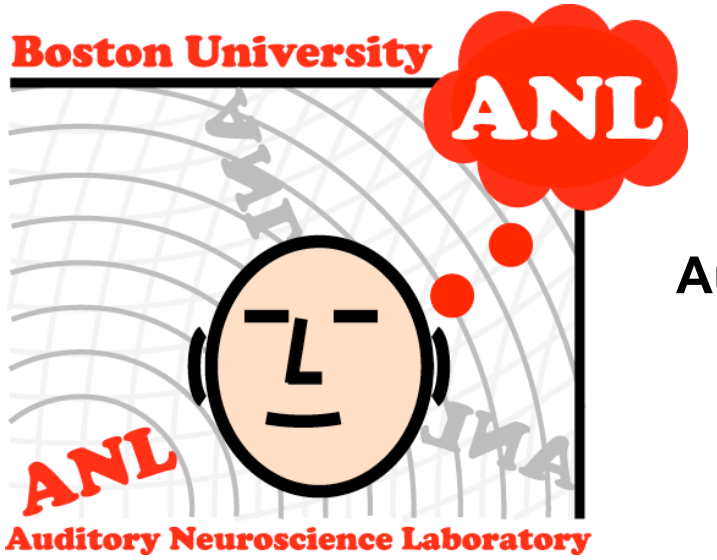
The costs of switching auditory attention

Prof. Barbara Shinn-Cunningham



Boston University
Hearing Research Center
Auditory Neuroscience Laboratory

Departments of
Cognitive and Neural Systems
Biomedical Engineering

Harvard-MIT Division
of Health Sciences and Technology

Acknowledgments

Funding from ONR, AFOSR, NSF, NIH

Post-docs / staff

(Gin Best)
Lorraine Delhorne
Norbert Kopco
Erol Ozmeral
Tim Streeter

Grad students

Scott Bressler
Liangqiang Kong
I-Fan Lin
Ross Maddox
Dorea Ruggles
Satyavarta

Andrew Schwartz
Eric Thompson
Lenny Verghese
Dali Wang
Jing Xia

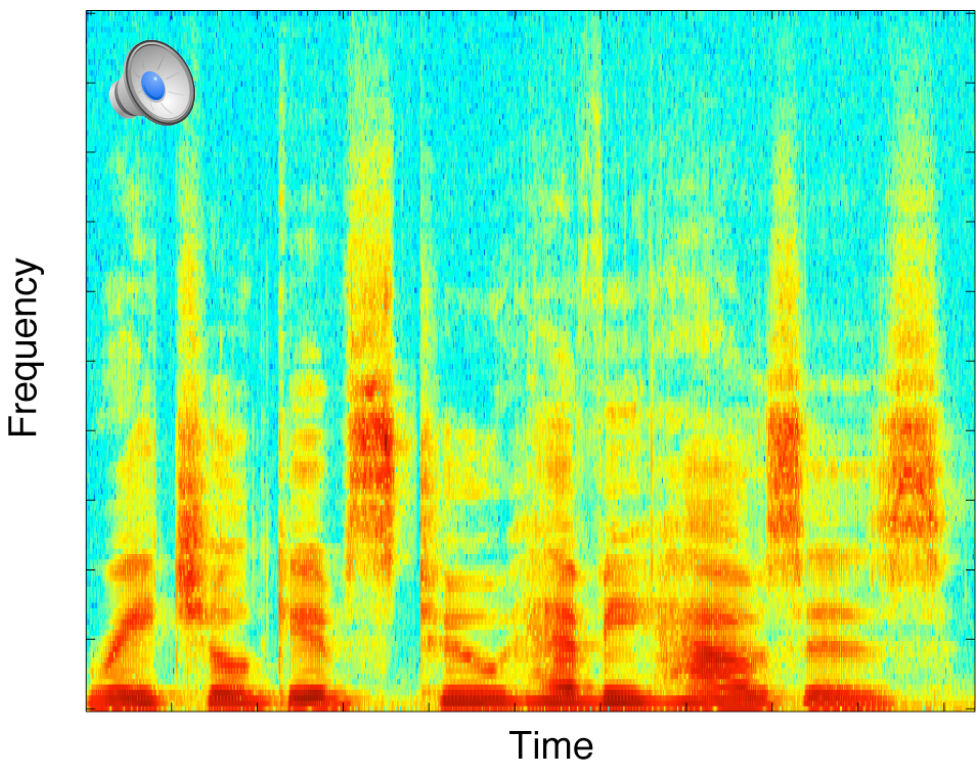
Undergrads

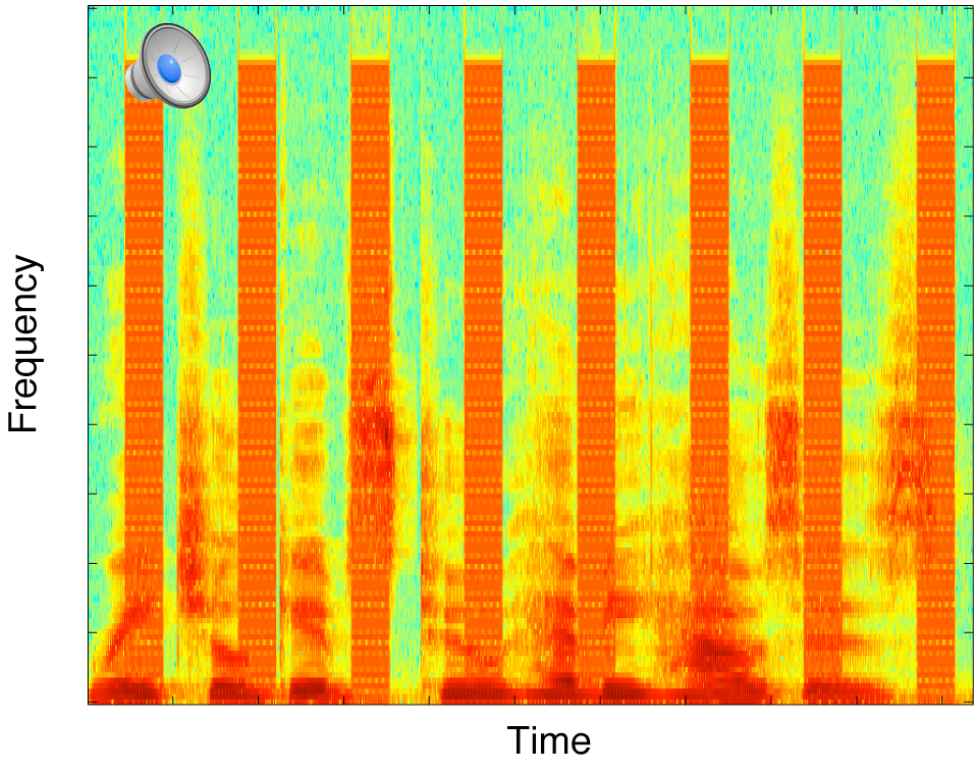
Brendan Hurd
Chloe McGuffin

Auditory objects are mainly determined by time-frequency relationships

Spectro-temporal structure determines "grouping" (object formation) as well as meaning

(Spatial cues only weakly influence "object" formation)



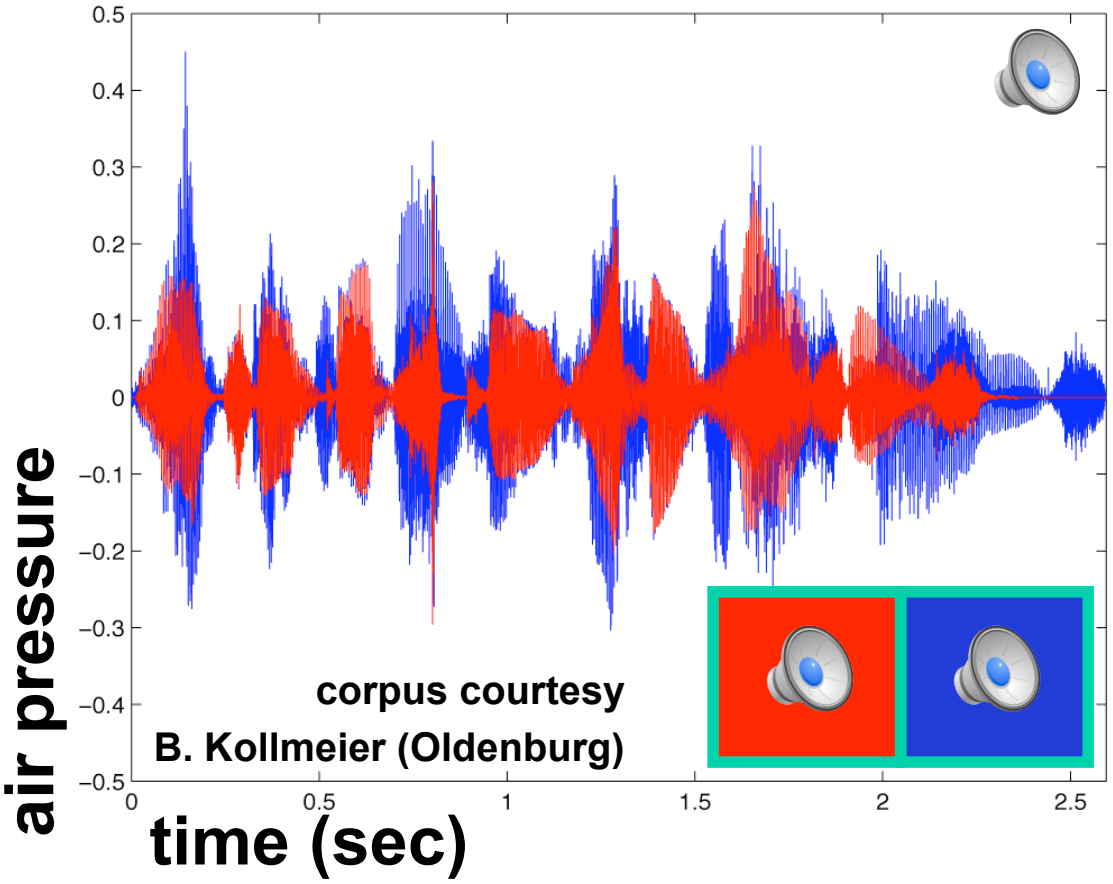


Syllables / words are heard; confusions occur between sources

Listen to the sentence starting with "Rachel"

Syllables are formed because of spectro-temporal structure

Need some feature to attend the correct syllables in the mixture



Features enable top-down selection

See Ihlefeld & S-C, JASA (accepted-B); Maddox et al., ARO

Listeners attend to a source based on some attribute or feature

Listen for the telephone number from the male, metallic voice

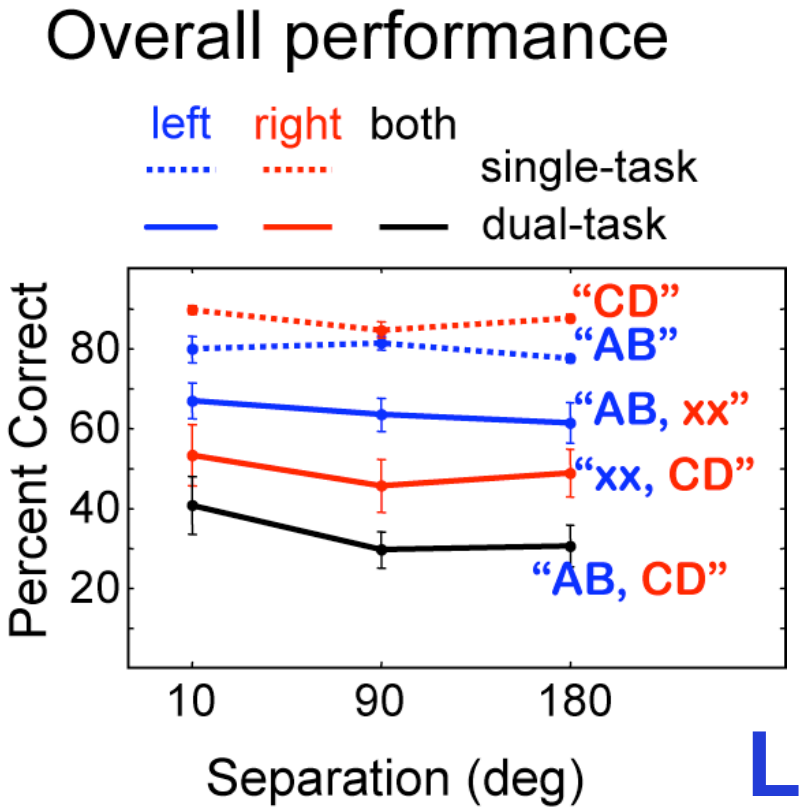
Because the two talkers sound different, there is little problem hearing out the number…

BUT WHAT WAS THE OTHER SIGNAL?

We naturally attend to one and only one object at a time

See Best et al., JASA (2006); Ihlefeld & S-C, JASA (2008)

Spatial separation hurts performance on a divided-attention task

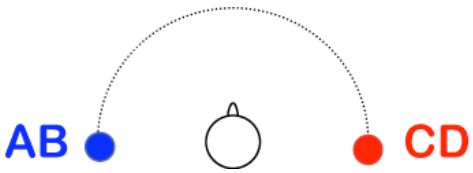


Selective performance equated (left-right asymmetry favors right source)

Divided performance is worse and decreases slightly with separation

Performance is

- **good for "report first" left**
- **worse for right**



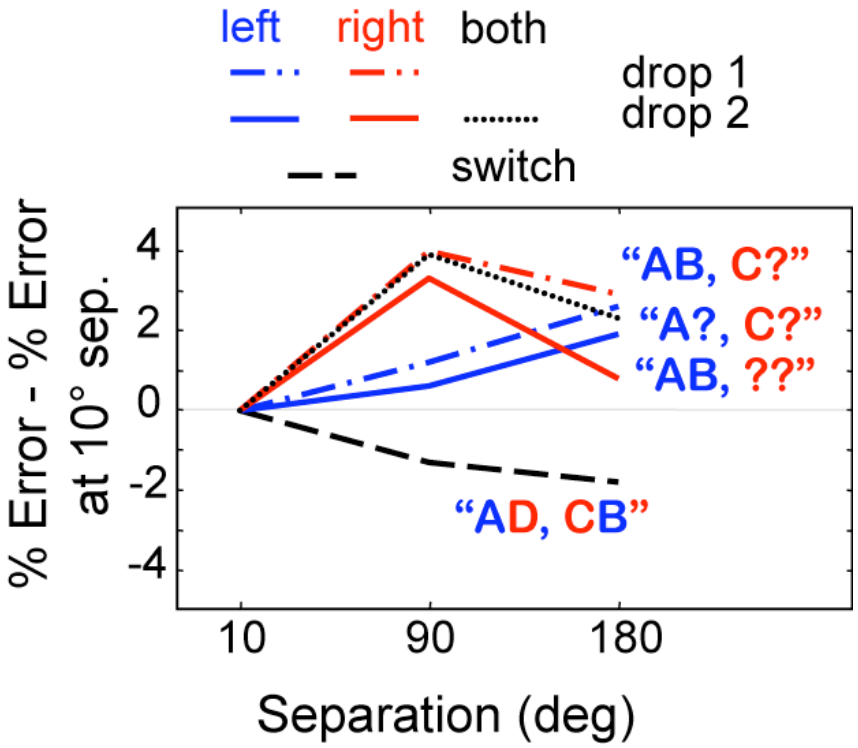
Listeners prioritize streams in divided attention (switching, not dividing?)

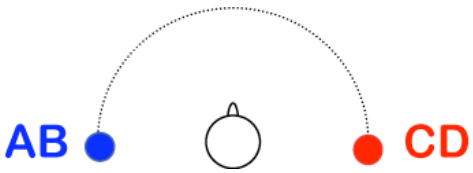
see also Ihlefeld & S-C, JASA (2008)

Best et al., JARO (2006)

Listeners are less likely to confuse messages as separation increases

Dual-task errors





Separation *increases* distinctiveness of competing streams

Listeners are more and more likely to *miss* part of one stream as they switch attention back and forth

Listeners are more and more likely to hear one (and only one) at a time

Best et al., JARO

In divided attention tasks, spatial separation has two effects

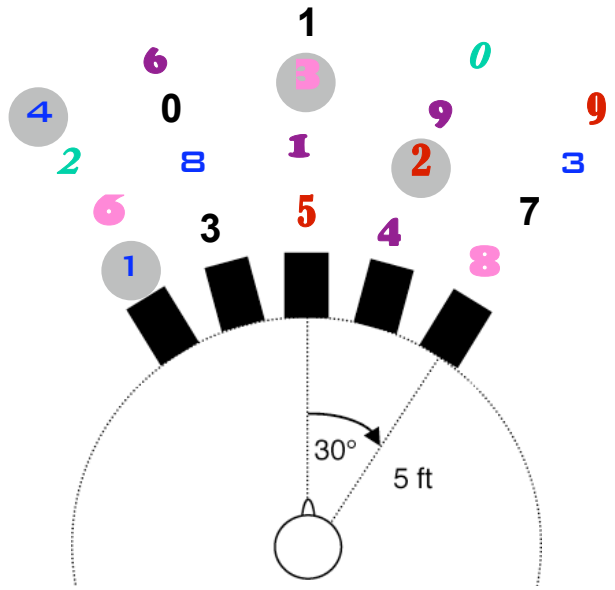
Listeners

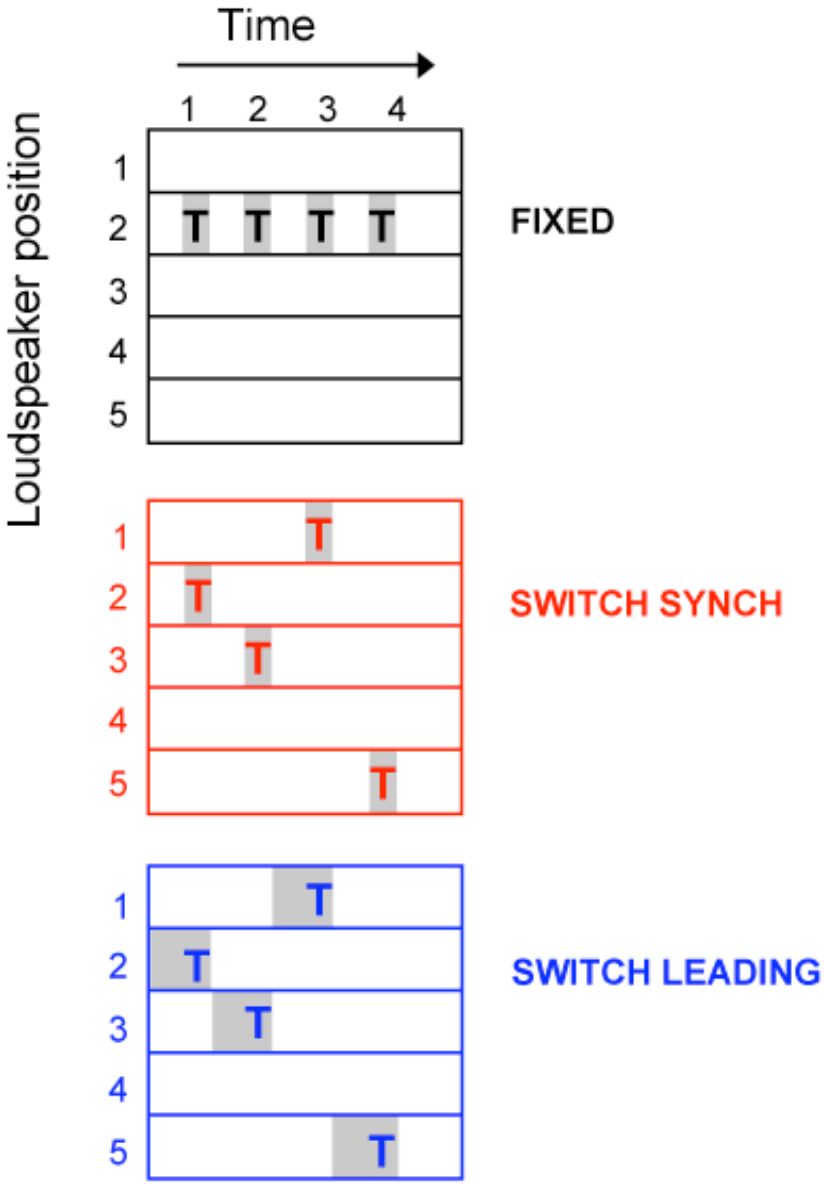
- **attend to one source, suppressing the other**
- **then switch attention to what was in the background**
  - **(see also Ihlefeld and Shinn-Cunningham, JASA, 2008)**

With spatial separation, more complete

- **suppression of background => more words dropped**
- **object formation => fewer confusions across streams**
  - **(see also Xia et al., ARO; Maddox et al, ARO)**

What is the cost of switching attention?





Report 4 digit sequence

Five loudspeakers, each on simultaneously

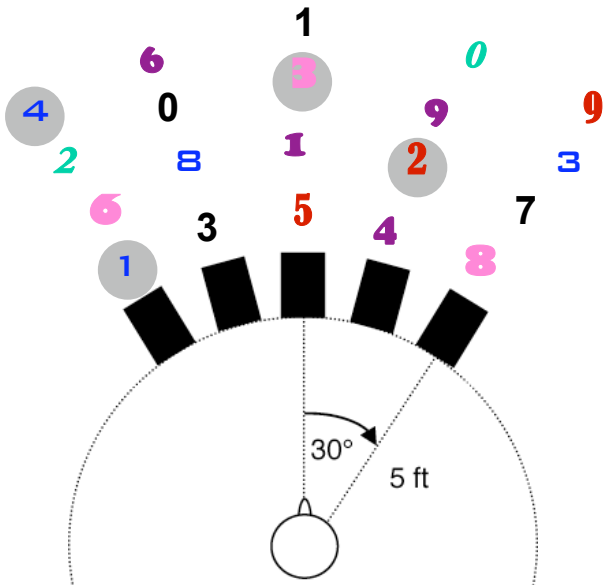
Light on loudspeaker directs spatial attention

Ozmeral et al., ARO

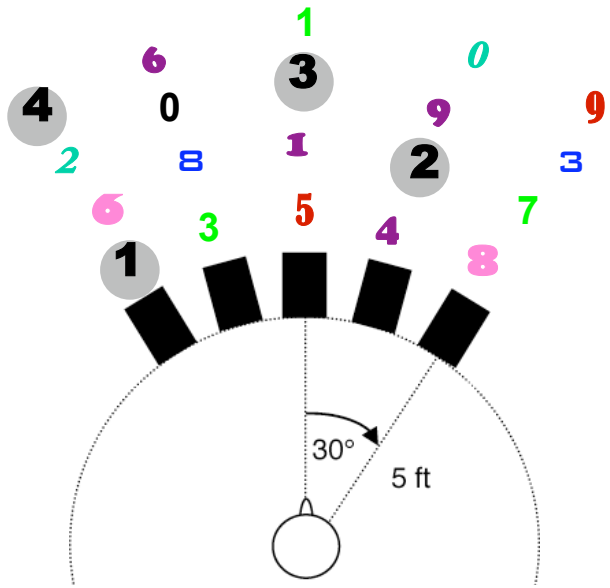
Examine effects of inter-digit delay, voice continuity

Hypothesize that at longer delays, sufficient time to

- **dis- and re-engage attention**
- **recover from cost of switching attention**



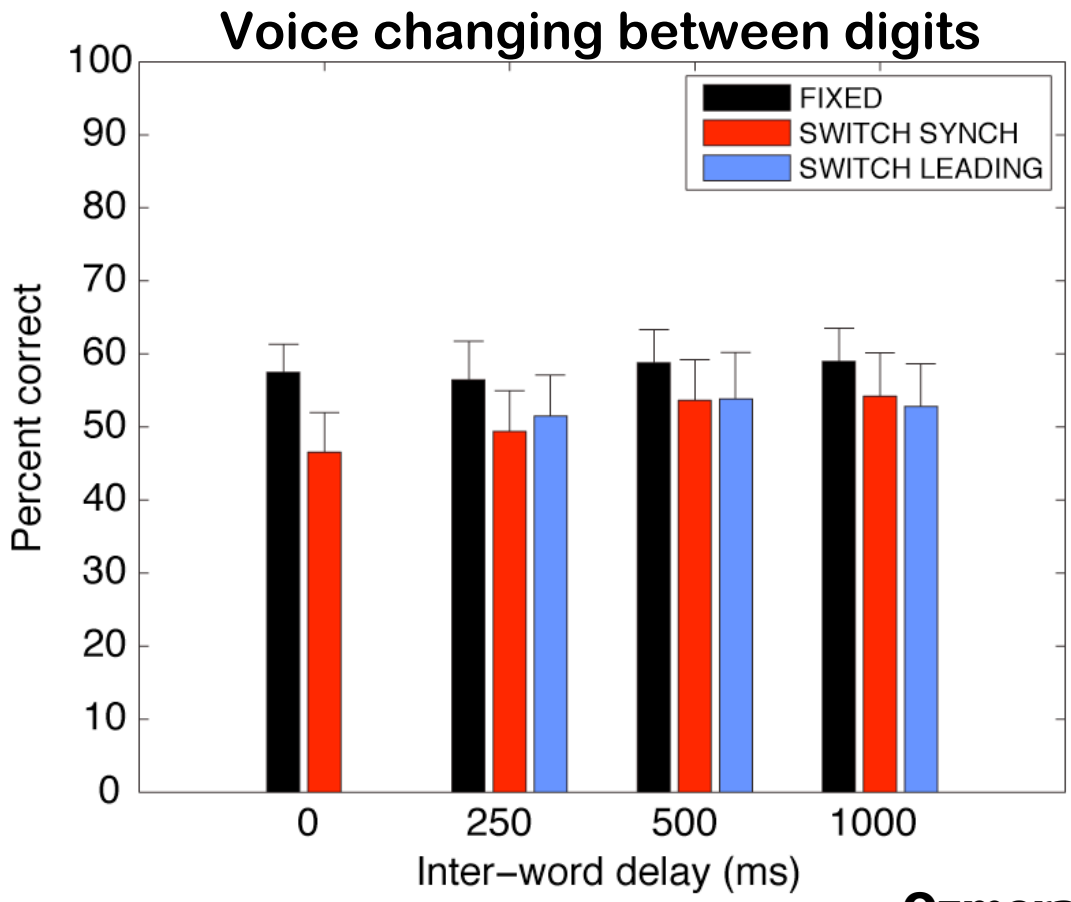
voice changing between digits



voice continuous between digits

Ozmeral et al., ARO

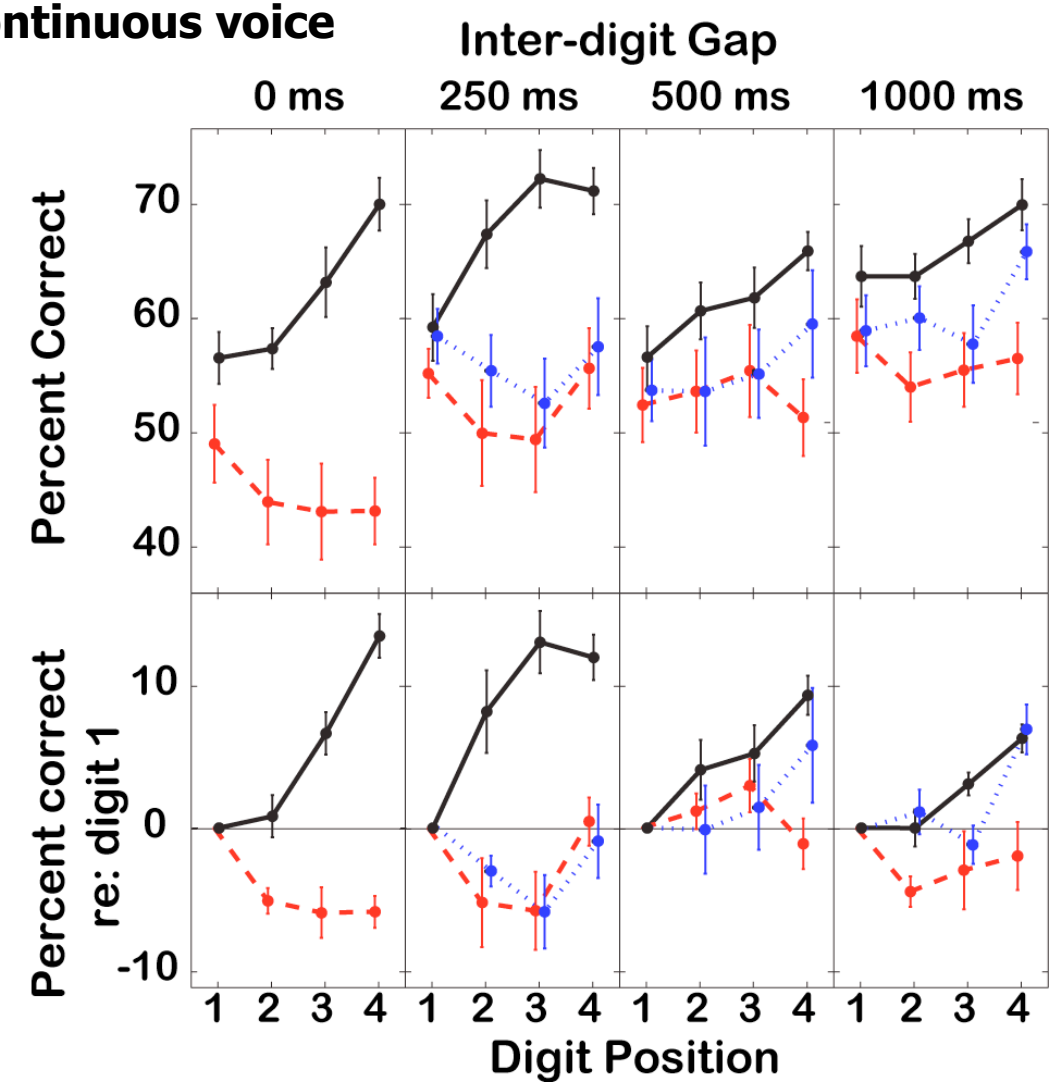
Advance knowledge (up to one sec!) does not improve performance

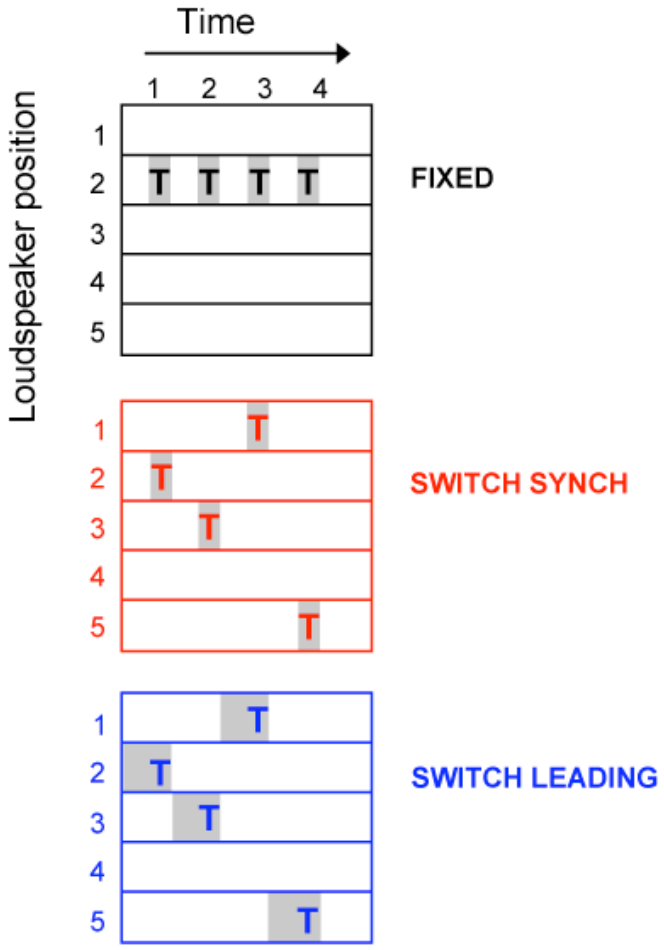


Ozmeral, et al., ARO, 2008

Advance knowledge of where to listen does not improve performance

Continuous voice

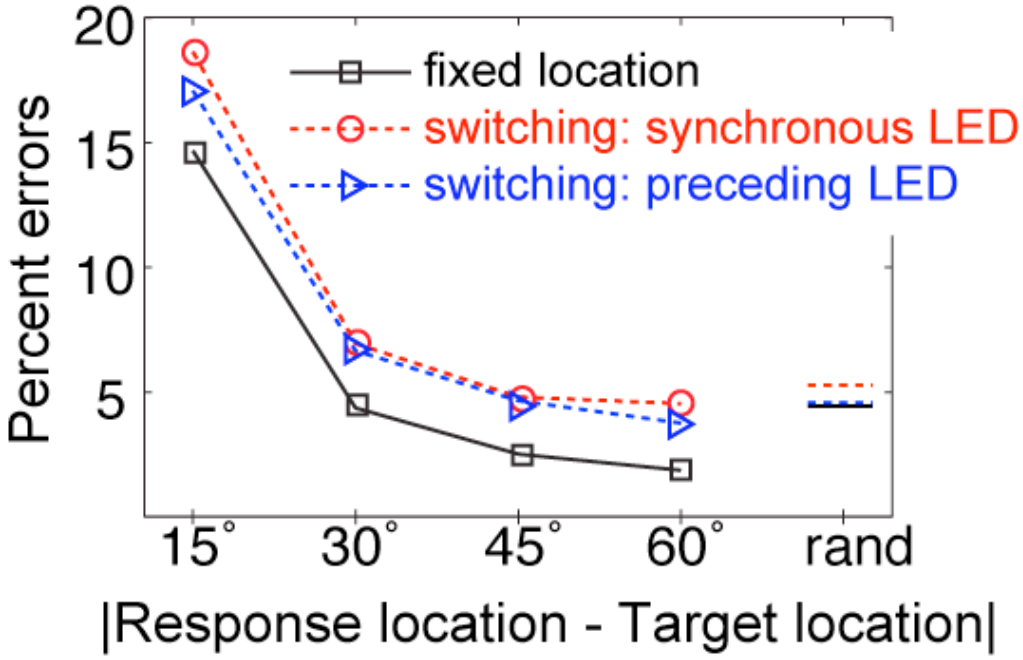


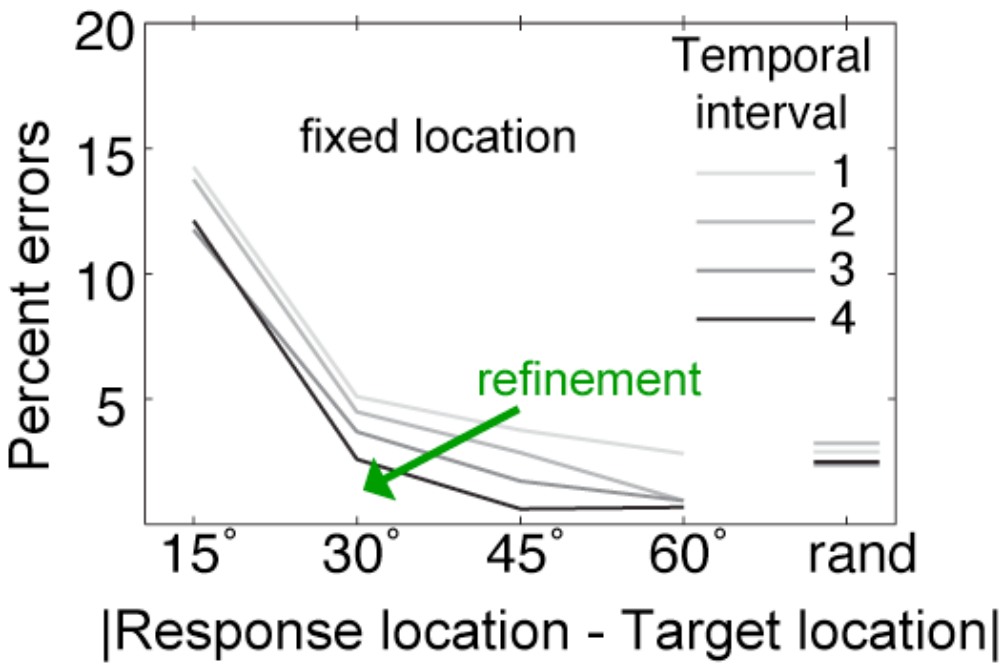


Ozmeral et al., ARO

This spatial filter becomes more refined from digit to digit when location is fixed

Continuous voice





Errors correspond to reporting digits from locations near the target

Errors grow fewer and become more tightly clustered in space

Switching attention is costly

Maintaining attention allows the listener to refine their "attentional filter"

Disengaging and re-engaging attention requires time

Advance knowledge of where to attend does not help much, as it does not restore refinement of attention

Features that make the stream cohere lead to refinement (continuity in location, voice quality, timing)

Attentional refinement may

- play a key role in real-world listening (hearing impairment: slower, bigger costs?)

- affect visual perception in complex scenes