Spatial Release from Speech-on-Speech Masking in Reverberation Evaluated Using the Portable Automated Rapid Test

Udbhav Singhal[1,3], Marusa Laure[2,3], Alyson Ditmore[4],
William C. Bologna[4], Aaron R. Seitz[3], Frederick J. Gallun[6], G. Christopher Stecker[5], Norbert Kopco[1]

1. Institute of Computer Science, P.J. Safarik University in Kosice, Slovakia
2. University of Maribor, Slovenia
3. Department of Psychology, Center for Cognitive and Brain Health, Northeastern University, Boston, MA
4. Department of Speech-Language Pathology and Audiology, Towson University, Towson, MD
5. Center for Hearing Research, Boys Town National Research Hospital, Omaha NE
6. Department of Otolaryngology, Oregon Health & Science University, Portland OR

BASH: 11/6/2025

[Work supported by HORIZON-MSCA-2022-SE-01 Grant No. 101129903, APVV-23–0054 and NIH R01 NIH R01-DC018166]

MOTIVATION

- **Spatial hearing** enables humans to extract meaning from complex soundscapes, supporting speech perception, localization, and situational awareness (Yost, 2006; Blauert, 1997; Litovsky et al., 2009).
- This ability relies on **binaural cues** interaural time differences (ITD), interaural level differences (ILD) and **spectrotemporal features** to construct a spatial map of the auditory scene (Wightman & Kistler, 1992; Ellinger et al., 2017; Baltzell et al., 2020; Caplan et al., 2024).
- **Reverberation**, caused by reflections from room surfaces, distorts ITD and ILD cues, reducing speech intelligibility and localization accuracy (Shinn-Cunningham et al., 2005; Kuttruff, 2016; Ihlefeld & Shinn-Cunningham, 2011).
- These distortions impair **spatial release from masking (SRM)** the benefit of spatial separation between target and masker speech by reducing audibility, smearing temporal features, and degrading spatial resolution (Bronkhorst, 2000; Hawley et al., 2004; Misurelli & Litovsky, 2012).
- SRM is typically reduced by ~5 dB in reverberant conditions compared to anechoic settings (Bronkhorst & Plomp, 1992; Lavandier & Culling, 2007; Marrone et al., 2008, Momtaz et al., 2025).

CURRENT STUDY

Main Goal

To investigate how reverberation affects **spatial release from masking (SRM)** for speech stimuli in a complex, multi-talker environment using the **Portable Automated Rapid Testing (PART)** system.

- Primary objective was to isolate and examine the **informational masking** component of SRM by employing a **closed-set corpus** (Coordinate Response Measure, CRM) and **symmetrically placed maskers**, thereby minimizing energetic masking effects.

ADDITIONAL GOALS:

I. **Evaluate Psychometric Function Slopes:** Determine whether reverberation alters the steepness of the psychometric functions differently for **co-located** versus **spatially separated** conditions.

II. Examine whether performance **improves over individual words** in target sentence.

III. **Analyze Confusion Patterns:** Identify how often and in what ways listeners confuse target and masker speech under varying spatial and reverberant conditions

WHY P.A.R.T?

PART: Portable Automated Rapid Testing is a tablet-based tool developed by the **UCR/Northeastern Brain Game Center** to rapidly assess **auditory processing** across multiple dimensions such as **pitch, timing, spatial cues** in both **clinical** and **research** settings.

PART Benefits for SRM (Spatial Release from Masking)

- **High-quality auditory assessments**: PART includes a wide variety of tasks validated for laboratory use, allowing precise measurement of auditory processing abilities.
- **Portable and automated:** Designed to run on tablets and laptops, PART enables testing in diverse environments without requiring expert supervision.
- **Multi-dimensional sound processing:** The tool evaluates how listeners use multiple auditory dimensions, which is essential for understanding SRM in complex listening scenarios.
- **Remote testing capability**: PART supports remote administration, making it easier to study SRM across large and varied populations.

Key Benefits

- Portable and self-guided
- Scientifically validated tasks
- Ideal for aging populations and hearing-impaired users

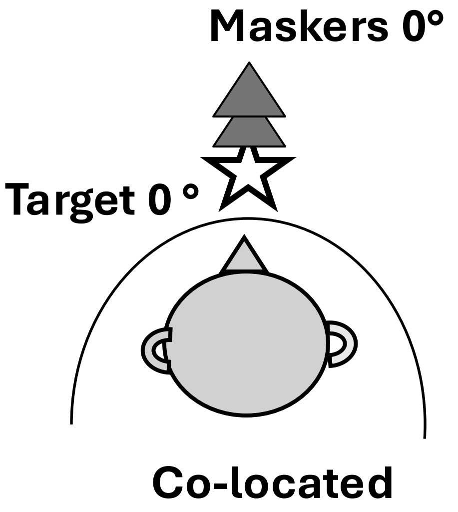
METHODS

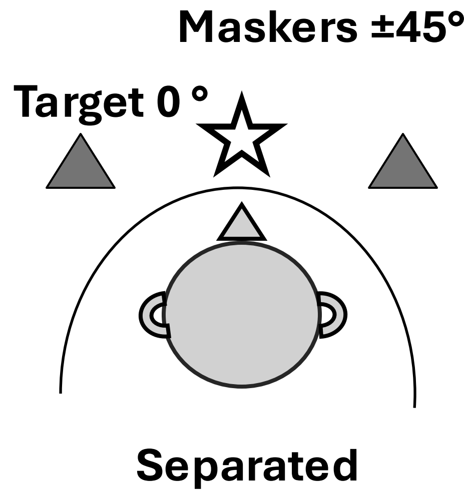
Participant & Apparatus

- **N =10** (ages 22-32, median= 25; 6 female) with **self-reported normal hearing.**
- Coordinate Response Measure (CRM) speech perception task under two acoustic conditions **anechoic** and **reverberant (room size: 10*10m, RT60 = .4s).**
- **Target** always at 0° azimuth; **maskers** at 0° or ±45° (co-located vs. spatially separated) in a multitasker talker configurations.
- The experiment divided into **two sessions** with order counterbalanced: **five participants** (1 male, 4 female) experienced the **anechoic** condition first followed by **reverberant**, while the remaining five completed the study in **reverse order.**

Sessions	Condition A	Condition B	
1	CRM task with reverberation	CRM task without reverberation (i.e. Anechoic)	Participants used Tablet A; completed the CRM task first with reverb, then anechoic
2	CRM task without reverberation (i.e. Anechoic)	CRM task with reverberation	Participants used Tablet B; completed the CRM task first with anechoic, then with reverb

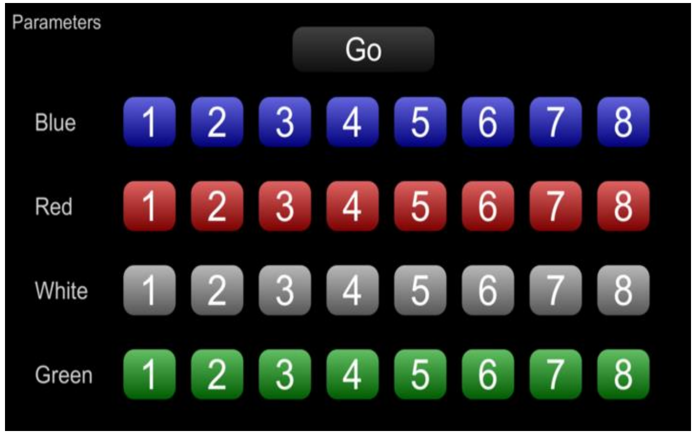
- **Equipment:** Two apple iPad + Sennheiser HD 280 Pro headphones





Stimuli:

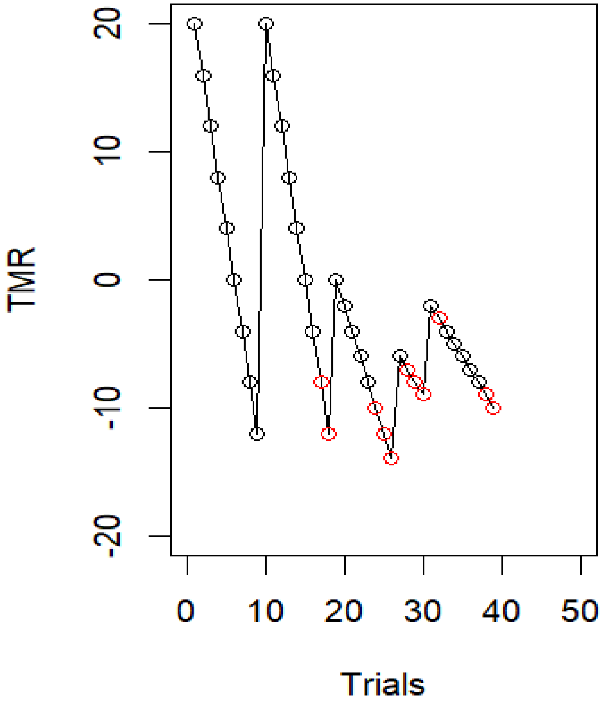
- Each trial included a sentence like: "Ready [**call sign**], go to [**color**] [**number**] now".
- Target talker identified by a unique **call sign** (e.g., "Charlie"), and two simultaneously talkers were presented as maskers.
- Grid displayed **8 digits (1–8)** and **4 colors** (e.g., red, blue, green, white)
- Participants selected the correct **color-digit pair** from a visual grid in P.A.R.T



Adaptive Scans (Lelo de Larrea-Mancera et al., 2024):

Adaptive Scan improves threshold estimation **by sampling across** a wide stimulus range, not just near threshold. It's **faster** and **more robust** than traditional staircases, making it ideal for clinical and remote testing.

- Each participant completed **five scans per condition**: **First two scans are non-adaptive.**
- Each scan (adaptive/non-adaptive) included **9 trials**
- Initial step size is **4dB,** after non-adaptive track it narrows to **2dB,** then **1dB**
- **Termination rule:** scan ended after **3 incorrect responses.**
- Signal and noise adapt with **TMR**s ranging from **20 to -20 dB**, where at least one component **(target or masker) remained at 70 dB SPL** throughout each trial.



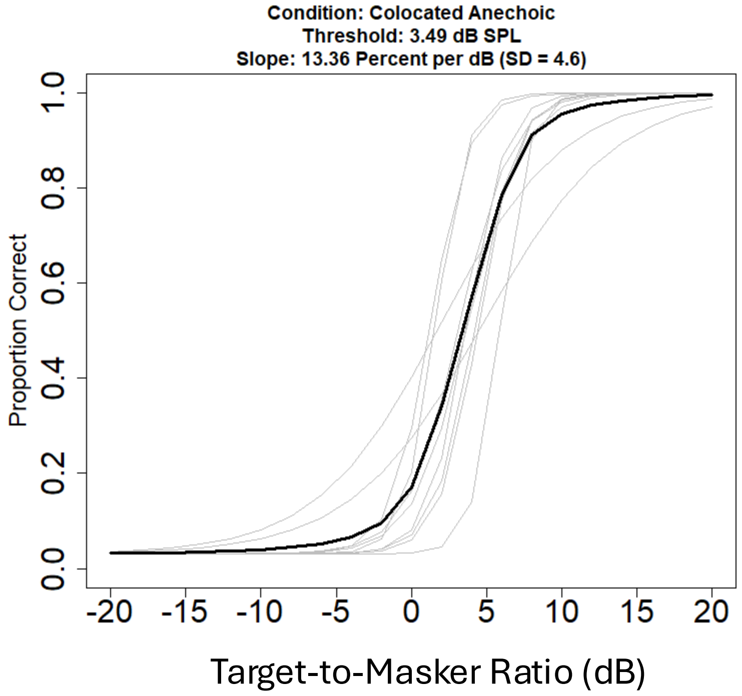
- **Threshold Estimation**= $(hits – nTrials * guessRate)/(1 – guessRate)$
  (Lelo de Larrea-Mancera et al., 2024)

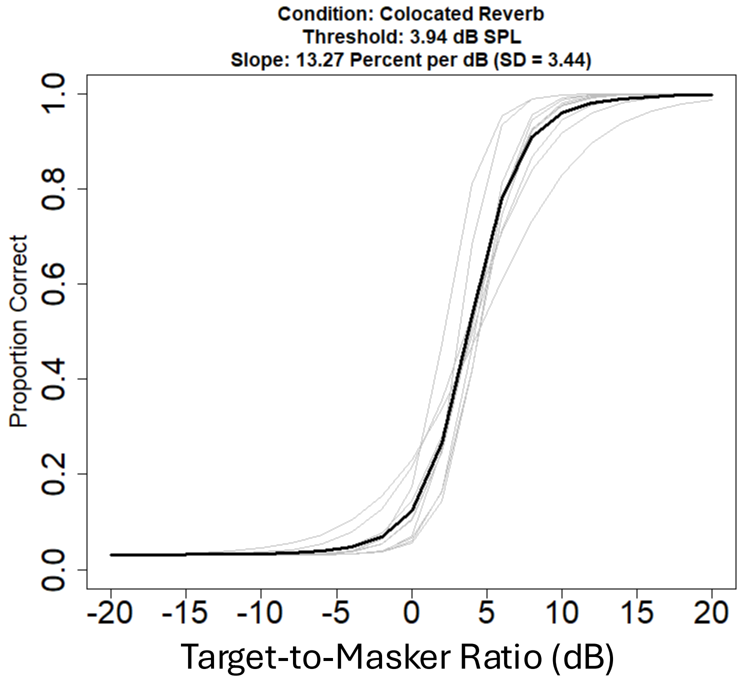
https://braingamecenter.ucr.edu/games/p-a-r-t/#

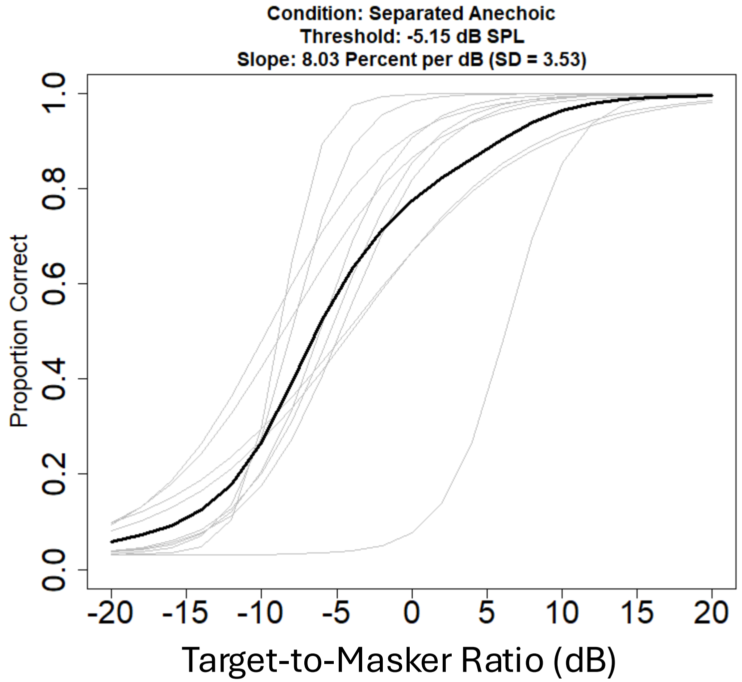
PERFORMANCE MEASURES

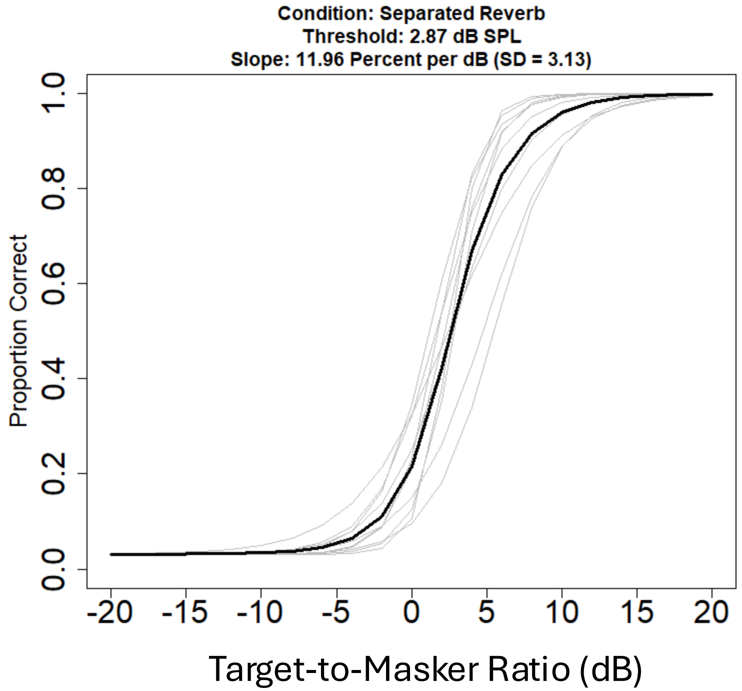
Psychometric functions fitted to data, TMRs and slopes estimated.

Results consistent across participants
(One participant failed to recognize the spatial cues in anechoic separated condition)

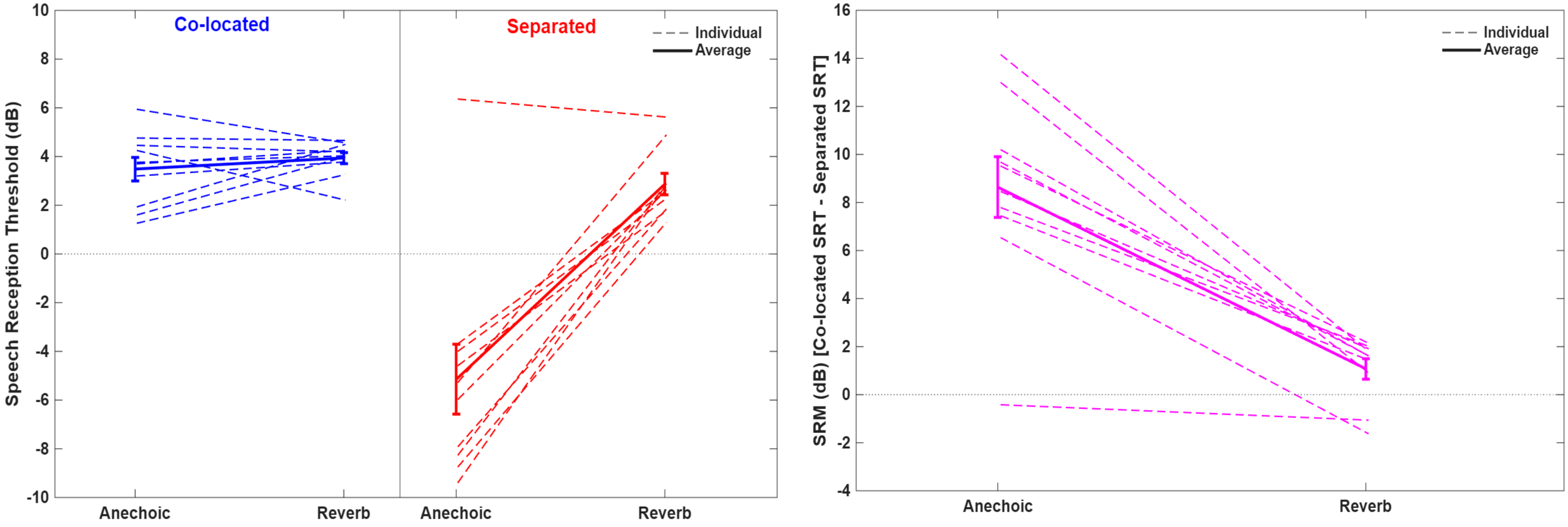








SRT AND SRM



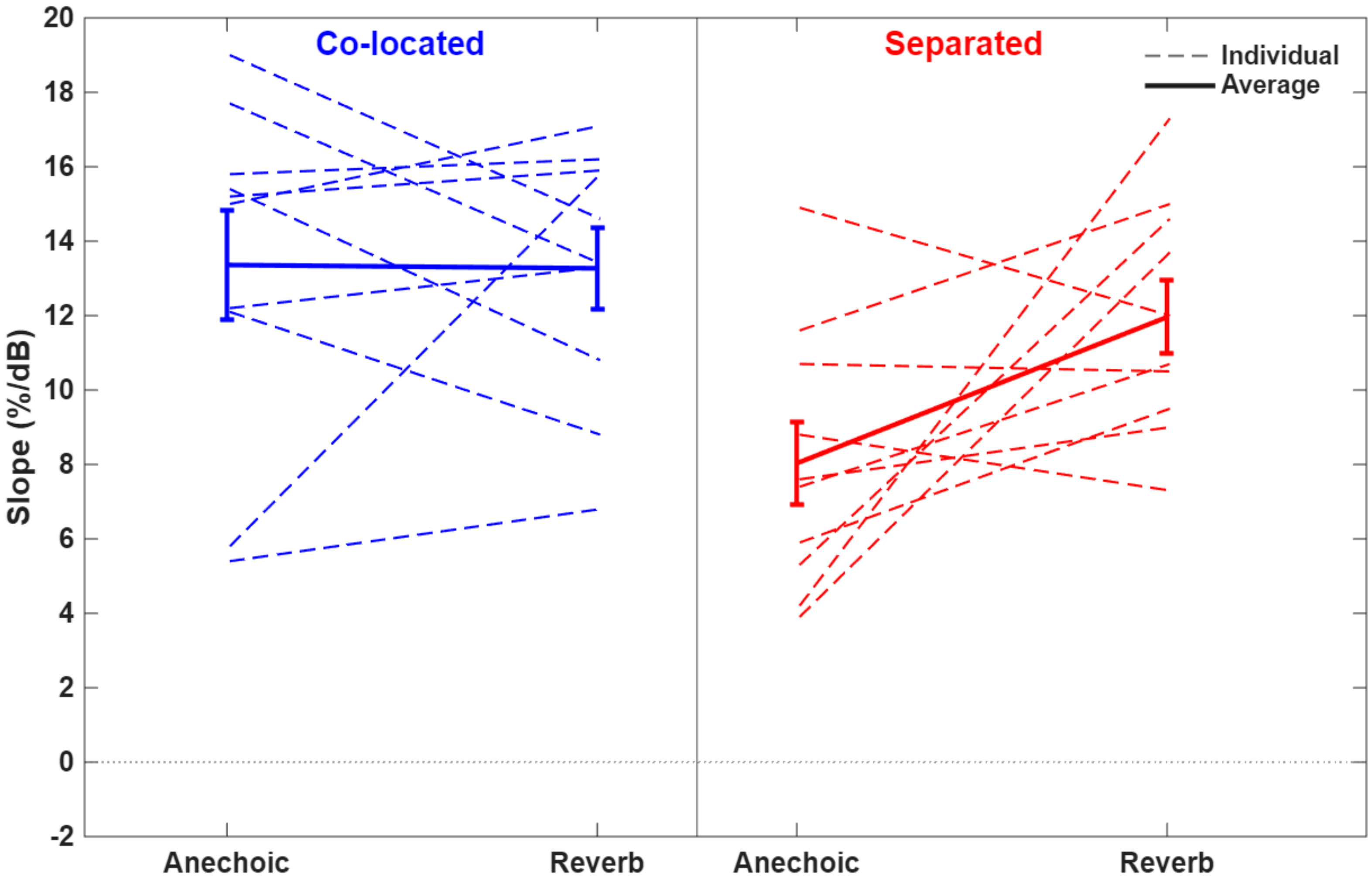
A 2-way ANOVA showed significant effects of **location**, **environment**, and their interaction (***p*** **< .001**). Post-hoc comparisons revealed that **Separated Anechoic** had significantly **lower SRTs** than Co-located Anechoic (**8.64 dB**), Co-located Reverb (**−9.10 dB**), and Separated Reverb (**−8.02 dB**), all ***p*** **< .001.**

SRM was **8.64 dB** in the anechoic environment, reduced to only **1.08 dB** in reverberation .
Spatial Release from Masking (SRM) was significantly greater in anechoic conditions compared to the reverberant (**MeanDiff = 7.56, p<.001**), indicating that spatial separation provided more perceptual benefit in the absence of reverberation

SLOPES OF PSYCHOMETRIC FUNCTIONS

The slopes of the psychometric curves were comparable in the co-located conditions (**anechoic: 13.4 %/dB, reverberant 13.3 %/dB**), becoming shallower in the separated anechoic condition (**8 %/dB vs. 12 %/dB in reverberation**).

A 2-way ANOVA showed a significant main effect of **Location** (co-located, separated) (F = 7.99, p = .007). The average slope across environments was **13.81 %/dB** for co-located and **10 %/dB** for separated condition. Although the slope appeared steeper in the reverberant condition, this difference was **not statistically significant.**



Shallower slopes indicate that non-energetic factors are more dominant in determining the separated threshold

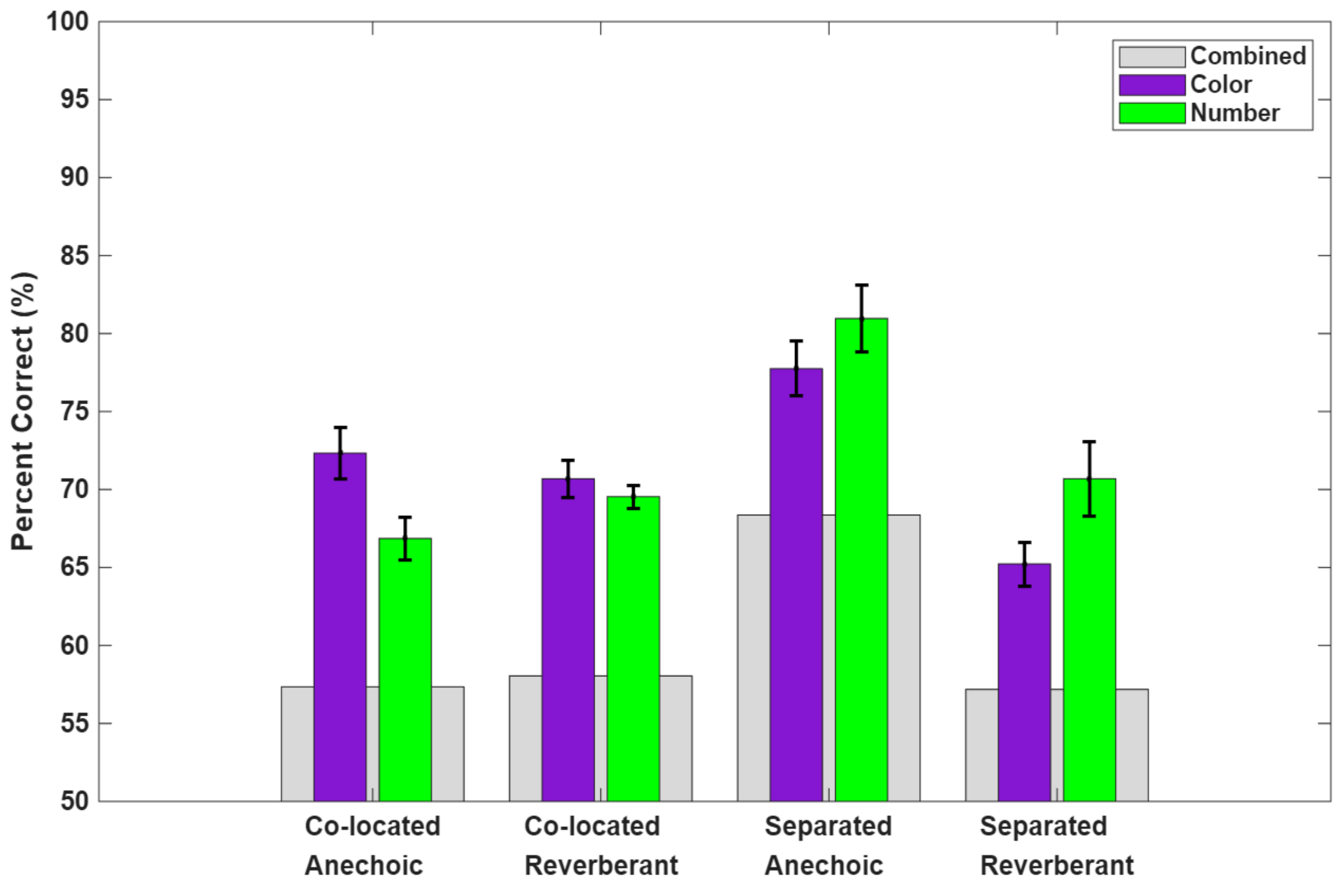
CORRECT RESPONSE RATE FOR INDIVIDUAL WORDS

PCC COLLAPSED ACROSS TMRs, COMPUTED SEPARATELY FOR COLOR AND NUMBER:

3-Way ANOVA: The **item type** (**color vs. number**) demonstrated significant interactions with both **environment** (anechoic, reverberant) (F(7.669), p = .021) and **location** (co-located, separated) (F(6.02), p = .036).

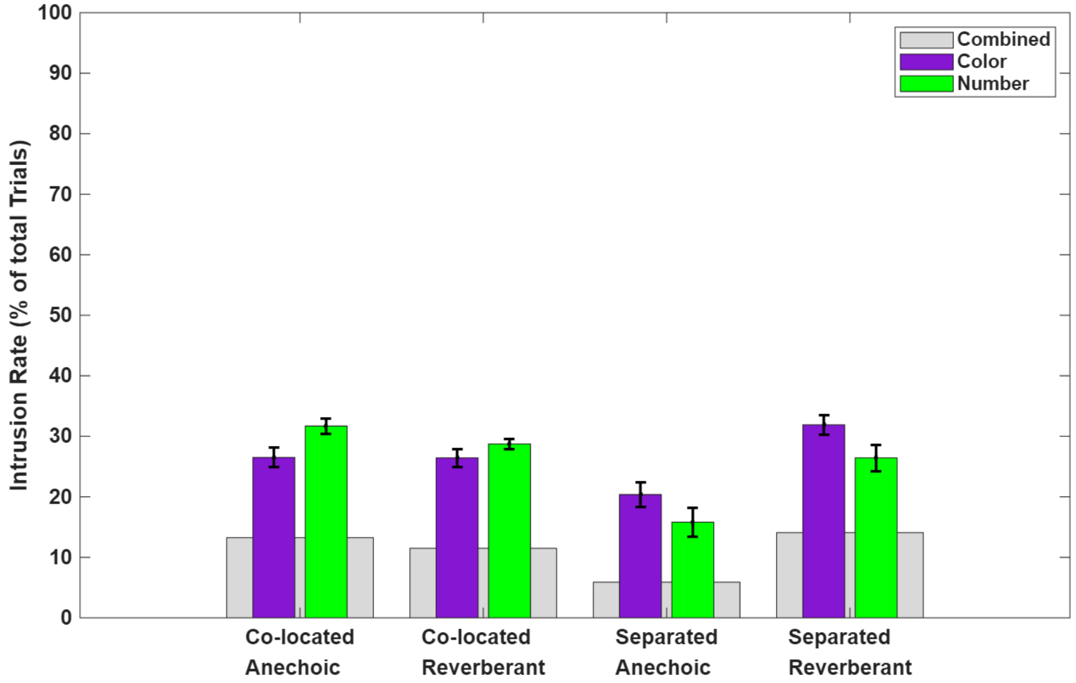
Number performed better than color for separated (both Anechoic and Reverberant)
Number performed worse or equal to color for collocated (mainly anechoic)

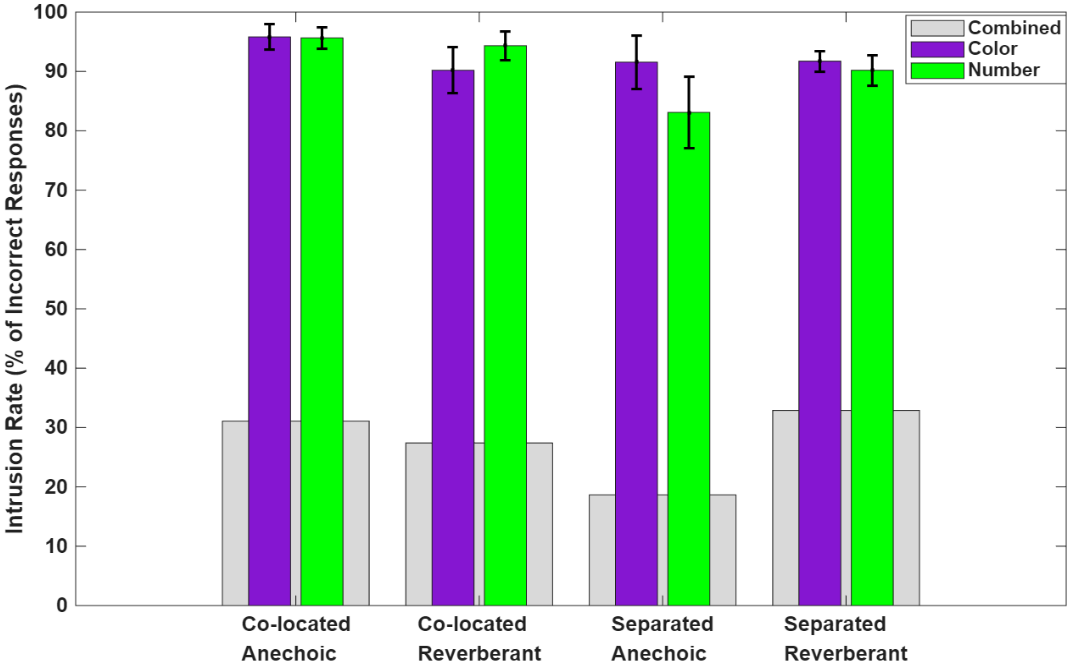
When Target is separated from Maskers, slow tuning to the target location occurs.



COLOR Vs. NUMBER INTRUSION

Intrusion : Response color and number matches a single masker. Separating color and number intrusions helps identify which part of the Stimuli is most affected by masking.





Significant interactions: **Environment × Location** (F = 28.02) and **Location × Item** (F = 12.72), both p < .001

As a proportion of all trials:
Independent of environment, number intrusion rate is higher than color intrusion rate in co-located condition, and lower in the separated condition.

Only location had a **significant** main effect **(F = 6.17, p = .01)**.

As a proportion of incorrect trials:
Intrusion rate
- **similar for color and number,**
- **lower in separated than co-located,**
- **in both environments.**

In combined analysis, intrusions lowest in Separated Anechoic Cond

INTRUSION BASED ON TMR (TARGET TO MASKER RATIO)

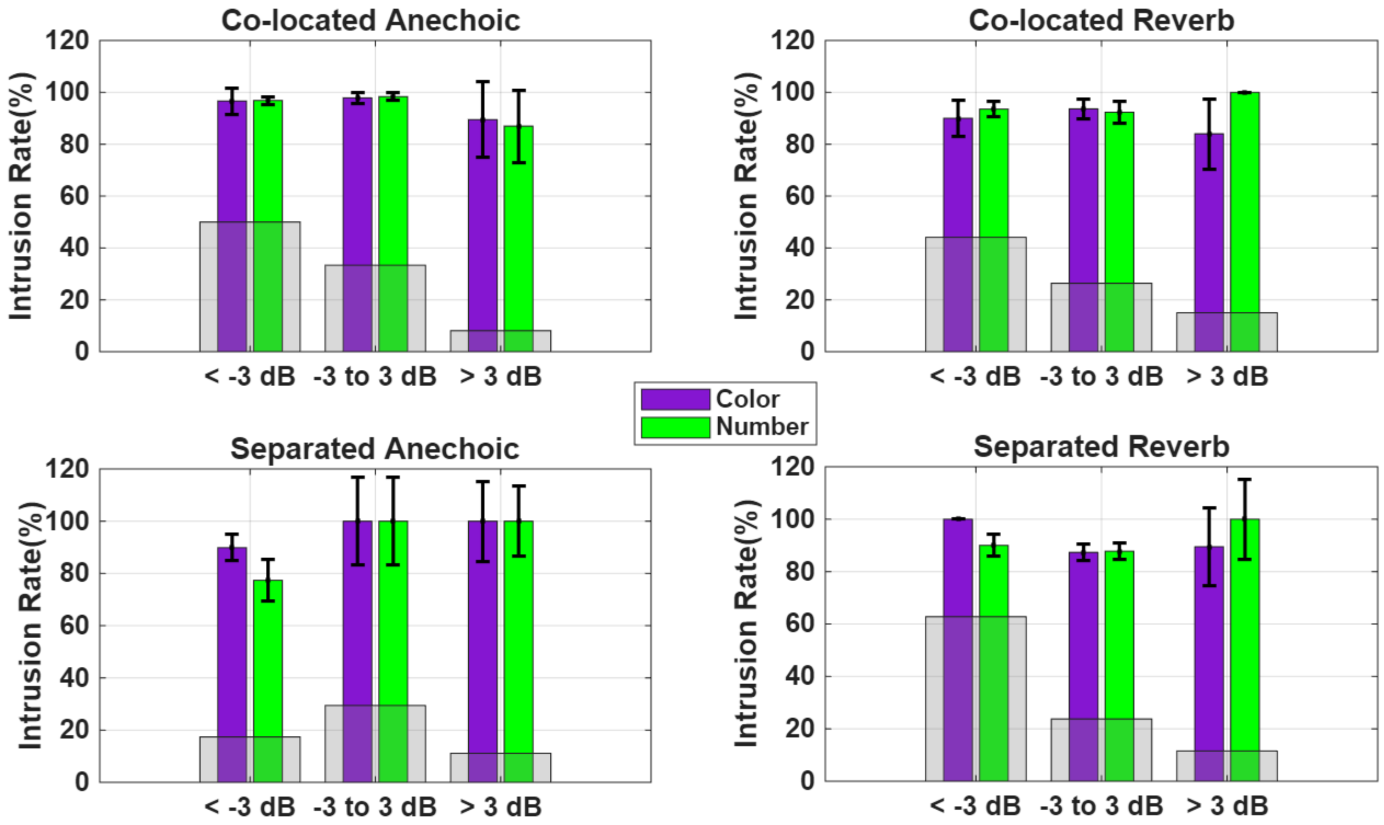
We grouped trials based on target to masker ratio (TMR) into three ranges:

<-3 dB, -3 to 3 dB and >3dB

This allows us to examine how different levels of masking affect intrusion rates across conditions

No strong trends for word-based intrusion analysis.

In, combined color+number analysis, intrusion rate decreases with TMR, except in separated anechoic condition.



CONCLUSIONS AND DISCUSSION

- SRM was **8.64 dB** in the anechoic environment, reduced to only **1.08 dB** in reverberation.
- This was driven mainly by an increase in the separated reverberant cost **8.02 dB** higher while the **co-located** staying close to **0.46 dB** in reverberation.
- The slopes of the psychometric curves were comparable in the co-located conditions (**anechoic: 13.4 %/dB, reverberant 13.3 %/dB**), becoming shallower in particular in the separated anechoic condition (**8 %/dB vs. 12 %/dB in reverberation**).
- Performance improved for the second vs first word in separated conditions, suggesting that tuning to space is slow even when only two spatial conditions are tested.
- Intrusion rates were lower in separated conditions (particularly separated anechoic) and decreased with TMR.
- The study showed that PART can be used to obtain measures of SRM consistent with published approaches.
- Results suggest that, in scenarios with high informational masking, the benefit of spatial separation can be dramatically reduced in reverberant vs anechoic environments, possibly due to increased difficulty segregating the target from maskers in reverberation.