Background and Motivation

- Prior exposure to consistent reverberation affects & improves speech perception (e.g., Brandewie & Zahorik, 2012, JASA; Zahorik & Brandewie, 2018, JASA; Srinivasan & Zahorik, 2013, JASA-EL; Watkins et al., 2005, JASA; Beeston et al., 2014, JASA)
- Here, we investigate the effects of prior exposure
  - to a consistent room → **buildup** of adaptation to reverberation
  - to an inconsistent room (anechoic or different reverb) → **breakdown**
- Focusing on
  - the unique effects of reverberation (i.e., without adding noise)
  - adaptation at the segmental **phoneme** & phonetic **feature** level, over a more representative range of speech sounds than what is typically used
- We also examine
  - the time course of adaptation to reverberation
  - effects of prior knowledge of target speech location

Ueno Kanako,
Meiji University,
Tokyo

Thank you!! ☺

BOSTON UNIVERSITY

Methods – Speech Stimuli & Simulated Rooms

Table 1. Phonetic feature classification. Consonants not used as stimuli are in grey. All consonants could be potential responses.

Feature		Consonants
Manner of articulation	Stop	k, t, p, g, d, b
	Fricative	f, v, ð, z, θ, s, ʃ
	Nasal	m, n, ŋ
Place of articulation	Labial	p, b, m, v, f
	Coronal	d, ð, θ, t, n, s, z
	Dorsal	k, g, ŋ, ʃ (post-alveolar)
Voicing	Voiced	g, d, b, v, ð, m, n, ŋ, z
	Unvoiced	k, t, p, f, θ, s, ʃ

- **R1** measured in elliptical church (distance from sound source 12 m).
  - **R2** in large concert hall (distance from source 33 m).
    - BRIRs equalized for overall energy.
      - T_{60} values larger for **R1** than **R2**.

Speech stimuli convolved with BRIRs from **two** different **rooms** with high levels of reverberation, and an **anechoic** room

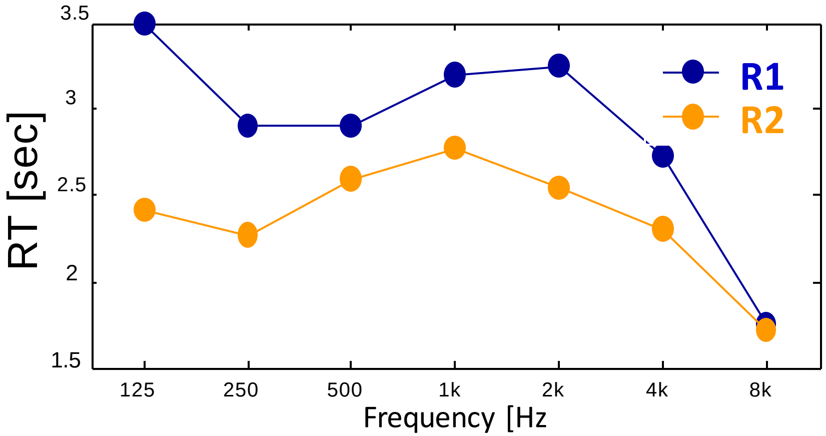


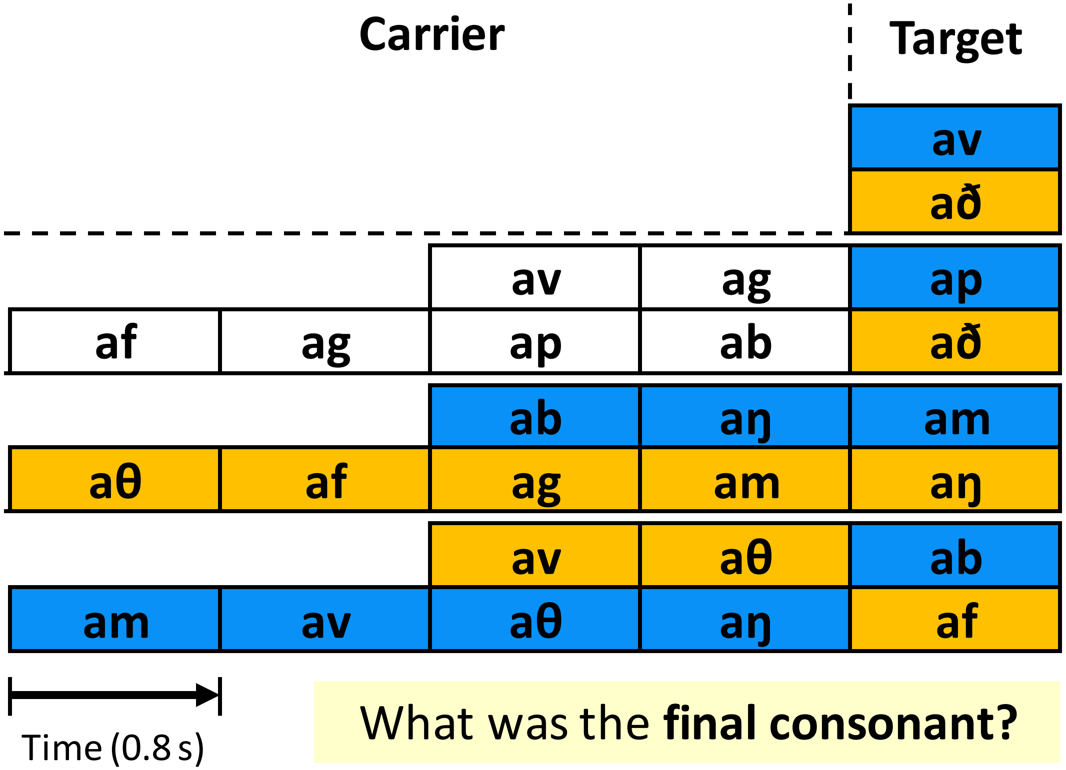
Figure 1. Acoustic characteristics of the BRIRs. Reverberation time (T_{60}) for each frequency band.

Methods – Task Schematics

Experiment 1

Carrier Type

Carrier Length



No Carrier, Target in **R1** or **R2**

0 - VC

Anechoic Carrier, Target in **R1** or **R2**

2 - VC

4 - VC

Same Carrier, both Carrier and Target in **R1** or **R2**

2 - VC

4 - VC

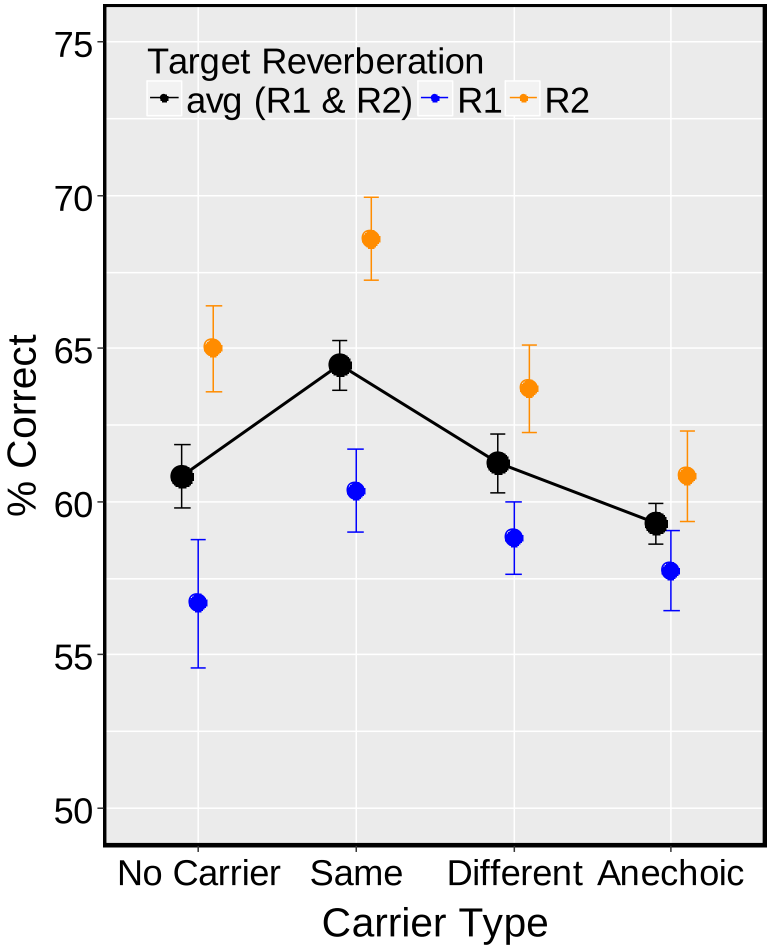
Different Carrier, Carrier in **R2** – Target in **R1** or Carrier in **R1** – Target in **R2**

2 - VC

4 - VC

What was the **final consonant?**

Experiment 1 – **Buildup** & **breakdown** of adaptation to reverberation



- Main effect of Carrier Type

  ($F_{(3,39)} = 5.38$, $p = 0.006$)

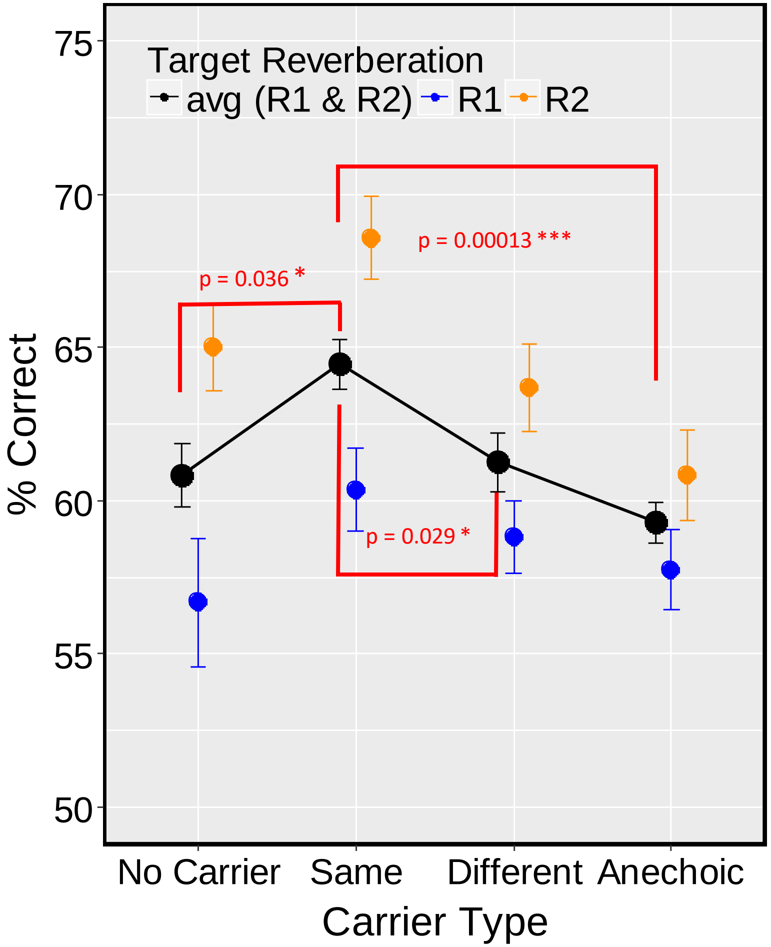
- Main effect of Target Reverberation

  ($F_{(1,13)} = 18.85$, $p < 0.001$)

- No interaction

  ($F_{(3,39)} = 1.78$, $p = 0.16$)

Experiment 1 – **Buildup** & **breakdown** of adaptation to reverberation



- Main effect of Carrier Type

  ($F_{(3,39)}$ = 5.38, p = 0.006)
- Main effect of Target Reverberation

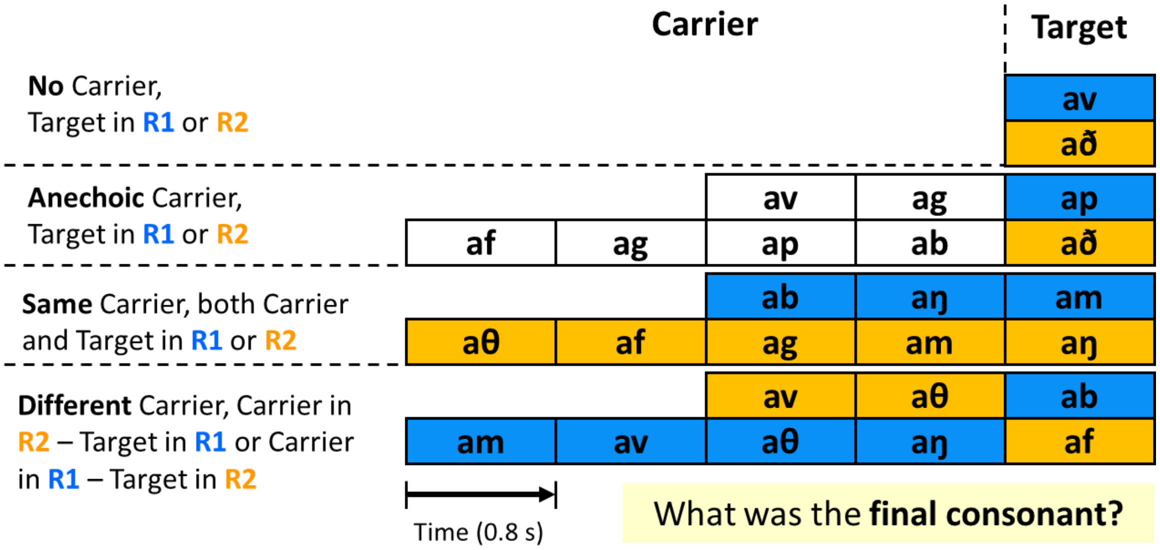
  ($F_{(1,13)}$ = 18.85, p < 0.001)
- No interaction

  ($F_{(3,39)}$ = 1.78, p = 0.16)

Paired t-tests
Same differs from all others, other comparisons ns

→ **Buildup** and **breakdown** of adaptation to reverberation

Methods – Task Schematics Exp. 1 & 2

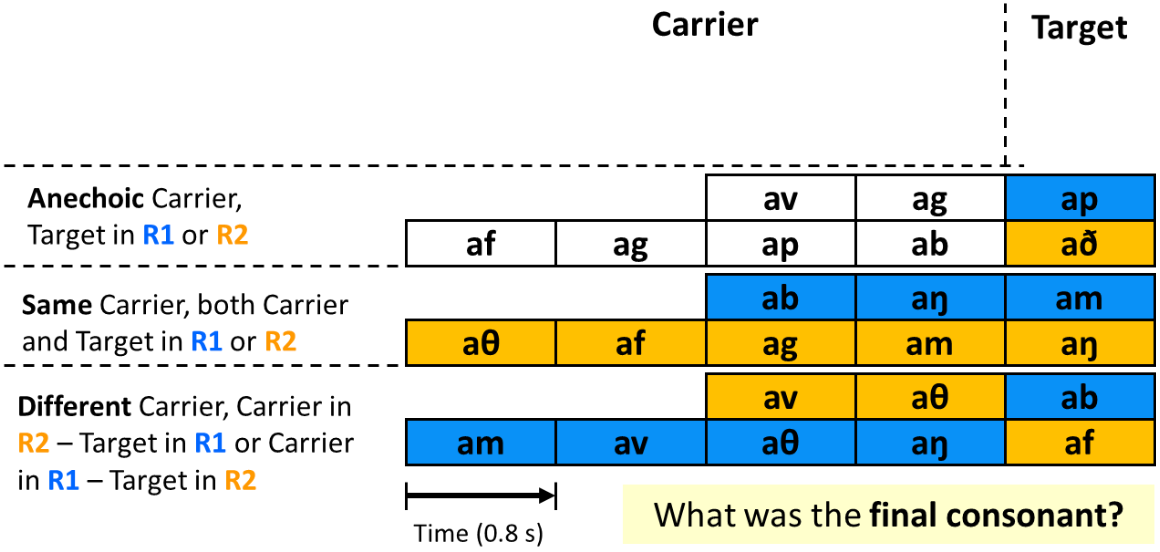


	Carrier				Target
No Carrier, Target in R1 or R2					av
					að
Anechoic Carrier, Target in R1 or R2			av	ag	ap
	af	ag	ap	ab	að
Same Carrier, both Carrier and Target in R1 or R2			ab	aŋ	am
	aθ	af	ag	am	aŋ
Different Carrier, Carrier in R2 – Target in R1 or Carrier in R1 – Target in R2			av	aθ	ab
	am	av	aθ	aŋ	af

What was the **final consonant?**

Exp 1

- → Carrier Length variable
- → No Carrier condition



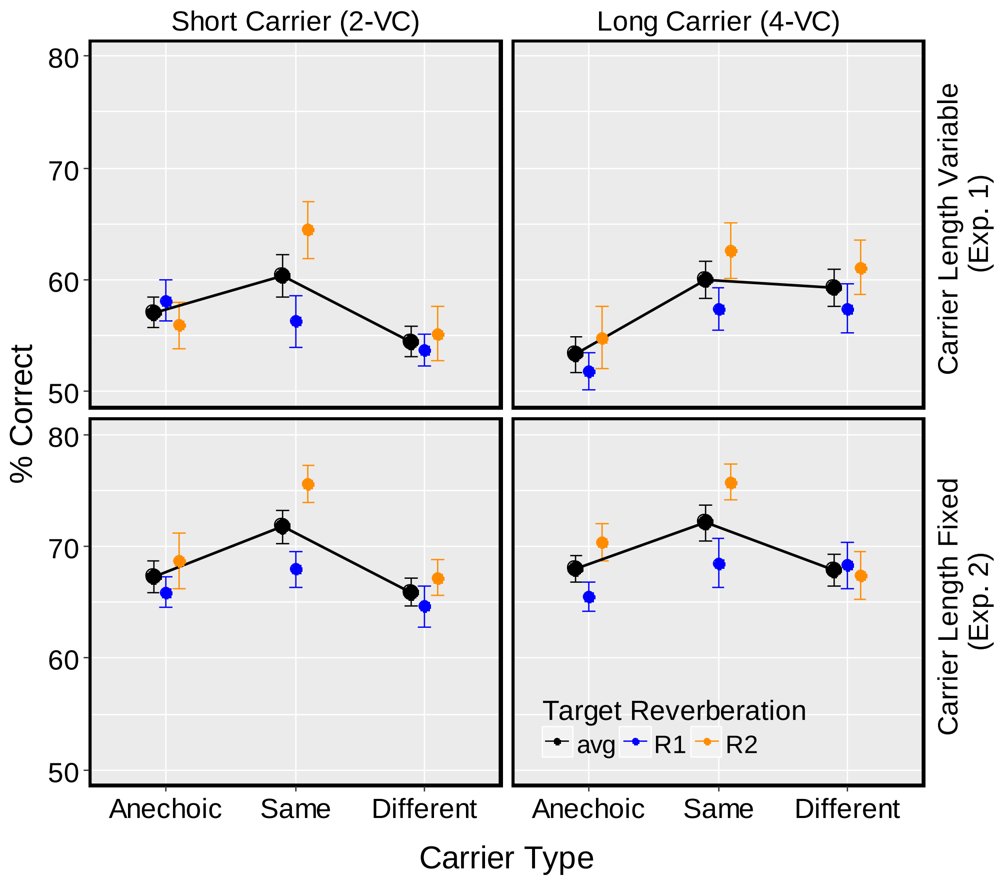
	Carrier				Target
Anechoic Carrier, Target in R1 or R2			av	ag	ap
	af	ag	ap	ab	að
Same Carrier, both Carrier and Target in R1 or R2			ab	aŋ	am
	aθ	af	ag	am	aŋ
Different Carrier, Carrier in R2 – Target in R1 or Carrier in R1 – Target in R2			av	aθ	ab
	am	av	aθ	aŋ	af

What was the **final consonant?**

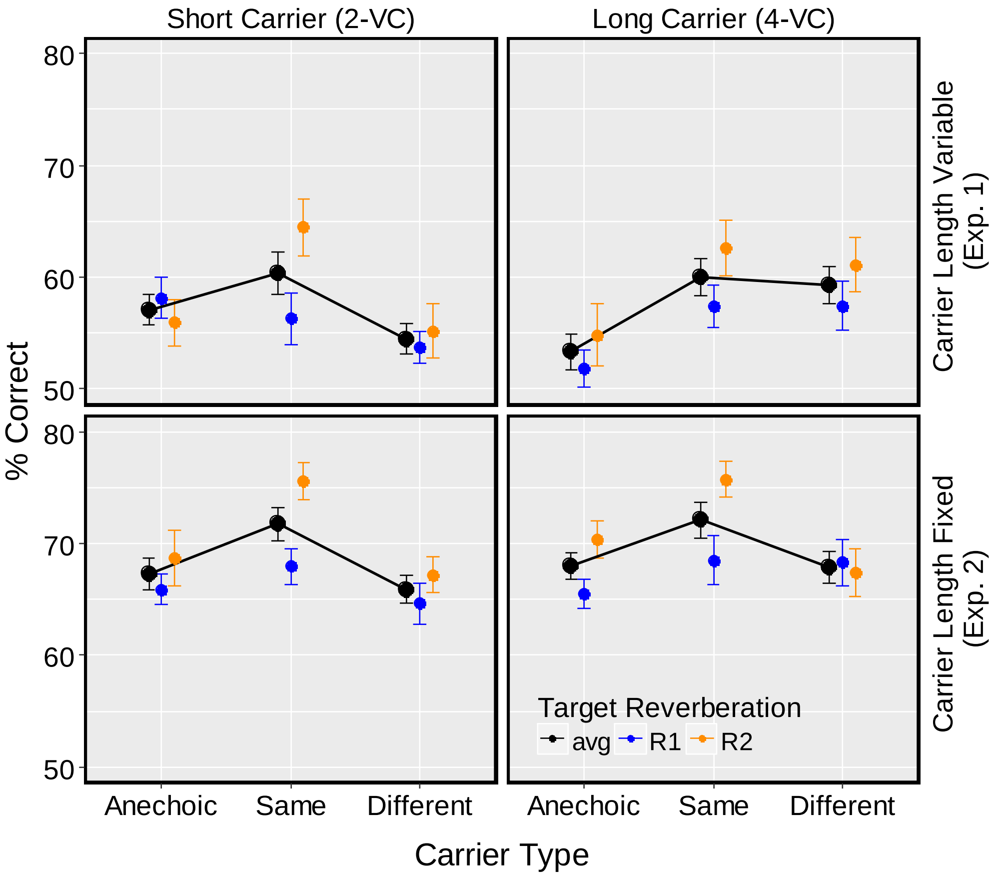
Exp 2

- → Carrier Length fixed

Experiments 1 & 2

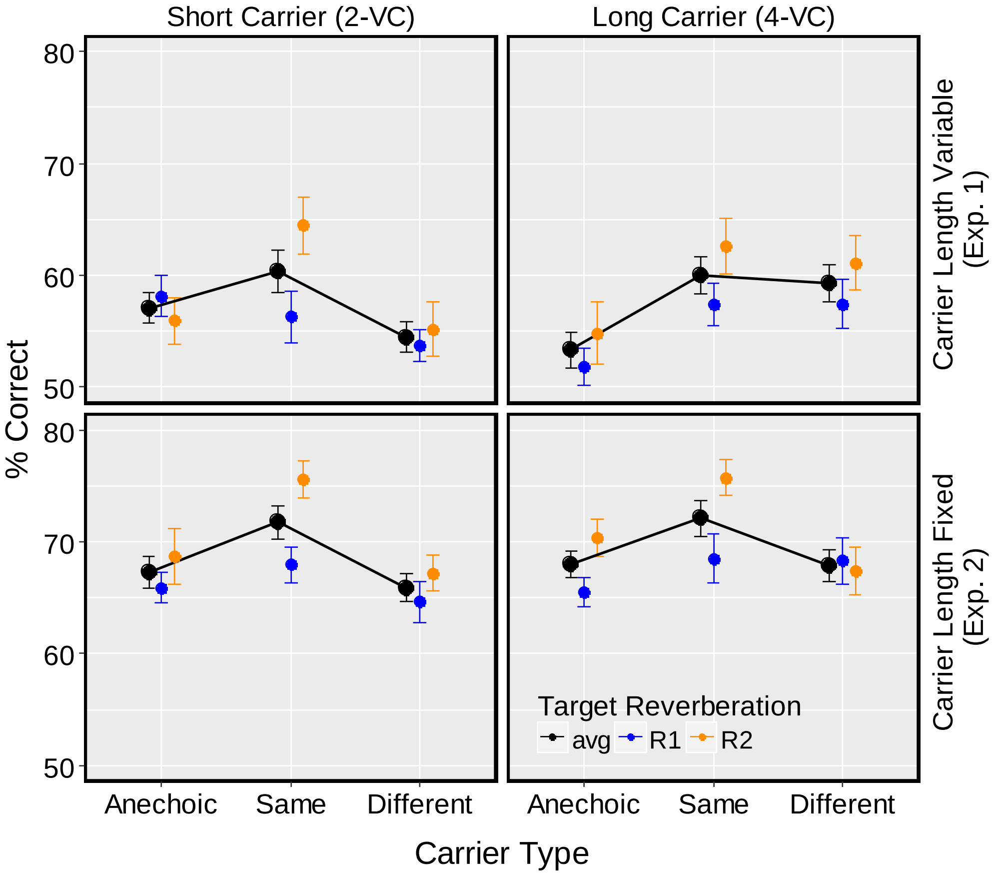


Experiments 1 & 2



- Main effect of Carrier Type ($F_{(2,32)}$ = 11.2, p = 0.0002),
- Target Reverberation ($F_{(1,16)}$ = 7.35, p = 0.015)
- Experiment ($F_{(1,16)}$ = 8.15, p = 0.012)
- No effect of Carrier Length ($F < 1$, ns)
- Significant interactions
  - Carrier Length X Carrier Type ($F_{(2,32)}$ = 5.34, p = 0.015)
  - **Carrier Type & Target Reverb** ($F_{(2,32)}$ = 7.48, p = 0.005)
- Trend Carrier Type X Carrier Length X Exp. ($F_{(2,32)}$ = 3.06, p = 0.07)
- All other interactions ns (p's > 0.30)

Experiments 1 & 2

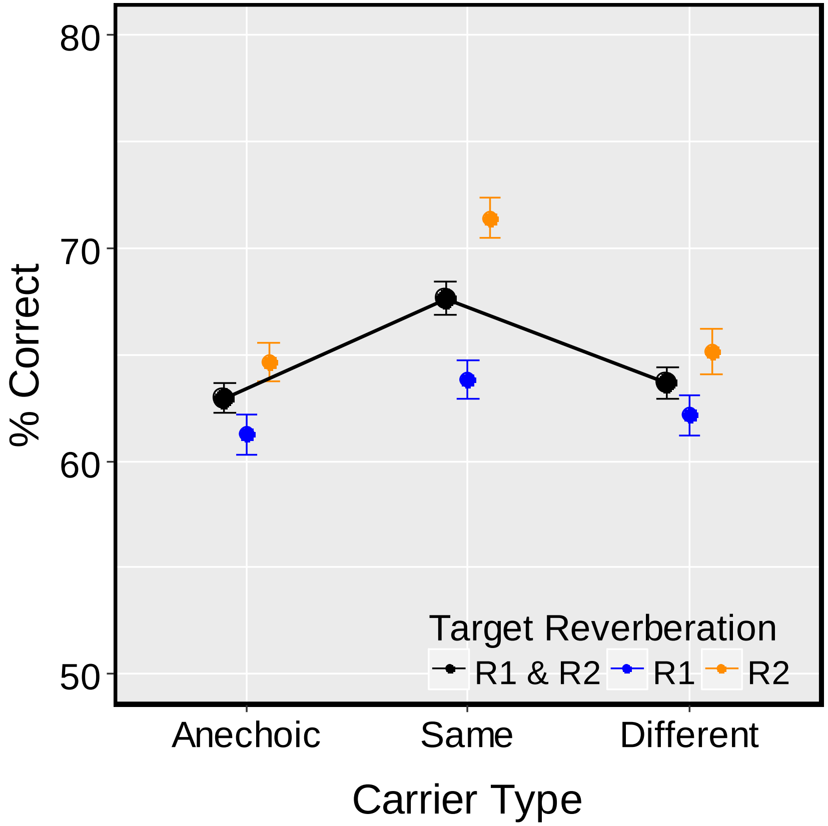


- Main effect of Carrier Type ($F_{(2,32)}$ = 11.2, *p* = 0.0002),
- Target Reverberation ($F_{(1,16)}$ = 7.35, *p* = 0.015)
- Experiment ($F_{(1,16)}$ = 8.15, *p* = 0.012)
- No effect of Carrier Length ($F < 1$, ns)
- Significant interactions
    - Carrier Length X Carrier Type ($F_{(2,32)}$ = 5.34, *p* = 0.015)
    - **Carrier Type & Target Reverb** ($F_{(2,32)}$ = 7.48, *p* = 0.005)
- Trend Carrier Type X Carrier Length X Exp. ($F_{(2,32)}$ = 3.06, *p* = 0.07)
- All other interactions ns (*p*'s > 0.30)

→ Exposure time (Carrier Length) & prior knowledge of target location (Experiment) do not affect magnitude of adaptation to reverberation

→ **Target Reverberation** important

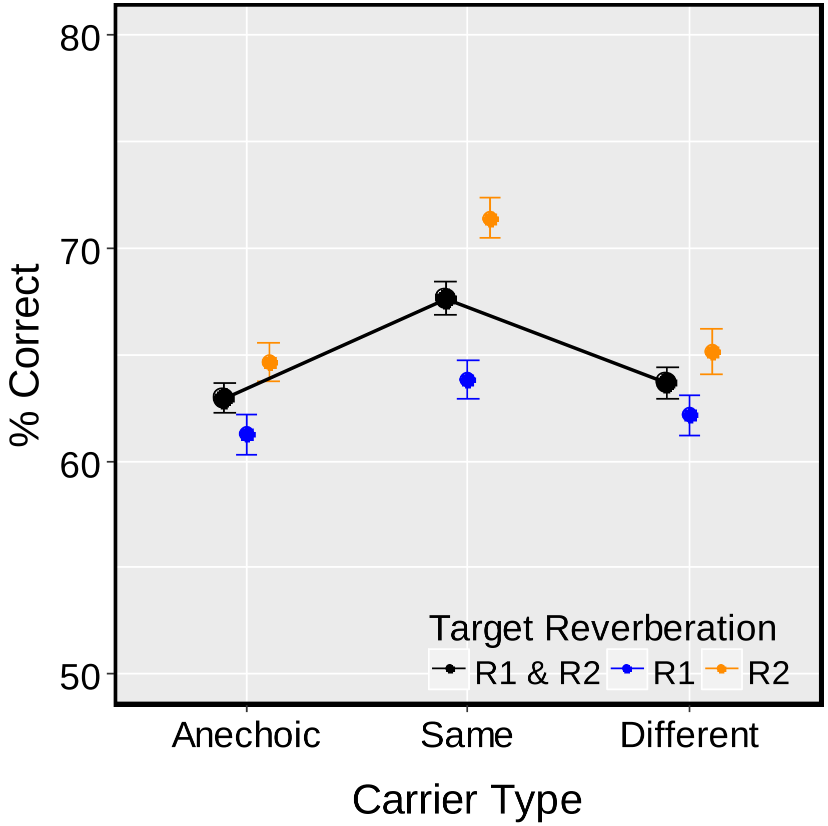
Pooled Data



Data from long and short carriers and from Exp. 1 & Exp. 2 pooled, replotted partial ANOVAs, separately for each Target Reverberation

- For R2, Main effect of Carrier Type; $F_{(2,34)} = 16.57$, $p < 0.001$)
- For R1 effect weaker ($F_{(2,34)} = 1.58$, $p =$ 0.221).

Pooled Data



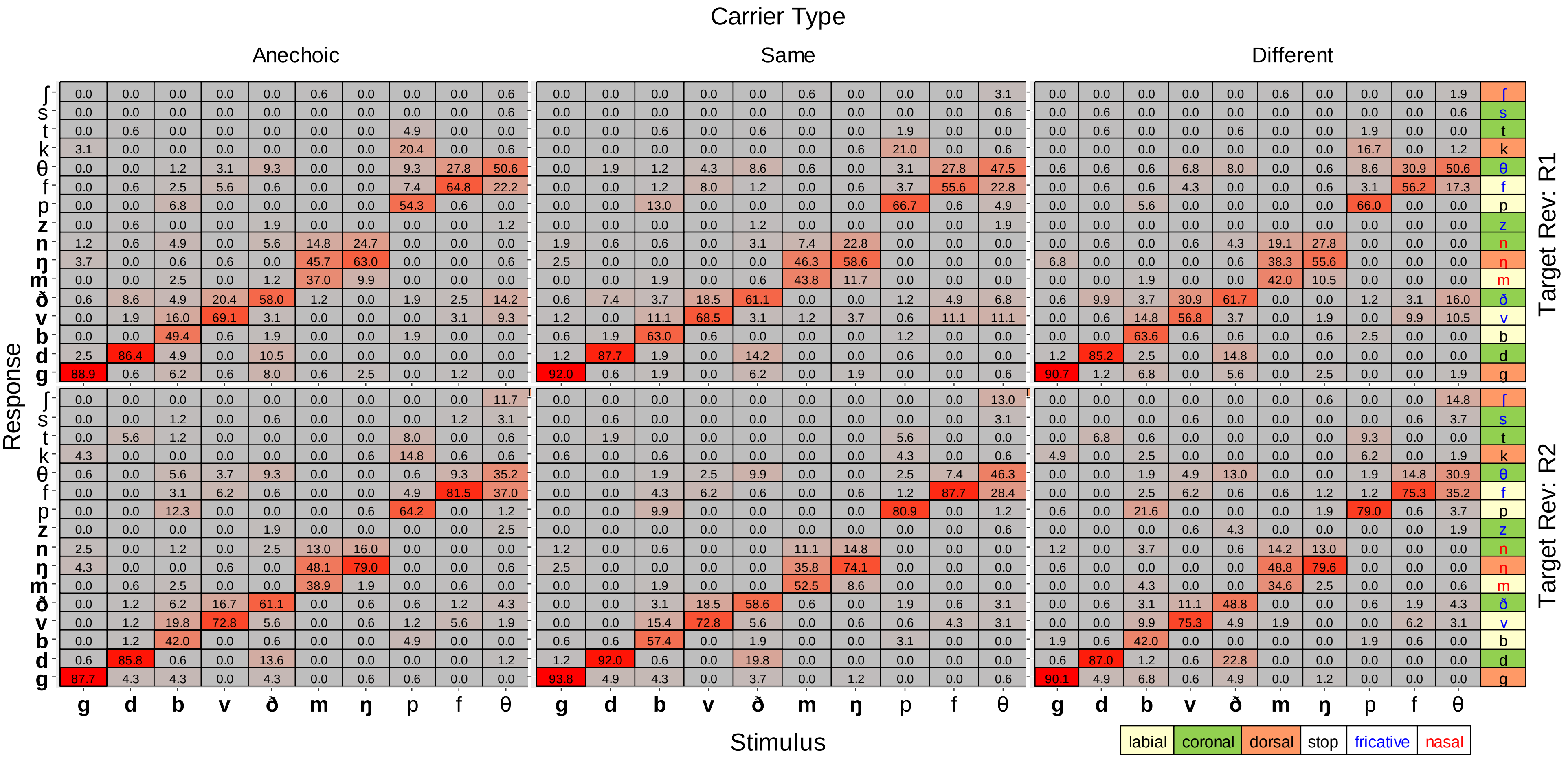
Data from long and short carriers and from Exp. 1 & Exp. 2 pooled, replotted partial ANOVAs, separately for each Target Reverberation

- For R2, Main effect of Carrier Type; $F_{(2,34)} = 16.57$, $p < 0.001$)

- For R1 effect weaker ($F_{(2,34)} = 1.58$, $p = 0.221$).

→ Improved speech perception after exposure to **consistent** reverberation
→ **Weaker** effect in the more reverberant environment

Pooled Data – Confusion Matrices



Carrier Type

Response

Stimulus

labial | coronal | dorsal | stop | fricative | nasal

Target Rev: R1

Anechoic

Response	g	d	b	v	ð	m	ŋ	p	f	θ
ʃ	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.6
s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
t	0.0	0.6	0.0	0.0	0.0	0.0	0.0	4.9	0.0	0.0
k	3.1	0.0	0.0	0.0	0.0	0.0	0.0	20.4	0.0	0.6
θ	0.0	0.0	1.2	3.1	9.3	0.0	0.0	9.3	27.8	50.6
f	0.0	0.6	2.5	5.6	0.6	0.0	0.0	7.4	64.8	22.2
p	0.0	0.0	6.8	0.0	0.0	0.0	0.0	54.3	0.6	0.0
z	0.0	0.6	0.0	0.0	1.9	0.0	0.0	0.0	0.0	1.2
n	1.2	0.6	4.9	0.0	5.6	14.8	24.7	0.0	0.0	0.0
ŋ	3.7	0.0	0.6	0.6	0.0	45.7	63.0	0.0	0.0	0.6
m	0.0	0.0	2.5	0.0	1.2	37.0	9.9	0.0	0.0	0.0
ð	0.6	8.6	4.9	20.4	58.0	1.2	0.0	1.9	2.5	14.2
v	0.0	1.9	16.0	69.1	3.1	0.0	0.0	0.0	3.1	9.3
b	0.0	0.0	49.4	0.6	1.9	0.0	0.0	1.9	0.0	0.0
d	2.5	86.4	4.9	0.0	10.5	0.0	0.0	0.0	0.0	0.0
g	88.9	0.6	6.2	0.6	8.0	0.6	2.5	0.0	1.2	0.0

Same

Response	g	d	b	v	ð	m	ŋ	p	f	θ
ʃ	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	3.1
s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
t	0.0	0.0	0.6	0.0	0.6	0.0	0.0	1.9	0.0	0.0
k	0.0	0.0	0.0	0.0	0.0	0.0	0.6	21.0	0.0	0.6
θ	0.0	1.9	1.2	4.3	8.6	0.6	0.0	3.1	27.8	47.5
f	0.0	0.0	1.2	8.0	1.2	0.0	0.6	3.7	55.6	22.8
p	0.0	0.0	13.0	0.0	0.0	0.0	0.0	66.7	0.6	4.9
z	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0	0.0	1.9
n	1.9	0.6	0.6	0.0	3.1	7.4	22.8	0.0	0.0	0.0
ŋ	2.5	0.0	0.0	0.0	0.0	46.3	58.6	0.0	0.0	0.0
m	0.0	0.0	1.9	0.0	0.6	43.8	11.7	0.0	0.0	0.0
ð	0.6	7.4	3.7	18.5	61.1	0.0	0.0	1.2	4.9	6.8
v	1.2	0.0	11.1	68.5	3.1	1.2	3.7	0.6	11.1	11.1
b	0.6	1.9	63.0	0.6	0.0	0.0	0.0	1.2	0.0	0.0
d	1.2	87.7	1.9	0.0	14.2	0.0	0.0	0.6	0.0	0.0
g	92.0	0.6	1.9	0.0	6.2	0.0	1.9	0.0	0.0	0.6

Different

Response	g	d	b	v	ð	m	ŋ	p	f	θ	
ʃ	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	1.9	ʃ
s	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	s
t	0.0	0.6	0.0	0.0	0.6	0.0	0.0	1.9	0.0	0.0	t
k	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.7	0.0	1.2	k
θ	0.6	0.6	0.6	6.8	8.0	0.0	0.6	8.6	30.9	50.6	θ
f	0.0	0.6	0.6	4.3	0.0	0.0	0.6	3.1	56.2	17.3	f
p	0.0	0.0	5.6	0.0	0.0	0.0	0.0	66.0	0.0	0.0	p
z	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	z
n	0.0	0.6	0.0	0.6	4.3	19.1	27.8	0.0	0.0	0.0	n
ŋ	6.8	0.0	0.0	0.0	0.6	38.3	55.6	0.0	0.0	0.0	ŋ
m	0.0	0.0	1.9	0.0	0.0	42.0	10.5	0.0	0.0	0.0	m
ð	0.6	9.9	3.7	30.9	61.7	0.0	0.0	1.2	3.1	16.0	ð
v	0.0	0.6	14.8	56.8	3.7	0.0	1.9	0.0	9.9	10.5	v
b	0.0	0.0	63.6	0.6	0.6	0.0	0.6	2.5	0.0	0.0	b
d	1.2	85.2	2.5	0.0	14.8	0.0	0.0	0.0	0.0	0.0	d
g	90.7	1.2	6.8	0.0	5.6	0.0	2.5	0.0	0.0	1.9	g

Target Rev: R2

Anechoic

Response	g	d	b	v	ð	m	ŋ	p	f	θ
ʃ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.7
s	0.0	0.0	1.2	0.0	0.6	0.0	0.0	0.0	1.2	3.1
t	0.0	5.6	1.2	0.0	0.0	0.0	0.0	8.0	0.0	0.6
k	4.3	0.0	0.0	0.0	0.0	0.0	0.6	14.8	0.6	0.6
θ	0.6	0.0	5.6	3.7	9.3	0.0	0.0	0.6	9.3	35.2
f	0.0	0.0	3.1	6.2	0.6	0.0	0.0	4.9	81.5	37.0
p	0.0	0.0	12.3	0.0	0.0	0.0	0.6	64.2	0.0	1.2
z	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0	2.5
n	2.5	0.0	1.2	0.0	2.5	13.0	16.0	0.0	0.0	0.0
ŋ	4.3	0.0	0.0	0.6	0.0	48.1	79.0	0.0	0.0	0.6
m	0.0	0.6	2.5	0.0	0.0	38.9	1.9	0.0	0.6	0.0
ð	0.0	1.2	6.2	16.7	61.1	0.0	0.6	0.6	1.2	4.3
v	0.0	1.2	19.8	72.8	5.6	0.0	0.6	1.2	5.6	1.9
b	0.0	1.2	42.0	0.0	0.6	0.0	0.0	4.9	0.0	0.0
d	0.6	85.8	0.6	0.0	13.6	0.0	0.0	0.0	0.0	1.2
g	87.7	4.3	4.3	0.0	4.3	0.0	0.6	0.6	0.0	0.0

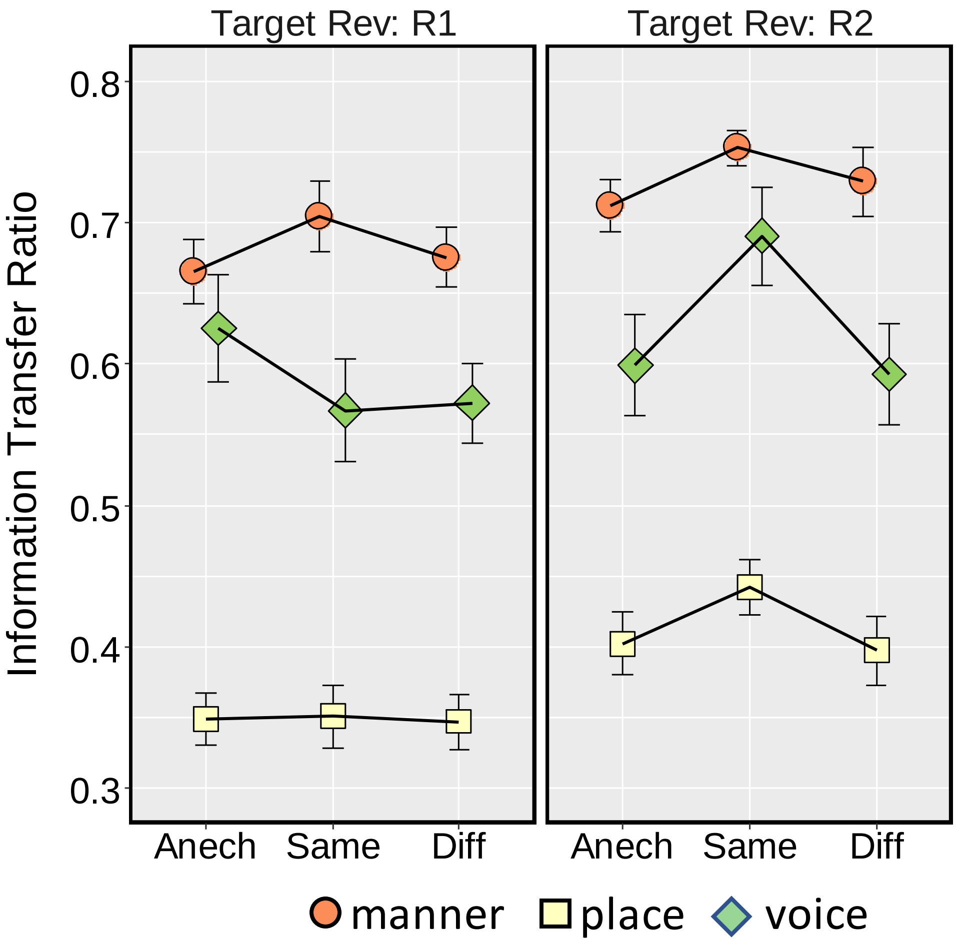
Same

Response	g	d	b	v	ð	m	ŋ	p	f	θ
ʃ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	13.0
s	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1
t	0.0	1.9	0.0	0.0	0.0	0.0	0.0	5.6	0.0	0.0
k	0.6	0.0	0.6	0.0	0.0	0.0	0.0	4.3	0.0	0.6
θ	0.0	0.0	1.9	2.5	9.9	0.0	0.0	2.5	7.4	46.3
f	0.0	0.0	4.3	6.2	0.6	0.0	0.6	1.2	87.7	28.4
p	0.0	0.0	9.9	0.0	0.0	0.0	0.0	80.9	0.0	1.2
z	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
n	1.2	0.0	0.6	0.0	0.0	11.1	14.8	0.0	0.0	0.0
ŋ	2.5	0.0	0.0	0.0	0.0	35.8	74.1	0.0	0.0	0.0
m	0.0	0.0	1.9	0.0	0.0	52.5	8.6	0.0	0.0	0.0
ð	0.0	0.0	3.1	18.5	58.6	0.6	0.0	1.9	0.6	3.1
v	0.0	0.0	15.4	72.8	5.6	0.0	0.6	0.6	4.3	3.1
b	0.6	0.6	57.4	0.0	1.9	0.0	0.0	3.1	0.0	0.0
d	1.2	92.0	0.6	0.0	19.8	0.0	0.0	0.0	0.0	0.0
g	93.8	4.9	4.3	0.0	3.7	0.0	1.2	0.0	0.0	0.6

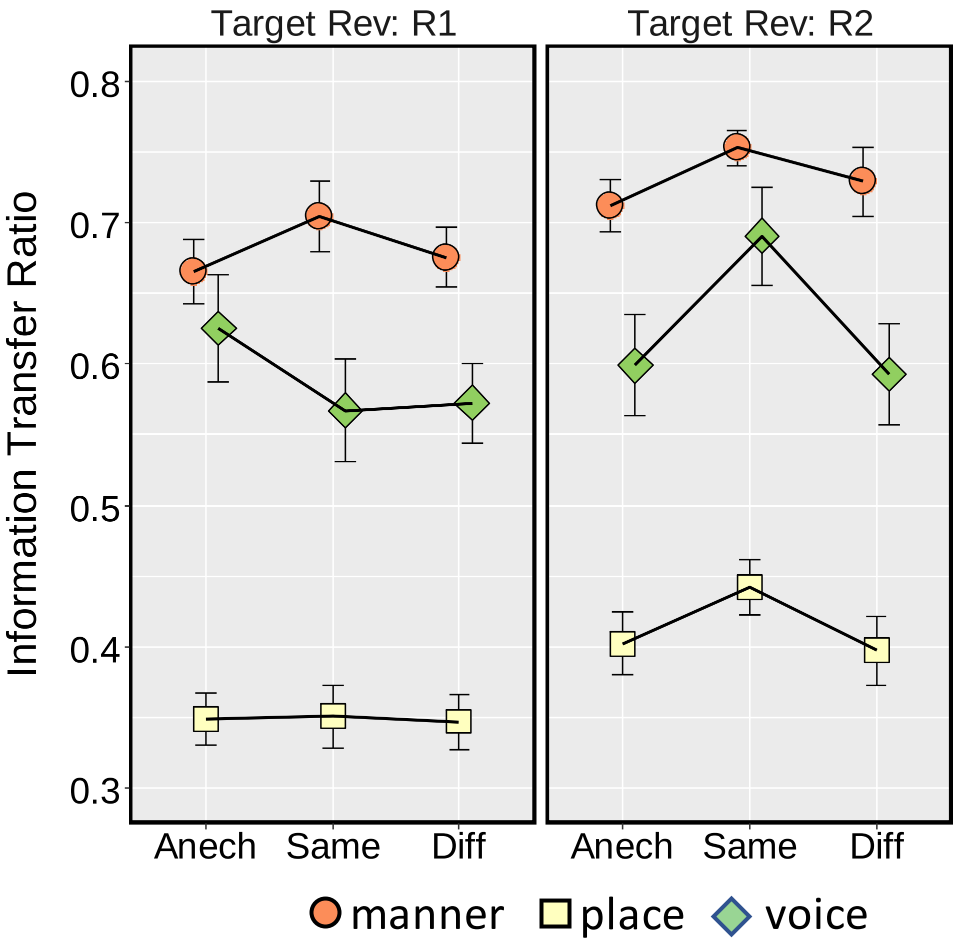
Different

Response	g	d	b	v	ð	m	ŋ	p	f	θ	
ʃ	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	14.8	ʃ
s	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.6	3.7	s
t	0.0	6.8	0.6	0.0	0.0	0.0	0.0	9.3	0.0	0.0	t
k	4.9	0.0	2.5	0.0	0.0	0.0	0.0	6.2	0.0	1.9	k
θ	0.0	0.0	1.9	4.9	13.0	0.0	0.0	1.9	14.8	30.9	θ
f	0.0	0.0	2.5	6.2	0.6	0.6	1.2	1.2	75.3	35.2	f
p	0.6	0.0	21.6	0.0	0.0	0.0	1.9	79.0	0.6	3.7	p
z	0.0	0.0	0.0	0.6	4.3	0.0	0.0	0.0	0.0	1.9	z
n	1.2	0.0	3.7	0.0	0.6	14.2	13.0	0.0	0.0	0.0	n
ŋ	0.6	0.0	0.0	0.0	0.0	48.8	79.6	0.0	0.0	0.0	ŋ
m	0.0	0.0	4.3	0.0	0.0	34.6	2.5	0.0	0.0	0.6	m
ð	0.0	0.6	3.1	11.1	48.8	0.0	0.0	0.6	1.9	4.3	ð
v	0.0	0.0	9.9	75.3	4.9	1.9	0.0	0.0	6.2	3.1	v
b	1.9	0.6	42.0	0.0	0.0	0.0	0.0	1.9	0.6	0.0	b
d	0.6	87.0	1.2	0.6	22.8	0.0	0.0	0.0	0.0	0.0	d
g	90.1	4.9	6.8	0.6	4.9	0.0	1.2	0.0	0.0	0.0	g

Pooled Data – Information Transfer Ratio

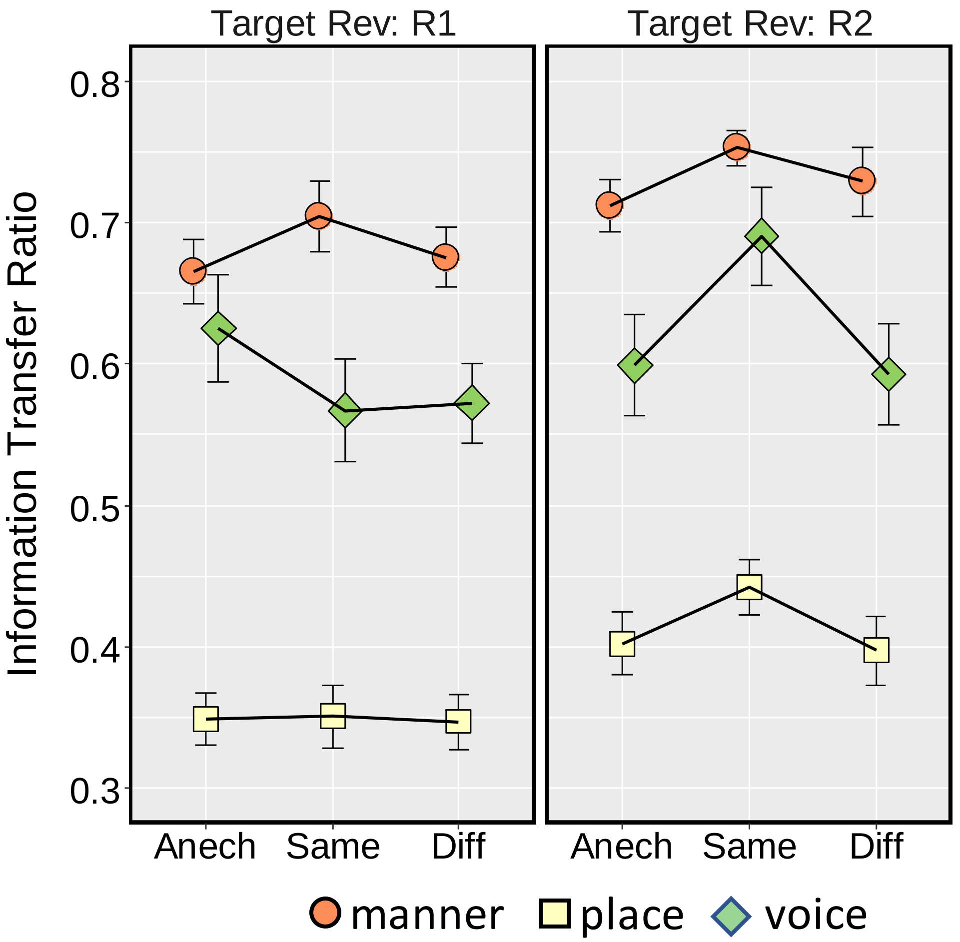


Pooled Data – Information Transfer Ratio



- Main effect of Feature ($F_{(2,34)}$ = 33.40, p < 0.0001),
- Target Reverberation ($F_{(1,17)}$ = 11.95, p = 0.003)
- Carrier Type ($F_{(2,34)}$ = 5.37, p = 0.013)
- Significant interactions
  - Carrier Type X Target Rev ($F_{(2,34)}$ = 8.83, p = 0.0012)
  - Carrier Type X Target Rev X Feature ($F_{(4,68)}$ = 4.14, p = 0.0140)

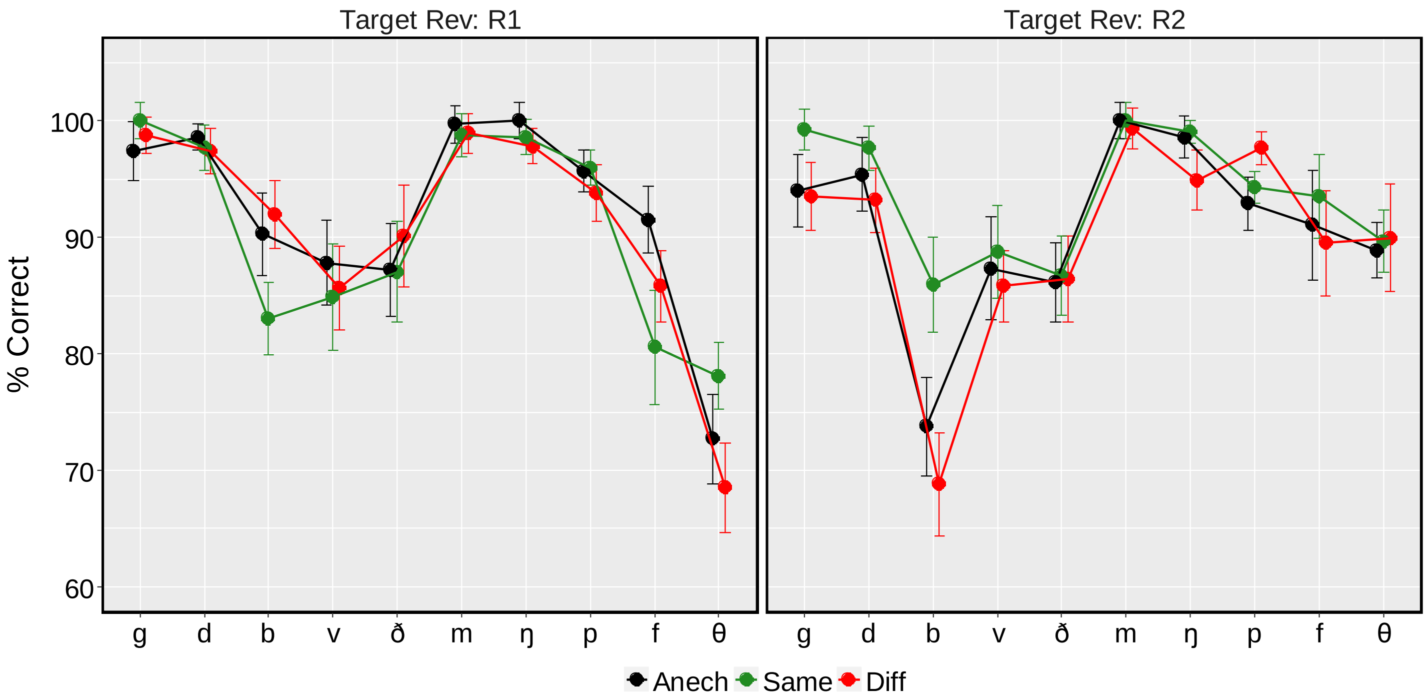
Pooled Data – Information Transfer Ratio



- Main effect of Feature ($F_{(2,34)}$ = 33.40, p < 0.0001),
- Target Reverberation ($F_{(1,17)}$ = 11.95, p = 0.003)
- Carrier Type ($F_{(2,34)}$ = 5.37, p = 0.013)
- Significant interactions
  - Carrier Type X Target Rev ($F_{(2,34)}$ = 8.83, p = 0.0012)
  - Carrier Type X Target Rev X Feature ($F_{(4,68)}$ = 4.14, p = 0.0140)

Improved **phonetic feature perception** after exposure to **consistent** reverberation

For **manner** or articulation effect more **resistant** in more challenging listening environments. For other features, it might be diminished



Performance on voicing dimension across the different phonemes, as a function of Carrier Type, separately for Target Reverberation R1 (left panel) and R2 (right panel).

Summary and Conclusions

Replicating and extending previous work, we showed…

  - strong adaptation to reverberation after exposure to a consistent room
  - breakdown of adaptation after exposure to a different environment (anechoic or different reverb)
  - rapid effect (after ~1600 ms an additional exposure of ~1600ms did not improve performance)
  - Effect largely independent of prior knowledge of target speech location

- The effect generalizes across different phonemes, with diverse spectrotemporal properties

- The effect varies substantially across different rooms and phonemes, and may be diminished for some sounds & phonetic features in very challenging listening environments