THE RELATION BETWEEN MULTIMODAL BEHAVIOUR AND ELABORATIONS IN EXPLANATIONS

**Stefan Lazarov & Angela Grimminger, Paderborn University**
*Presented at the 1st International Multimodal Communication Symposium,
26-28 April 2023, University of Pompeu Fabra, Barcelona, Catalonia
Special acknowledgements to Kai Biermeier*

UNIVERSITÄT PADERBORN

TRR 318 · **CONSTRUCTING EXPLAINABILITY**

1. Introduction

- Monitoring explainees' multimodal behaviour [1] is essential for supporting understanding via **elaborations** initiated by explainers [2].
- Yet there is no account about the **relation between elaborations in explanations initiated by explainers and explainees' multimodal behaviour**.

Fig. 1. Data segmentation example
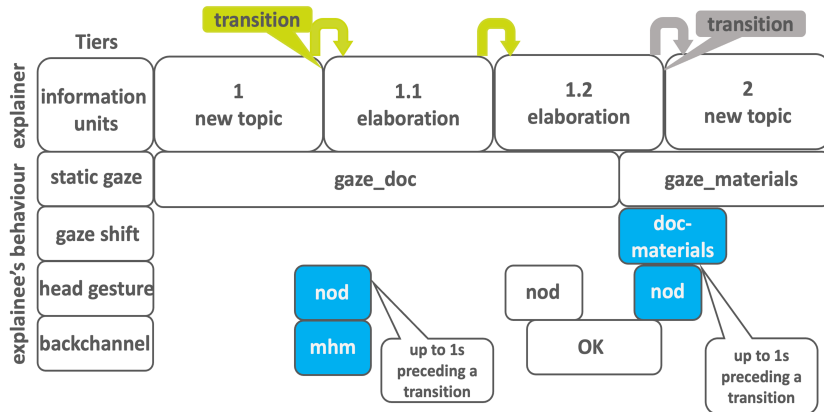
2. Previous research & motivation

Monitoring is defined as a continuous exchange between interlocutors about the state of an explanation in the form of an identification of spontaneous problems or a request for further information [1], e.g., via turn adaptation by explainers to explainees' behaviour [3].

Gaze behaviour may be related to indications of turn keeping/taking [4]. Also, **gaze aversions** may indicate increased cognitive processing load [5, 6].

Head nodding & backchannelling indicate either understanding or mental presence of the addressee [7], and they often occur bimodally [8, 9].
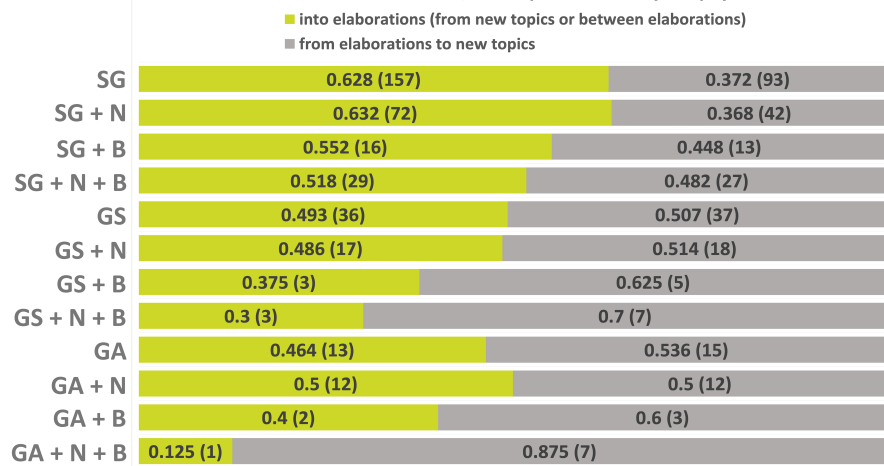
3. Research methods

- Explorative analysis of 10 physician - caregiver interactions during an explanation of an upcoming surgery [10]
- Coding transitions **into elaborations (from new topics or between elaborations)** and **from elaborations to new topics** initiated by explaining physicians (Fig. 1) [11]
- Coding explainees' **static gaze** (directed towards the physician / materials), **gaze shifts** (changes of the viewing direction), **gaze aversions** from the physicians, **head nods** and **backchannels** with an **onset** preceding the coded transitions by max. 1 s [12]
- **Conditional probabilities:** $P(A|B) = P(A \cap B)/P(B)$ [13]
  Example: $P(\text{elaboration}|\text{static gaze}) = P(\text{elaboration} \cap \text{static gaze})/P(\text{static gaze})$

4. Results

- **333 elaborations** & **242** related **topic changes** initiated by the explainers
- 575 instances of explainees' multimodal behaviour: unimodal (gaze only), bimodal (gaze & nodding / backchannelling, and multimodal (gaze, nodding & backchannelling), resulting in 12 forms of multimodal behaviour.
- **Higher probabilities** for explainees' unimodal, bimodal and multimodal behaviour with **static gaze** related to transitions **into elaborations**
- **Lower probabilities** for explainees' unimodal, bimodal and multimodal behaviour with **gaze shifts and aversions from the explainers** related to transitions **into elaborations**
- **Equal probabilities for both transition types** for explainees' **gaze aversions** from the explainers accompanied by **nodding**
- **Nearly equal probabilities** for explainees' **gaze shifts** occurring **unimodally** and **gaze shifts** accompanied by **nodding**

Fig. 2 Conditional probabilities of caregivers' multimodal behaviour preceding transitions to elaborations/new topics initiated by the physicians

- **into elaborations (from new topics or between elaborations)**
- **from elaborations to new topics**

	into elaborations	from elaborations to new topics
SG	0.628 (157)	0.372 (93)
SG + N	0.632 (72)	0.368 (42)
SG + B	0.552 (16)	0.448 (13)
SG + N + B	0.518 (29)	0.482 (27)
GS	0.493 (36)	0.507 (37)
GS + N	0.486 (17)	0.514 (18)
GS + B	0.375 (3)	0.625 (5)
GS + N + B	0.3 (3)	0.7 (7)
GA	0.464 (13)	0.536 (15)
GA + N	0.5 (12)	0.5 (12)
GA + B	0.4 (2)	0.6 (3)
GA + N + B	0.125 (1)	0.875 (7)

Labels: SG = static gaze; GS = gaze shift; GA = gaze aversion; N = nodding; B = backchannelling
* Total frequencies are given in the brackets.

5. Discussion

- Multimodal behaviour is related to different transition types:
  - static gaze (+ other modalities) → transitions into elaborations
  - gaze shifts and gaze aversions (+ other modalities) → transitions from elaborations to new topics; associated with a completion of cognitive tasks [5, 6]
- The co-occurrence of either nodding or backchannelling, or both with any type of gaze behaviour supports previous findings on the ambiguous function of these modalities [7-9].

→ Next steps:
- Analysing individual differences of caregivers' multimodal behaviour
- Including other transition types in the analysis, e.g., between multiple new topics in order to analyse the whole explanation structure

References

[1] Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*, 62–81.
[2] Dingemanse M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S. (2015). Universal Principles in the Repair of Communication Problems. *PLoS ONE, 10*(9), Article e0136100.
[3] Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50*, 696-735.
[4] Jongerius, C., Hillen, M. A., Romijn, J., A., Smets, E. M. A., & Koole, T. (2022). Physician gaze shifts in patient-physician interactions: functions, accounts and responses. *Patient Education and Counceling, 105*(7), 1-14.
[5] Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26*, 651–658.
[6] Morency, L. P., Christoudias, C. M., & Darell, T. (2006). Recognizing gaze aversion gestures in embodied conversational discourse. *Proceedings of the 8th International Conference on Multimodal Interfaces, NY, USA*, 298-294.
[7] Gander, A. G., & Gander, P. (2020). Micro-feedback as cues to understanding in communication. Dialogue and Perception – Extended Papers from DaP2018. In C. Howes, S. Dobnik, & E. Breitholtz (Eds.) *CLASP Papers in Computational Linguistics* (pp. 1-11). Gothenburg University.
[8] Allwood, J. & Cerrato, L. (2003). A study of gestural feedback expressions. *Proceedings of the 1st Nordic Symposium on Multimodal Communication*. Copenhagen, 7-22.
[9] Buschmeier, H., & Kopp, S. (2014). A dynamic minimal model of the listener for feedback based dialogue coordination. *DialWatt - SemDial 2014: Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue*. Edinburgh, UK, 17-25.
[10] Fisher, J. B., Lohmer, V., Kern, F., Barthlen, W., Gaus, S., & Rohlfing, K. J. (2022). Exploring Monological and Dialogical Phases in Naturally Occurring Explanations. *KI - Künstliche Intelligenz, 36*, 317-326.
[11] Roscoe, R. & Chi, M.T.H. (2008). Tutor learning: the role of explaining and responding to questions. *Instructional Science, 36*(4), 321-350.
[12] Louwerse, M. M., Dale, R., Bard, E. G. & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science, 36*, 1404-1426.
[13] Dekking, F. M., Kraaikamp, C., Lopuhää, H. P. & Meester, L. E. (2005). *A modern introduction to probability and statistics*. Springer.