

Generalized Turn-Taking–Based Conversational Partner Identification in Multi-Talker Scenarios using Voice Activity

→ Introduction

- Group conversations are more complex than dyads and therefore pose a greater challenge for people with hearing loss, especially in noise. Conversation partner identification could therefore be very beneficial for hearing aid users (HAUs).
- Prior research has developed the Bayesian Minimum Overlap and Gap (BMOG) method for identifying a single conversation partner [1]. However, this method does not extend beyond two-talker conversations.

→ Aim

To identify conversation partners by developing a **Multi-Talker BMOG (MT-BMOG)** framework that applies a **Conversational Alignment Score (CAS)** to measure turn-taking coordination in groups of arbitrary size.

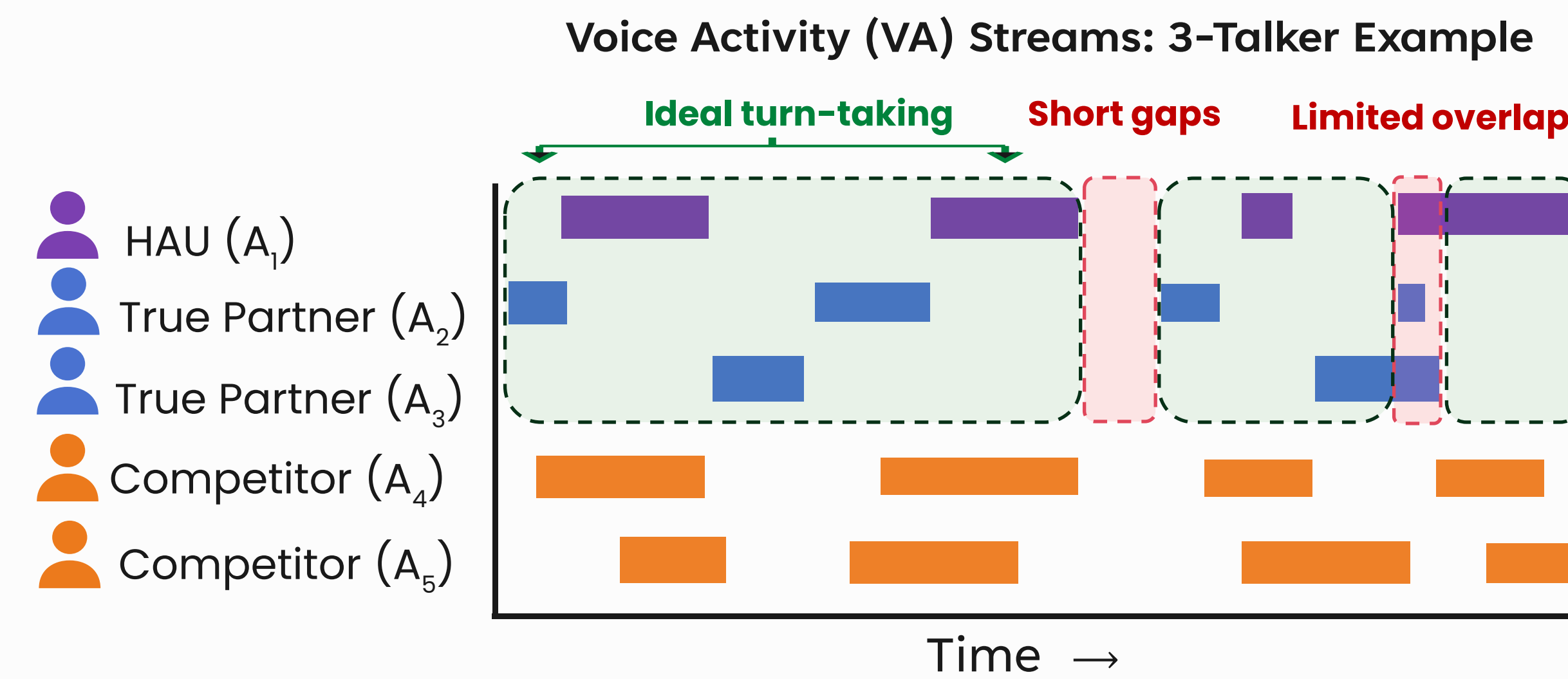
→ Evaluation Setup

- Real conversational scenarios of **2-talker**[2], **3-talker**[3], and **4-talker**[4] were evaluated.
- For each dataset and condition, 100 random realizations of conversation and competing groups were generated, with the HAU randomly selected from the conversation group.
- Compared with two extended BMOG baselines: **BMOG-IR**, which treats each HAU–talker pair as a separate dyadic conversation and selects the highest-ranked talkers as partners, and **BMOG-GE**, which evaluates candidate groups using merged voice-activity streams and selects the highest-ranked group as conversational partners.

→ Hypothesis & Conversation Science Motivation

HYPOTHESIS:

In multi-talker conversations, the HAU's true conversational partners should show stronger turn-taking coordination with minimal gaps and minimal overlap than competing talkers.

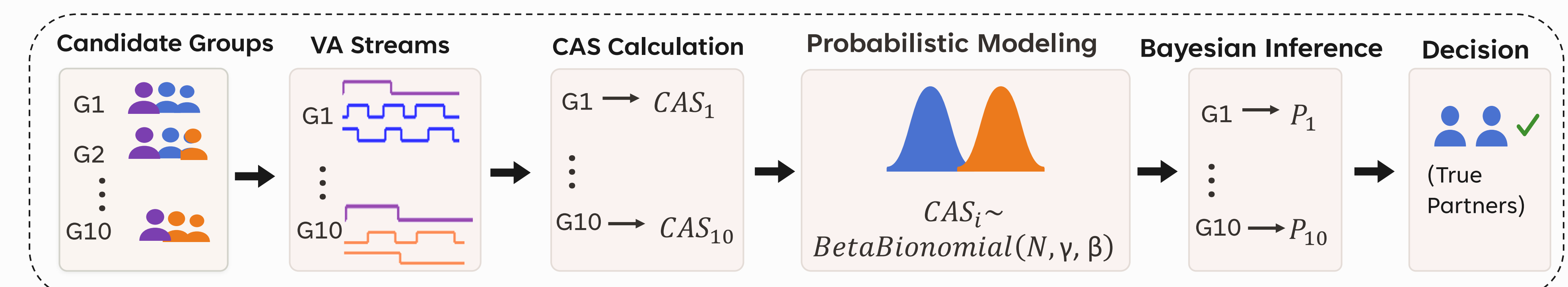


Key assumption

- True partners show **stronger coordination** (shorter gap and less overlap).
- Competitors exhibit weaker or conflicting turn patterns.

→ Method Overview

MT-BMOG algorithm pipeline



Conversational Alignment Score (CAS)

$$f_i^{(M)}(n) \triangleq \sum_{p,q \in Q_i, p < q} (A_p(n) - A_q(n))^2 - (M-2) \sum_{z \in Q_i} A_z(n) + \sum_{m=3}^M m(-1)^{m+1} \sum_{U \subseteq Q_i, |U|=m} \prod_{v \in U} A_v(n)$$

$$CAS_i^{(M)}(n; N) = \frac{1}{N} \sum_{r=n-N+1}^n f_i^{(M)}(r)$$

$A_k(n)$: binary voice activity Q_i : candidate group
 M : group size N : window length

Decision Rule

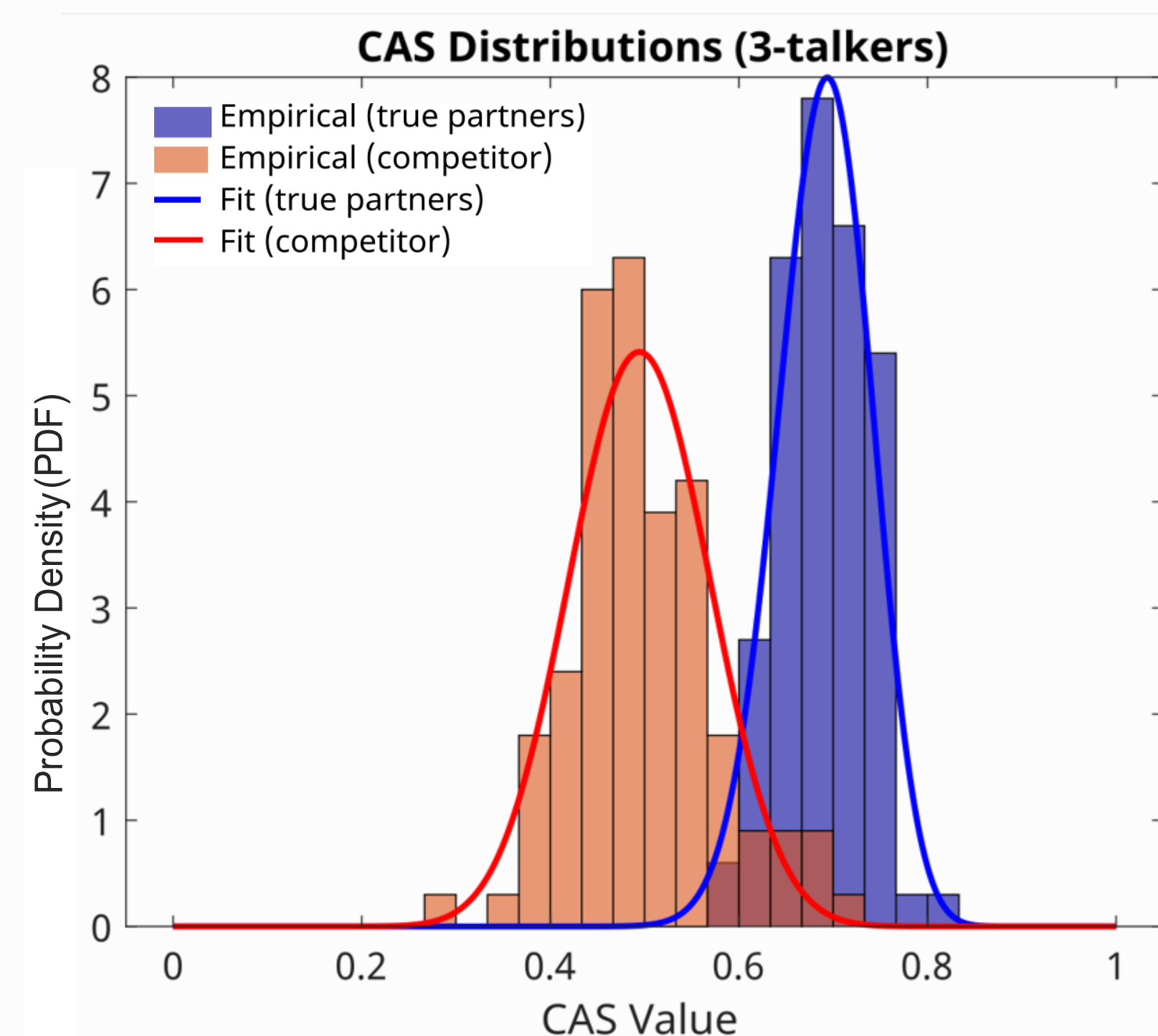
$$\widehat{TG} = \arg \max_{i \in \{1, \dots, I\}} P(\mathcal{H}_i | CAS_1, \dots, CAS_I)$$

Select the candidate group $\widehat{TG}$ with the highest posterior probability.

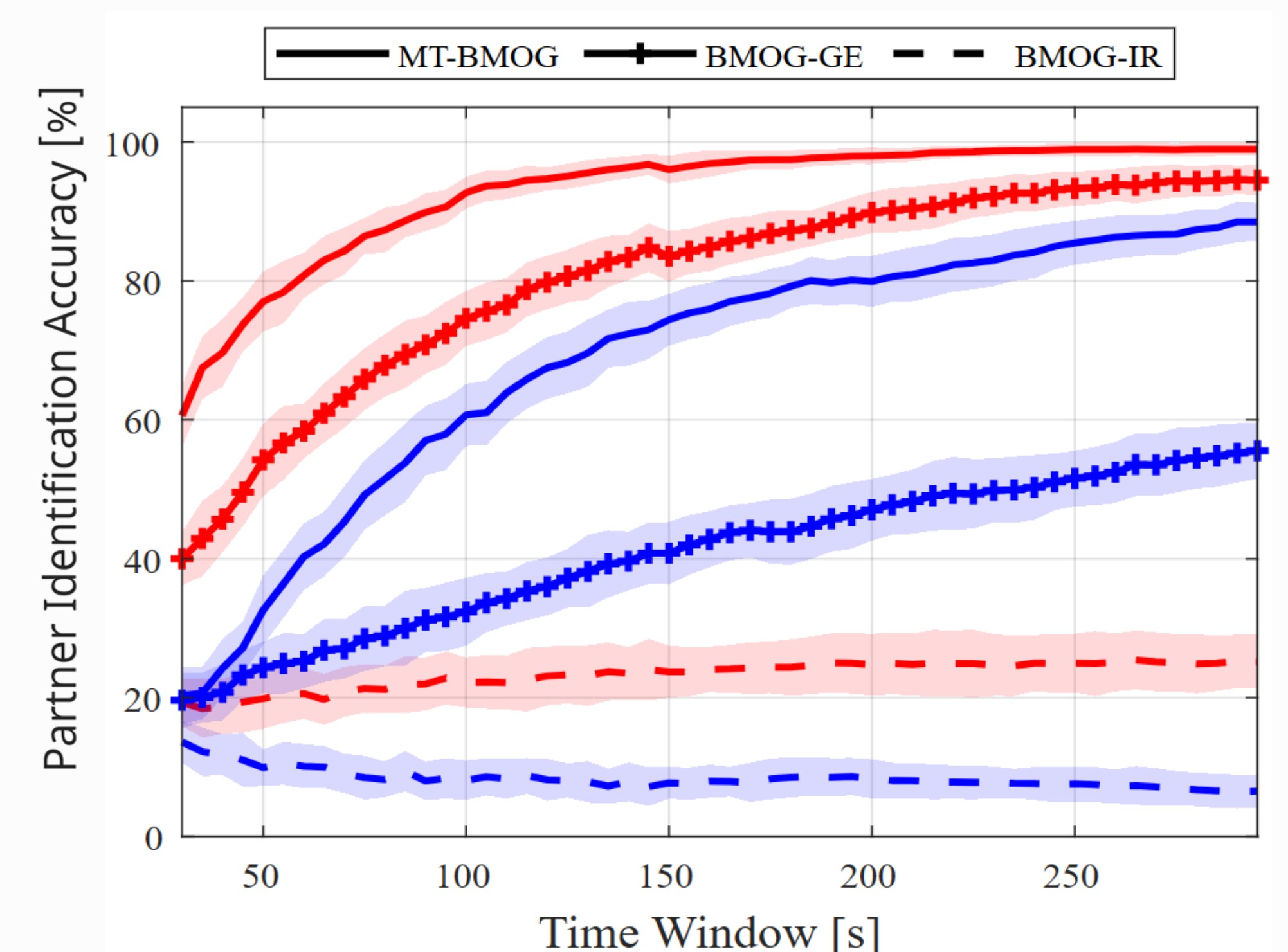
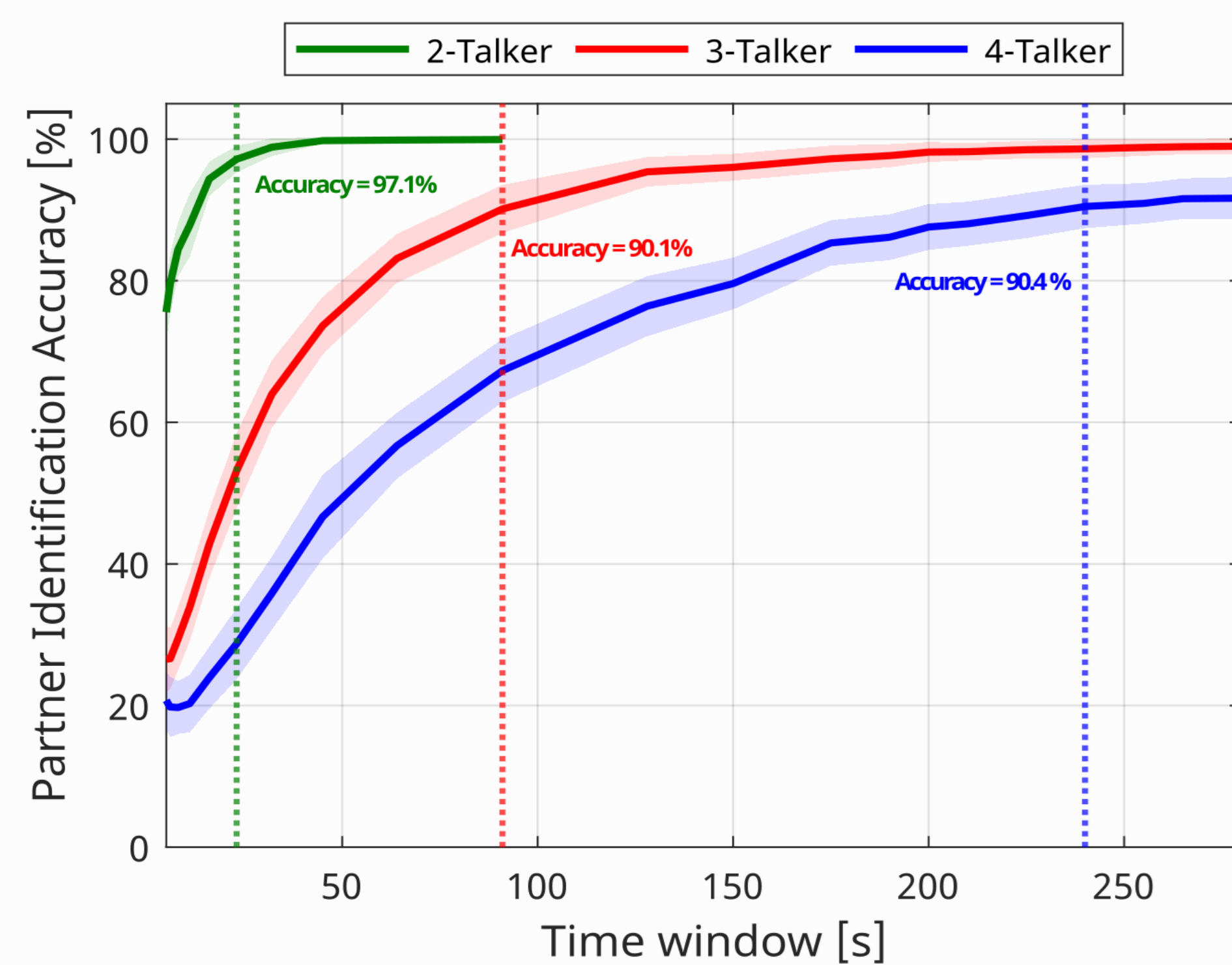
$\mathcal{H}_i$: hypothesis that candidate group i is the true group

→ Results

➤ True partners achieve higher CAS values.

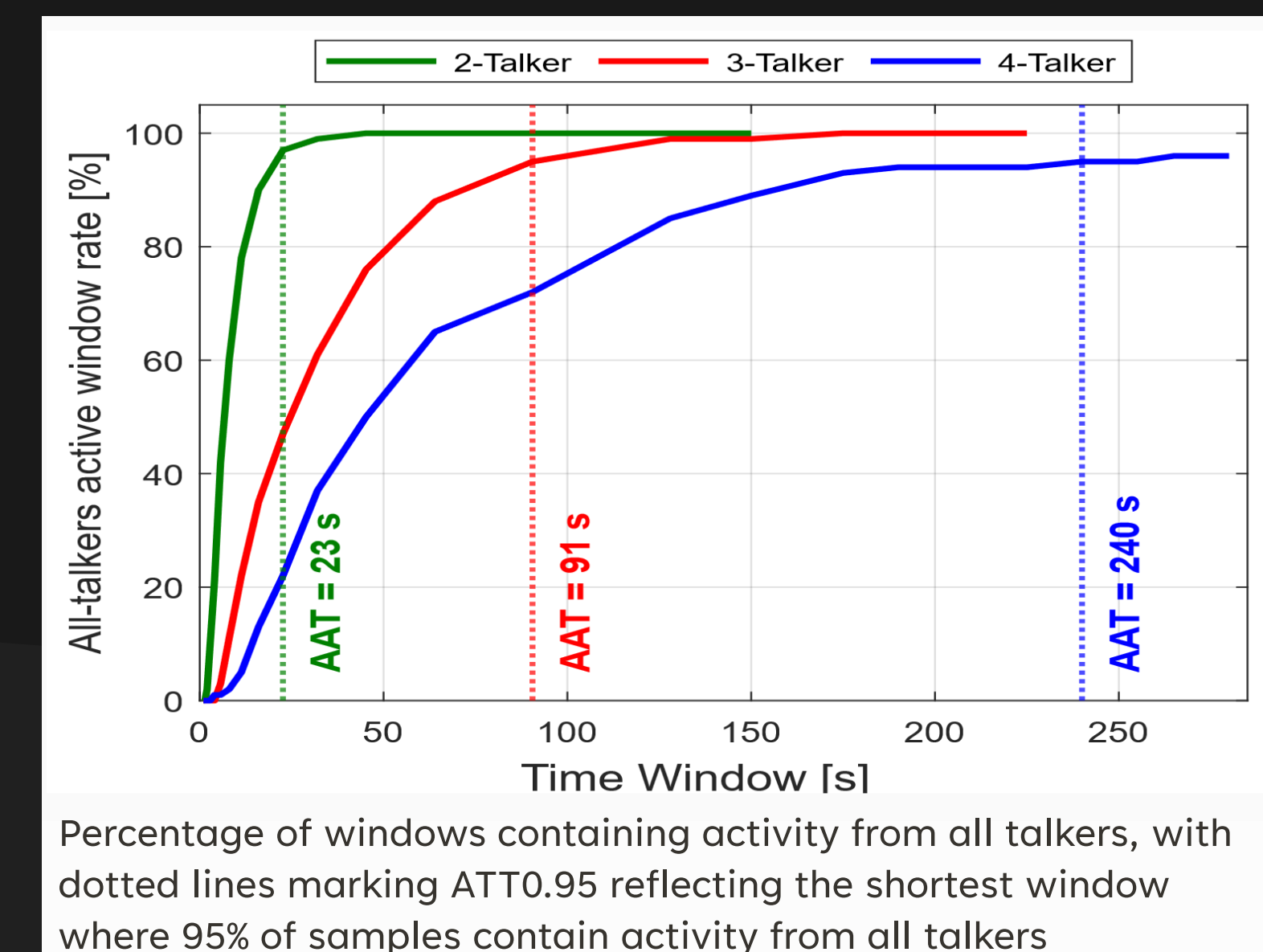


➤ Accuracy improves with integration time and follows talker-activity availability.



→ Conclusion

- MT-BMOG extends turn-taking-based partner identification to multi-talker conversations using CAS.
- As shown in the plot, high accuracy is reached once all partners are sufficiently active, indicating that lower accuracy in short windows is mainly due to insufficient group interaction.
- BMOG-IR ignores overlaps and gaps between non-HAU talkers, while BMOG-GE loses interaction information by merging talker activity. MT-BMOG preserves these group-level turn-taking dynamics, resulting in better partner identification.
- Next step: reduce observation time using complementary spatial/behavioral cues and improve inaccuracies in the VA detection.



- Hoang, P. et al. (2022). Minimum overlap-gap algorithm for speech enhancement. IEEE Access.
- Sørensen, A.J. (2021). DiapixDK picture materials, adapted Danish version. Zenodo.
- Petersen, E.B. et al. (2022). Effects of hearing-aid amplification and noise on conversational dynamics. Trends in Hearing.
- Nicoras, R. et al. (2025). Effective design for experiments on small-group conversation. American Journal of Audiology.