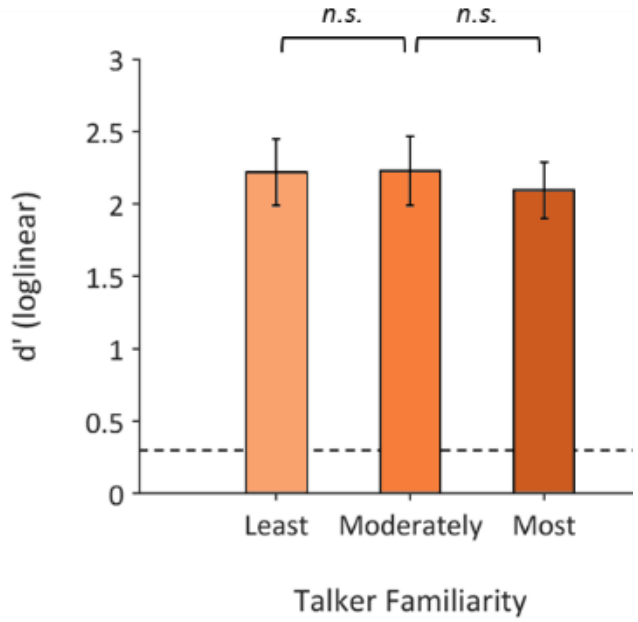
- **Both Unfamiliar**
- **Most Familiar Target**
- **Moderately Familiar Target**
- **Least Familiar Target**

Note: different sentences were presented during training and testing

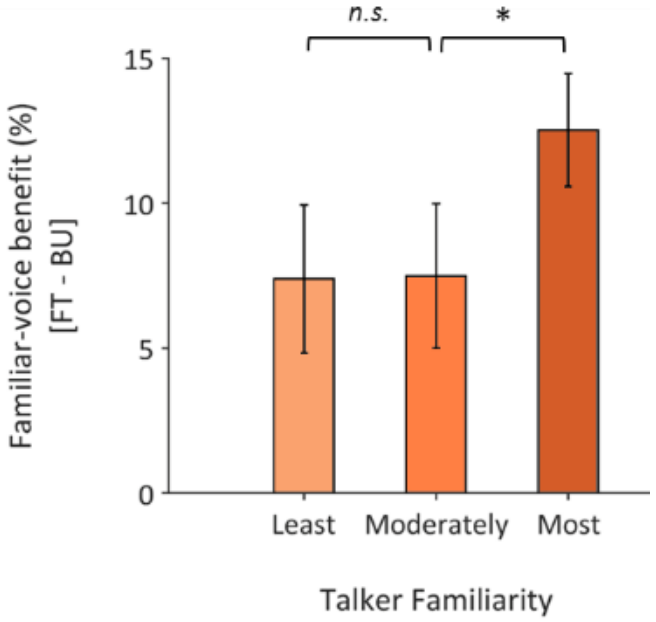
Holmes, To, & Johnsrude, 2021, Psych Science

How rapidly does voice training improve intelligibility?

Voice Identification



Speech Intelligibility Benefit

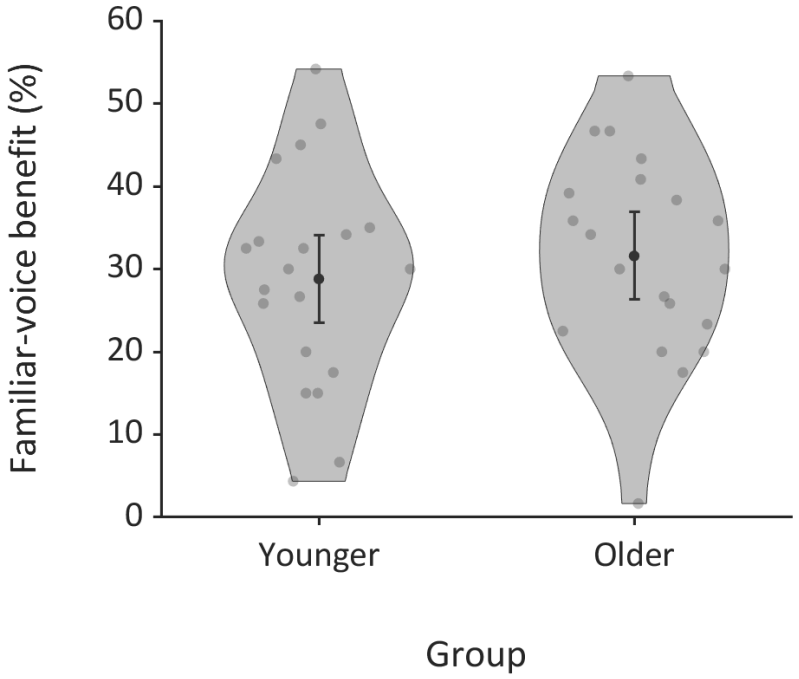


Speech intelligibility improves rapidly within 10 minutes of voice training

Holmes, To, & Johnsrude, 2021, Psych Science

Online voice training in older and younger adults

20 older (aged 55–73) and 20 younger (aged 18–34) participants



Older participants benefit as much as younger adults from voice training delivered online

Zhu & Holmes, 2024, PsyArXiv

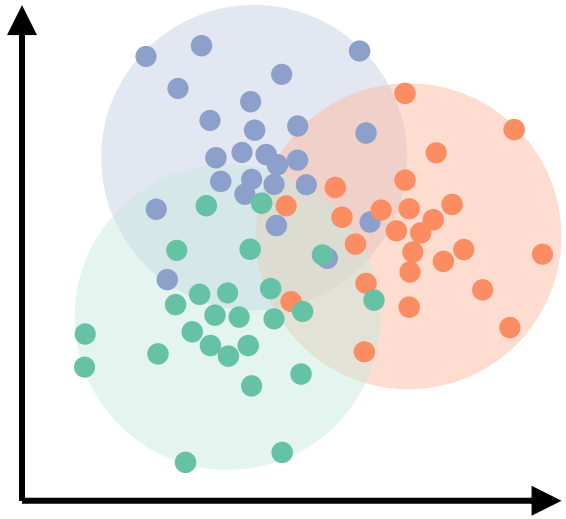
What are people learning?

Cognition
Volume 193, December 2019, 104026

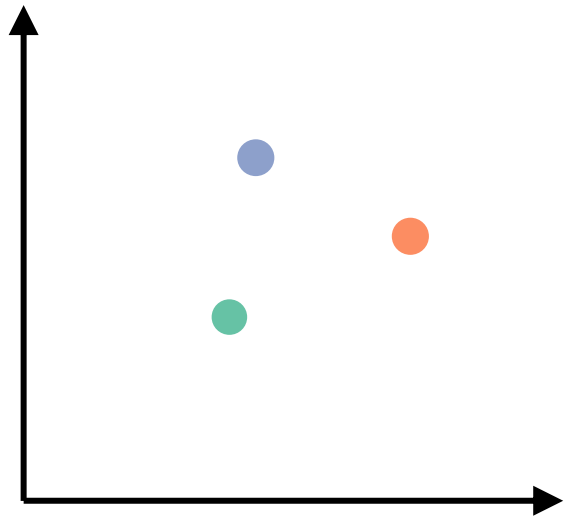
Original Articles

The effects of high variability training on voice identity learning

Nadine Lavan [a b], Sarah Knight [a], Valerie Hazan [a], Carolyn McGettigan [a b]



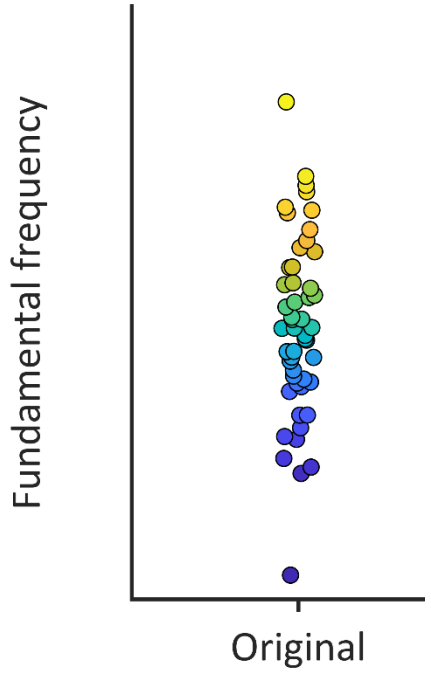
Voice 1
Voice 2
Voice 3

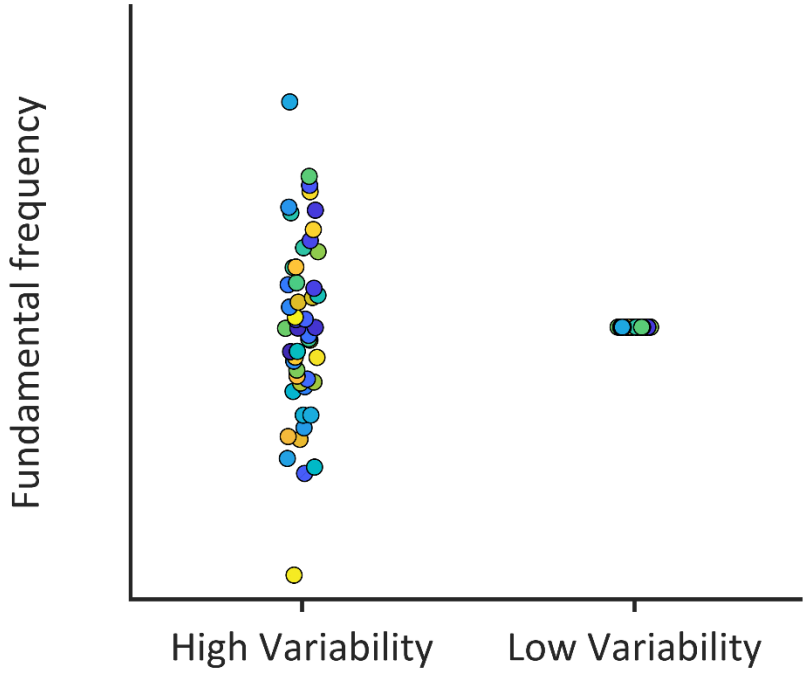


Voice 1
Voice 2
Voice 3

Are people learning about the variability of a voice?

Is variability in f0 during training critical for the speech intelligibility benefit?

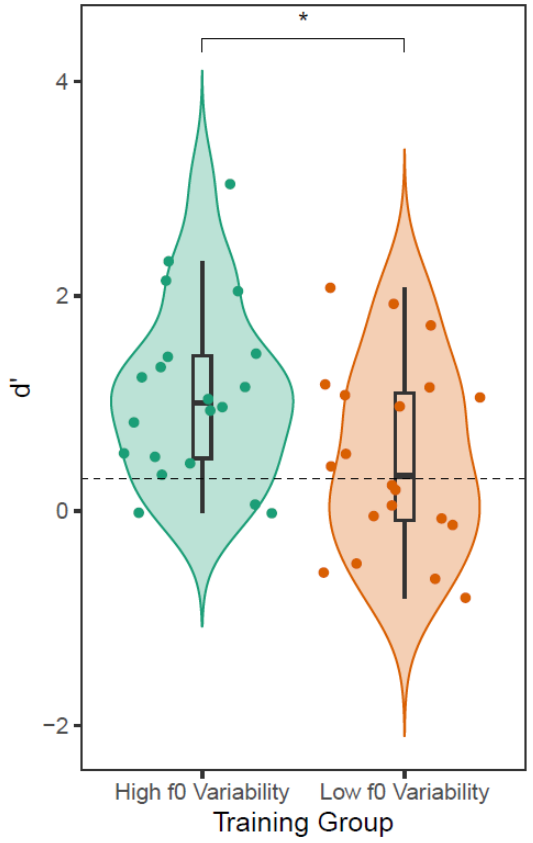




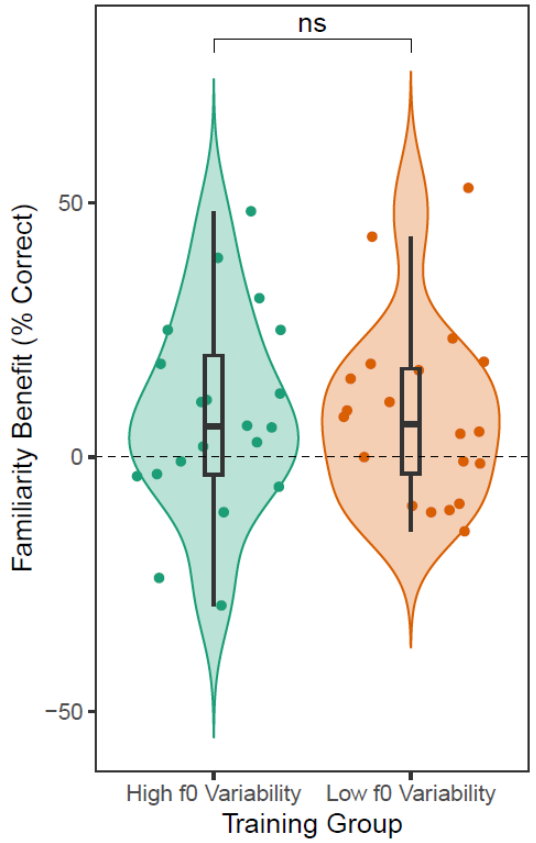
Test stimuli were not manipulated in either group (and were different sentences than the training stimuli)

Are people learning about the variability of a voice?

Voice Identification



Speech Intelligibility Benefit



People use variability in fundamental frequency for voice identification but don't rely on this variability for improving intelligibility

Conclusions

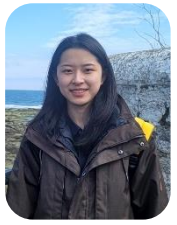
1) Familiar voices are more intelligible, even when they are not recognised as familiar. This benefit is due to reduced linguistic interference when we listen to sentences spoken by familiar people

2) Even short durations of training are sufficient to produce better intelligibility than novel voices, and can benefit older and younger adults


Ingrid Johnsrude


Ysabel Domingo


Grace To


Wansu Zhu


Harriet Smith

NSERC CRSNG
Fondation Brain Canada Foundation
Western BrainsCAN Transforming brain research.
CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada
EPS