# Self-adaptors while listening about novel and repeated topics in explanations

**Stefan Lazarov (stefan.lazarov@upb.de)[1,3], Olcay Türk[2,3], Petra Wagner[2,3] & Angela Grimminger[1,3]**
[1]Faculty of Arts and Humanities, Psycholinguistics Group, Paderborn University
[2]Faculty of Linguistics and Literary Studies, Phonetics Group, Bielefeld University
[3]TRR 318 "Constructing Explainability"

**Abstract**

In this study, we investigated the production of self-adaptors by explainees ($N$ = 53) while listening about novel and repeated topics conveyed by explainers ($N$ = 20) during board game explanations. Based on previous research linking self-adaptors to increased cognitive load and non-understanding in novel stimuli, we hypothesized that self-adaptors would occur more frequently during novel than repeated topics. To add granularity to topic repetition frequency, we considered topics that repeated once, twice, or three or more times. Contrary to our prediction, self-adaptors occurred significantly more often during topics repeated two or more times than during novel topics. This suggested that frequent topic repetitions could be associated with increased cognitive load, consistent with prior research linking self-adaptors to non-understanding. Further, we tested whether explainees' self-adaptors differed by topic initiator. Similar distributions of explainees' self-adaptors across topics initiated by themselves or the explainers indicated no effect compared to topic repetitions.

**Keywords:** self-adaptors; understanding processes; explanations

## Introduction

In research on multimodal interaction, adaptors (Cienki, 2024; Ekman & Friesen, 1969; Freedman, 1972) — self-oriented bodily movements that have no referential function compared to referential gestures—are yet among the multimodal behaviors that have been less investigated. Previous empirical research on adaptors has mainly focused on non-interactional settings to investigate the effect of self-adaptors (a subgroup of adaptors), for example, in lexical retrieval tasks (Cienki et al., 2023; Sekine & Hotta, 2025), memory tasks (Mueller et al., 2019; Spille, Grunwald, et al., 2022; Spille, Mueller, et al., 2022), and stress condition tests (Dreisoerner et al., 2021). These studies have demonstrated that self-adaptors indicate cognitive processing and attention regulation (Harrigan, 1985; Lopez-Ozieblo, 2025). In a recent interactional study, Lopez-Ozieblo (2025) suggested that future empirical studies should investigate the function of self-adaptors more differentially, specifically regarding the interlocutor's conversational role (speaker vs. listener).

In the present study, we investigated self-adaptors in interactional settings. More specifically, we analyzed dyadic explanatory interactions of a board game between less knowledgeable explainees and more knowledgeable explainers who explain the objects and processes relevant in the game (Kang et al., 2015; Kotthoff, 2009; Rohlfing et al., 2021). We focused on explainees' self-adaptors in the context of understanding the explanations. With regard to the organic flow of explanation topics and the way they emerge within dyadic explanations (Fisher et al., 2022), we explored the relation of explainees' self-adaptors to novel and repeated explanation topics, as well as to the interlocutor initiating them. Our research was motivated by results from recent studies on explanatory interactions showing that explainees' adaptors were related to moments of non-understanding (Türk et al., 2024), and that the topical structure of explanations was adapted to explainees' multimodal behavior (Lazarov, 2025; Lazarov et al., 2024). In this exploratory study, we investigated two questions: (1) How frequently do explainees produce self-adaptors while listening to explanations of novel and repeated topics? (2) How are these frequencies distributed with regard to the interlocutor who initiates the topic during the interaction?

## Characteristics of self-adaptors

Self-adaptors form a group of self-oriented bodily movements (often articulated by the hands) that have no overt speech-related function in the same way co-speech gestures (e.g., deictics, iconics, beats) do (Cienki, 2024; Ekman & Friesen, 1969; Freedman, 1972; Kendon, 2004; Lopez-Ozieblo, 2025). There are also other groups of adaptors, namely object-adaptors and alter-adaptors which remained out of the scope of the present study. In terms of form, self-adaptors represent movements oriented toward the body, such as self-touch, self-scratch, fidgeting, etc. (Ekman & Friesen, 1969; Harrigan, 1985; Kimura, 1976; Lopez-Ozieblo, 2025; Türk et al., 2024). In addition, clothes adjustment could be categorized as either self-adaptor (Harrigan, 1985; Türk et al., 2024) or object-adaptor (Ekman & Friesen, 1969; Maricchiolo et al., 2012), depending on the annotation procedure.

Unlike co-speech gestures (Kendon, 2004; McNeill, 2005), adaptors are not related to semantic or syntactic units of speech, but rather to a person's emotional or cognitive state (Ekman & Friesen, 1969; Kendon, 2004). In states of increased levels of anxiety or cognitive load, adaptors may serve as 'regulators' (Ekman & Friesen, 1969; Kendon, 2004). In non-interactional studies, fidgeting movements have been shown to indicate individuals' temporal disengagement from a task and correlate with the increasing level of task boredom (Ricciardi et al., 2019; Seli et al., 2014). Facial self-touches have also been reported to reduce cognitive load (Mueller et



In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* 

al., 2019) and stress (Dreisoerner et al., 2021; Iriskhanova et al., 2023), suppress negative emotions (Spille, Mueller, et al., 2022), and regulate attention and working memory (Grunwald et al., 2014; Spille, Grunwald, et al., 2022; Spille, Mueller, et al., 2022; Spille et al., 2021). For language production (also in non-interactional settings), individuals' self-adaptors have been shown to aid lexical retrieval during oral translation tasks (Cienki et al., 2023) and story retelling, yet not resolving tip-of-the-tongue states (Sekine & Hotta, 2025).

Research on the function of self-adaptors in interactional settings remains sparse, with few exceptions. In a study on physician–patient consultations, Harrigan (1985) investigated the production of self-adaptors (touching faces, hands, and clothing) during both language production and understanding explanations. The study revealed that patients produced significantly more self-adaptors while answering physician's questions and making health-related statements than while listening to physicians' questions, answers, and statements. Harrigan (1985) attributed this difference to the cognitive demands of producing and processing sensitive or unpleasant information. Conversely, physicians produced more self-adaptors while listening to patients' statements than while speaking (Harrigan, 1985). According to Harrigan (1985), physicians' self-adaptors during patients' medical and psychosocial accounts may have reflected challenges in understanding the patients' condition as presented by non-experts.

A more recent and longitudinal study on second language learning by Lopez-Ozieblo (2025) revealed that speakers' self-adaptors decreased over time as familiarity with monological tasks increased, suggesting reduced cognitive demands with practice. In the same study, Lopez-Ozieblo (2025) analyzed speakers' and listeners' self-adaptors in face-to-face dialogues referring to them as cognitive self-regulators, also emerging through synchrony. However, the results by Lopez-Ozieblo (2025) were not conclusive as the author recommended that future research should provide deeper understanding about the functions of adaptors of both speakers and listeners.

Another recent study on multimodal signals of (non-)understanding in dyadic human–human explanations revealed that self-adaptors are significantly associated with moments of explainees' non-understanding (Türk et al., 2024). Yet, the question how explainees' self-adaptors while listening to explainers' explanations are related to the dynamically evolving topical structure remains open.

## The dynamic topical structure of explanations

The topical structure of explanations is oriented toward resolving questions, such as *what*, *how* and *why* (Klein, 2009). These questions point to further sub-explananda, namely explanation topics involving novel and repeated (elaborated) information (Lazarov et al., 2024). One way to analyze topics within explanations is to apply an episodic segmentation (Roscoe & Chi, 2008), marked by semantic boundaries in interlocutors' spoken utterances or parts of utterances (Haas et al., 1999). Although this topical structure is largely organized by explainers as the more knowledgeable interlocutors, it is prone to dynamic changes based on explainees' continuous verbal and nonverbal feedback behavior (Fisher et al., 2022; Lazarov, 2025; Lazarov et al., 2024). Regarding nonverbal behavior, Lazarov et al. (2024) found that explainers were more likely to elaborate on previous topics when explainees showed only fixed gaze, without accompanying cues, such as head gestures or vocal backchannels. However, no study to date has examined self-adaptors as feedback in relation to changes in the topical structure of explanations, including shifts to novel or repeated topics.

## The present study

In the present exploratory study, we investigated self-adaptors produced by explainees in relation to the dynamic topical structure while the explainers were speaking. More specifically, we related explainees' self-adaptors to novel and repeated topics, and analyzed this distribution with regard to the interlocutor who initiated the topic. Based on previous research pointing out that self-adaptors regulate cognitive load (Ekman & Friesen, 1969; Kendon, 2004; Mueller et al., 2019), working memory and attention (Grunwald et al., 2014; Ricciardi et al., 2019; Seli et al., 2014; Spille, Grunwald, et al., 2022; Spille, Mueller, et al., 2022; Spille et al., 2021), and that they decrease with task familiarity (Lopez-Ozieblo, 2025), we hypothesized that (1) explainees produce more self-adaptors while listening to novel than repeated explanation topics. Novel explanation topics were considered to demand more cognitive resources (higher cognitive load), thus requiring explainees to regulate working memory capacity.

Based on the findings from the interactive study by Türk et al. (2024) showing that explainees' self-adaptors are associated with moments of non-understanding, we hypothesized that (2) explainees produce more self-adaptors when explanation topics are verbally initiated by themselves than by the explainers. We anticipated that verbal topic initiations by the explainees could be a signal of non-understanding.

### Methods

**Participants** In the current study, the data from 20 explainers and 53 explainees was analyzed (see below). All participants were German native speakers. Among the explainers ($M_{\text{Age}} = 23.35, SD_{\text{Age}} = 2.64$) 6 were male, 13 female, and 1 non-binary. Of the 53 explainees, only 36 provided sociodemographic data ($M_{\text{Age}} = 24.22, SD_{\text{Age}} = 7.43$, 12 male, 24 female). We consider the missing sociodemographic data of the other explainees as unproblematic because the research question did not address age or gender. All participants signed a consent form, and the study had been approved by the Ethics board of Paderborn University.

**Materials** We analyzed a subsample of 53 dyadic explanatory interactions from the MUNDEX corpus (Türk et al., 2023), which was compiled to investigate explainees' multimodal signals of (non-)understanding. The entire corpus contains 87 naturalistic explanatory interactions about a competitive board game named 'Deep Sea Adventure' (Sasaki & Sasaki, 2014) in German. The subsample for the current study was selected based on the current state of the annotations.

In the corpus design, one explainer explained the board game subsequently to either two or three explainees. Furthermore, the interactions followed a specific but also temporally varying three-phase structure: 1) explanation in the absence of the board game (game absent phase), 2) a continuation of the explanation with the board game (game present phase), and 3) an interactive game play. This structure was designed to examine how situational dynamics are related to explainees' understanding and multimodal behavior; however, it is not considered here, as the analyses are limited to the game-absent phase of the data. The duration of the selected interactions averaged 25:35 min (*SD* = 05:09), while the analyzed game-absent phase averaged 06:57 min (*SD* = 03:10).

The interactions were videotaped from six perspectives: two cameras directed at participants' face area, two cameras at participants' torso area, one side-positioned camera, and one camera positioned on the top. Additionally, participants' speech was audio-recorded for speech analysis using headsets.

**Procedure** Prior to the study, all participants were assigned either the role of explainer or explainee. The explainers were given the opportunity to learn and practice the board game on their own a few days prior to the study. None of the participants were informed about the main objective of the study prior to it in order to ensure that the interactions would remain as natural as possible, mirroring everyday communicative exchanges. The only instructions provided to the explainers were 1) the sequence in which the board game should be introduced to the explainees (i.e., the game phases), and 2) that their explanations should be sufficiently comprehensive to enable the explainee to play the game with confidence afterward.

**Data coding** All video data was annotated in ELAN (Wittenburg et al., 2006). Annotators manually segmented the explainers' and the explainees' utterances into explanation topics, based on previous semi-automatic transcription and segmentation of their speech into utterances split by pauses of ≥ 250 ms (Kisler et al., 2017). For the present study, only the explainers' utterances were analyzed. Further, explainees' self-adaptors were manually annotated.

*Explanation topics*. The segmentation into distinct explanation topics was based on previous approaches outlined by Klein (2009) and Roscoe and Chi (2008). The topic segments spanned full utterances or parts of utterances, defined by the semantic boundaries of the topics (Haas et al., 1999). The list of topics anticipated to emerge within the explanations was derived from the official game instructions. These instructions were given to the explainers, together with the game itself, as preparatory materials. For labeling the explanation topics, the coders applied a hierarchical numbering system provided in the annotation scheme (Lazarov et al., 2025). To ensure reliability, two annotators independently coded the same 10% of the data prior to the main annotation process (Fleiss' $\kappa$ = 0.79). Subsequently, both coders completed the remaining annotations, each working on different videos.

Subsequently, the speakers' participation within the specific topics was annotated—either as initiated by the explainer (EX) or by the explainee (EE). While all topics stemmed from the EX's utterances, the initiative to introduce a new topic or to repeat a previous one could come from either partner. The EXs introduced novel topics or topic repetitions spontaneously over the course of the interactions in line with their individual explaining strategies. EEs initiated topic repetitions through open or closed questions, paraphrases, or repetitions of previously given information to verify their understanding. Cases in which the EXs repeated or elaborated on a topic in response to such verbal requests by the EEs were coded as explainee-initiated (see Example (2) in Figure 1). In the analyzed subsample, 1238 topics were initiated by the explainers and 97 by the explainees.

In addition, the occurrence of topics as either novel or repeated, was annotated (Figure 1). For example, an explanation topic that occurred for the first time was annotated as novel. Explanation topics that occurred more than once (independent of who initiated it) were annotated as repeated. Repeated topics emerged from the natural flow of the explanatory interactions regarding the verbal participation of the explainers and the explainees (Fisher et al., 2022). All topic repetitions included elaborations, for example in the form of additional factual, procedural, and causal information related to the game rules, or corrections of incorrectly explained topics.

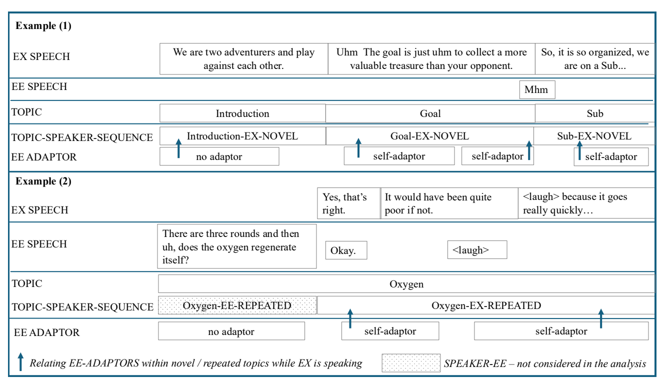


Figure 1: Linking explainees' self-adaptors to topics. Each line in both examples represents tier annotations in ELAN.

*Adaptors*. Two coders segmented and annotated the explainees' self-adaptors and object-adaptors. In this study, only self-adaptors were analyzed, which were identified as explainees' self-touching movements (on any body area), hand fidgeting, adjustments of sitting position, and adjustments of clothes as in Harrigan (1985) and Türk et al. (2024) (Figure 2). To ensure reliability, the two main coders and an additional third coder independently annotated 10% of the same material before the main annotation process (Fleiss' $\kappa$ = 0.77). Subsequently, the two main coders completed the remaining annotations, each working on different videos. In the selected subsample for the present study, in which the board game was absent, 1077 self-adaptors ($M = 20.32, SD = 16.85$) were annotated.

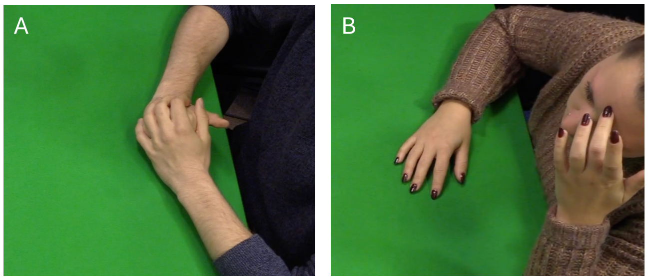


Figure 2: Self-adaptors: A) hand fidget, B) facial touch.

Coders created overlap tiers named “TOPIC_SPEAKER_SEQ_EE_ADAPTOR" from the tier containing the numeric annotations of the explanation topics, the speaking part (the explainer or the explainee) and the occurrence of topics (novel or repeated) “TOPIC_SPEAKER_SEQUENCE", and the tier containing annotations of the explainees’ adaptors “EE_ADAPTORS" (Figure 3). Explainees’ self-adaptors produced within the spoken part of the explainees were excluded from the analysis because the present study focused on self-adaptors produced while the explainees were listening.

Figure 3: Analyzing video data in ELAN.

## Data analysis

The dataset was imported to *RStudio* (RStudio Team, 2020) for statistical analysis. We fitted two generalized linear mixed effects models, one for the fixed effect addressed in hypothesis 1 (topic sequence) and one for that in hypothesis 2 (topic initiator) using the *lme4* package in R (Bates et al., 2015). In Model 1, the fixed effect ‘topic sequence’ contained four levels, representing the exact sequential order as they were mentioned during the individual explanations: ’Novel’ when topics occurred for the first time, and three levels of repeated topics, as those were sub-categorized into repeated once (’1st repetition’), twice (’2nd repetition’), or three or more times (’3rd+ repetition’). This subcategorization was done to provide more granularity into our exploration of the relation between EEs’ self-adaptors and their listening to topics repeated by the EXs.

**Model 1:**

```
glmer(SELF_ADAPTOR ~  TOPIC_SEQUENCE +
(1|EX:EE), family = binomial
```

In Model 2, the fixed effect TOPIC_INITIATOR indicated which interlocutor initiated a topic verbally, i.e., whether the explainer introduced or repeated a topic on their own or as a result of explainees’ request via questioning or repetitions. In the second case referring to explainees’ requests, the explainees were categorized as topic initiator. For example, if the explainees asked a question about the game chips (topic 2.4) and the explainer responded in relation to the same topic (2.4), the spoken part of the explainer covering this topic was defined as initiated by the explainee. Respectively, if the same topic was introduced by the explainer without the verbal contribution of the explainee, the topic was categorized as initiated by the explainer.

**Model 2:**

```
glmer(SELF_ADAPTOR ~  TOPIC_INITIATOR +
(1|EX:EE), family = binomial
```

In both models, SELF_ADAPTOR was the response binomial variable. Overall, 1934 cases in which the explainees did (computed as 1) and did not (computed as 0) produce self-adaptors while listening to novel and repeated topics conveyed by the explainers were observed. The random effect represented the individual interactions of each explainer with each explainee. For the descriptive part of the analysis, the binomial occurrences of explainees’ self-adaptors were converted to proportions for each level of TOPIC_SEQUENCE. Note that there was high imbalance in the distribution of topic initiation by either the explainers (1238) or the explainees (97) across the dataset. This imbalance was expected given the design of the present study where the explainers were more knowledgeable about the explanandum than the explainees. The model converged successfully without undersampling, thereby preserving ecological validity.

# Results

## Self-adaptors during novel vs. repeated topics

The first model indicated a slightly better fit ($AIC = 1929.8$) compared to a null model ($AIC = 1978.9$). This was also indicated by the low degree of variance of the fixed effect ‘topic sequence’ (Marginal $R^2 = 0.03$) and a substantial degree of variance including the random effect (Conditional $R^2 = 0.52$). The random effect indicated largely pronounced individual differences in explainees’ production of self-adaptors while listening across the 53 interactions ($\sigma^2 = 3.39, SD = 1.81$).

Regarding the fixed effect ‘topic sequence’, the model indicated topics being repeated for the second ($M = 0.7, SD = 0.46, est = 0.52, se = 0.2, z = 2.64, p < 0.01$) and the third+ time ($M = 0.84, SD = 0.37, est = 1.27, se = 0.21, z = 5.97, p < 0.001$) as significant predictors of explainees’ self-adaptors while listening compared to the intercept level ‘novel’ ($M = 0.62, SD = 0.49, est = 0.36, se = 0.26, z = 1.36, p > 0.05$). Topics being repeated for the second or the third+ time indicated a gradual increase in the production of self-adaptors by the explainees (Figure 4) compared to novel topics and topics repeated for the first time.

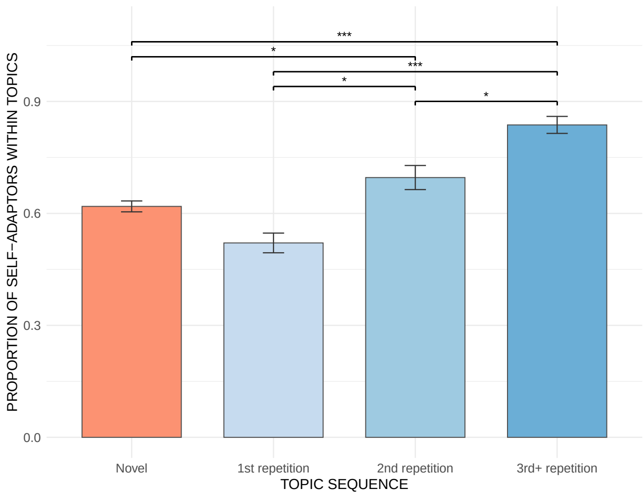


Figure 4: Proportions (*M* & *SE*) of explainees' self-adaptors while listening to novel and repeated topics.

To further verify the first hypothesis, we conducted Tukey-adjusted pairwise comparisons between the model predicted probabilities. The results indicated a significantly higher likelihood of the EEs' self-adaptors while listening to topics that were repeated by the EX for a second ($OR = 0.59, 95\% CI[0.36, 0.98], se = 0.2, z = -2.64, p < 0.05$) and third+ time ($OR = 0.28, 95\% CI[0.16, 0.48], se = 0.21, z = -5.97, p < 0.001$) compared to topics that were introduced by the EX as novel. No significant difference was found in self-adaptors during topics repeated only once compared to novel topics; in contrast, there was a gradual increase in the use of self-adaptors starting from the first repetitions of topics toward the third+ repetitions. More specifically, the results displayed a significantly higher likelihood of EEs producing self-adaptors stepwise for each increasing level of topic repetition: first vs. second repetitions ($OR = 0.45, 95\% CI[0.26, 0.8], se = 0.22, z = -3.6, p < 0.05$), second vs. third+ repetitions ($OR = 0.47, 95\% CI[0.24, 0.92], se = 0.26, z = -2.89, p < 0.05$), and first vs. third+ repetitions ($OR = 0.21, 95\% CI[0.12, 0.39], se = 0.23, z = -6.61, p < 0.001$). These results do not support the first hypothesis that explainees produce more self-adaptors while listening to novel compared to repeated topics.

### Self-adaptors related to the topic initiator

The second model did not indicate a better fit ($AIC = 1980.9$) compared to a null model ($AIC = 1978.9$), which was evident from the zero degree of variance of the fixed effect 'topic initiator' (Marginal $R^2 = 0.00$) and a moderate degree of variance including the random effect (Conditional $R^2 = 0.50$). As in the first model, the random effect indicated a high degree of individual differences in explainees' production of self-adaptors while listening across the 53 interactions ($\sigma^2 = 3.34, SD = 1.83$).

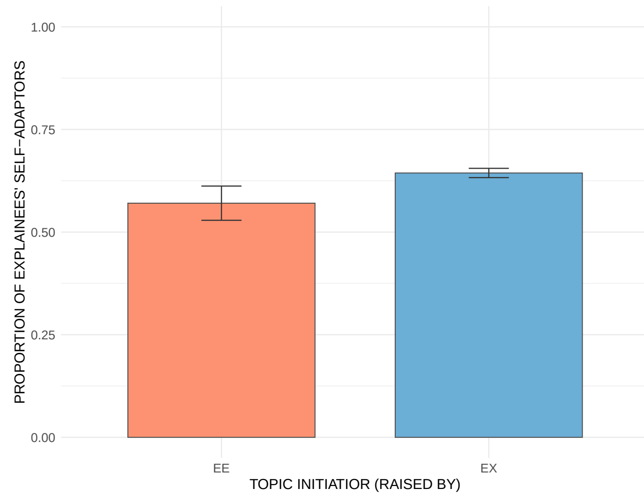


Figure 5: Explainees' self-adaptors (*M* & *SE*) while listening, by topic initiator.

The descriptive analysis of the proportional distribution of explainees' self-adaptors indicated that the explainees produced more self-adaptors while listening to topics being initiated by the explainers ($M = 0.64, SD = 0.48$) than by themselves ($M = 0.57, SD = 0.5$) (Figure 5). However, the model revealed no effect of the interlocutor initiating a topic — the EX or the EE — on the explainees' self-adaptors while listening. Given the results, the second hypothesis was not supported.

## Discussion

This study examined the production of self-adaptors by explainees while listening to novel and repeated explanation topics about a board game conveyed by explainers. In addition, it explored the distribution of explainees' self-adaptors while listening to topics initiated either by themselves or by the explainers. During the explanations, the board game was physically absent from the shared space. Given the scarcity of empirical research on self-adaptors in understanding processes during explanatory interactions, the current study investigated two exploratory hypotheses.

The first hypothesis that explainees produce more self-adaptors during novel compared to repeated explanation topics was not supported. Instead, the results indicated that explainees produce significantly more self-adaptors while listening to topics repeated for the second and the third+ time compared to novel topics. This hypothesis was based on previous findings that self-adaptors aid working memory, regulate attention (Grunwald et al., 2014; Ricciardi et al., 2019; Seli et al., 2014; Spille, Grunwald, et al., 2022; Spille, Mueller, et al., 2022; Spille et al., 2021) and reduce cognitive load (Ekman & Friesen, 1969; Kendon, 2004; Mueller et al., 2019). Contrary to our initial expectations, our results therefore suggest that the explainees experienced higher cognitive load and demand to free working memory capacity when the explainers repeated topics for the second, or third+ time. One possible explanation for this could be that some explanation topics were repeated more often than others, potentially in relation to moments of the explainees' non-understanding. This interpretation is supported by a recent study on multimodal predictors of understanding in explanatory interactions by Türk et al. (2024), who showed that explainees' moments of non-understanding can be predicted by the presence of adaptors. Taken together, our results indicate that repetitions of topics for two or more times mark a critical point of an interaction, displayed by increased self-adaptors produced by explainees. With this, we do not argue that novel topics impose no cognitive load on the explainees. Rather, our results point to substantial differences in the degree of cognitive load when novel topics are compared with topics repeated multiple times. One possible explanation for this is that repeated topics contained elaborations of factual, procedural, or causal information related to previously introduced topics (Lazarov et al., 2024). These repetitions were initiated spontaneously either by the explainers, who continuously monitor the explainees' multimodal behavior, or in response to verbal requests (Fisher et al., 2022; Lazarov et al., 2024). Consequently, the topic repetitions in the present study are likely to fill explainees' knowledge gaps about the board game or to correct misunderstood information (rather than being a sign of boredom due to familiarity, see below). Another explanation for the increased frequency of self-adaptors during second and third+ repetitions of topics could be that, as a result of non-understanding and uncertainty, the explainees' level of stress increased (Dreisoerner et al., 2021; Iriskhanova et al., 2023). To address this issue, future studies should consider linking frequently repeated topics to explainees' levels of understanding.

Regarding the effect of familiarity, our results did not align with those reported by Lopez-Ozieblo (2025). In their study, familiarity was identified as a factor influencing the production of adaptors in longitudinal settings with the focus on tasks. The present study did not include a longitudinal component, therefore familiarity, which could be considered as a factor contributing to or resulting from boredom (Ricciardi et al., 2019), does not apply to our interactional design.

Further, the finding that explainees' self-adaptors occur frequently in repeated explanation topics is in line with a recent study by Lazarov et al. (2024), which reported that explainees' fixed gaze direction was related to further elaborations of previous topics. A future study will examine explainees' self-adaptors co-occurring with the explainees' gaze direction (at the explainers, the shared space, or away) in relation to other behaviors signaling (non-)understanding in discourse, such as vocal feedback and head gestures. Addressing this limitation will extend our knowledge of the function of self-adaptors on cognitive load.

Our second hypothesis that explainees produce more self-adaptors while listening to explanation topics raised by themselves than by the explainers was also not supported. It was anticipated that when explainees raise a topic, for example by questioning or repeating previous information, they signal non-understanding and thus increased cognitive load (Türk et al., 2024). However, the results indicated that, regardless of which interlocutor raised a topic, the explainees' production of self-adaptors remained stable. Considering the functional characteristics of self-adaptors, these results suggest that the explainees experienced a similar degree of cognitive load in both ways of topic initiation within the dialogue. Given the knowledge gap between the explainers and the explainees inherent to the task (Kang et al., 2015; Kotthoff, 2009; Rohlfing et al., 2021), explanation topics are more likely to be initiated by the explainers (Lazarov, 2025). Before running our statistical analysis, we noted an imbalanced distribution of topic initiations according to which the explainers initiated considerably more topics than the explainees (see above). Even when the explainees initiated topics, their self-adaptors occurring during the speaking part of the explainers were analyzed. Given the existing knowledge gap between the interlocutors and the comparable frequency of self-adaptors regardless of who initiated a topic, we infer that the explainees' experienced a similar degree of cognitive load.

Our finding related to the second hypothesis was not in line with previous research by Harrigan (1985), who found that explainees (patients) produce more self-adaptors when explainers (physicians) initiate topics themselves rather than when the explainees (patients) initiate topics by asking questions. However, we acknowledge that our model might underestimate the effect of explainee initiations due to imbalance in distribution. Moreover, Harrigan (1985) attributed the patients' self-adaptors to possible distress experienced during the consultation with the physicians. In contrast to medical consultations, which address personally relevant and potentially sensitive issues, board game interactions concern abstract rule-based concepts that are less emotionally charged. Thus, our results do not contradict Harrigan's (1985) findings but rather extend the body of research on self-adaptors produced while listening by considering a different interactional context.

To our knowledge, this was the first study to address the production of listener's self-adaptors in understanding processes in explanatory interactions and relate them to the sequential occurrence of explanation topics. Our results showed that self-adaptors should not be associated only with novel but also with repeated topics, as the latter indicated an increased production of self-adaptors which may be related to moments of non-understanding or understanding-related issues. Future studies will investigate explainees' multimodal behaviors within those repeated topics by addressing the degree of topic complexity, which may explain the increased production of self-adaptors.

## Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): TRR 318/3 2026 – 438445824. We thank all participants for supporting this research and our research assistants for their help in transcribing and annotating the video data.

## Data availability

The annotation manual, materials, and statistical codes used for conducting the analysis are available at OSF. Following the data protection agreement between the researchers and the participants, audio and video data are not publicly available.

## References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Cienki, A. (2024). Self-focused versus dialogic features of gesturing during simultaneous interpreting. *Russian Journal of Linguistics*, *28*(2), 227–242. https://doi.org/10.22363/2687-0088-34572

Cienki, A., Leonteva, A. V., Agafonova, O. V., & Petrov, A. A. (2023). Numbers in simultaneous interpreting: A multimodal analysis. *Research Result. Theoretical and Applied Linguistics*, *9*(1), 82–98. https://doi.org/10.18413/2313-8912-2023-9-1-0-6

Dreisoerner, A., Munker, N., Scholtz, W., Heimrich, J., & Bloemeke, S. (2021). Self-soothing touch and being hugged reduce cortisol responses to stress: A randomized controlled trial on stress, physical touch, and social identity. *Comprehensive Psychoneuroendocrinology*, *8*, Article 100091. https://doi.org/10.1016/j.cpnec.2021.100091

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*, 49–98. https://doi.org/10.1515/semi.1969.1.1.49

Fisher, J. B., Lohmer, V., Kern, F., Barthlen, W., Gaus, S., & Rohlfing, K. (2022). Exploring monological and dialogical phases in naturally occurring explanations. *Künstliche Intelligenz*, *36*, 317–326. https://doi.org/10.1007/s13218-022-00787-1

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. W. Siegman & B. Pope (Eds.), *Studies in Dyadic Communication* (pp. 153–175, Vol. 7). Pergamon.

Grunwald, M., Weiss, T., Mueller, S., & Rall, L. (2014). Eeg changes caused by spontaneous facial self-touch may represent emotion regulating processes and working memory maintenance. *Brain Research*, *1557*, 111–126. https://doi.org/10.1016/j.brainres.2014.02.002

Haas, J., Warnke, V., Niemann, H., Cettolo, M., Corazza, A., Falavigna, D., & Lazzari, G. (1999). Semantic boundaries in multiple languages. *6th European Conference on Speech Communication and Technology*, 535–538. https://doi.org/10.21437/Eurospeech.1999-138

Harrigan, J. A. (1985). Self-touching as an indicator of underlying affect and language processes. *Social Science & Medicine*, *20*(11), 1161–.1168. https://doi.org/10.1016/0277-9536(85)90193-5

Iriskhanova, O., Kiose, M., Leonteva, A. V., & Agafonova, O. (2023). Vague reference in expository discourse: Multimodal regularities of speech and gesture. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2023´´*.

Kang, S., Tversky, B., & and, J. B. B. (2015). Coordinating gesture, word, and diagram: Explanations for experts and novices. *Spatial Cognition & Computation*, *15*(1), 1–26. https://doi.org/10.1080/13875868.2014.958837

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press. https://doi.org/10.1017/cbo9780511807572

Kimura, D. (1976). 4 - The Neural Basis of Language Qua Gesture. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in Neurolinguistics* (pp. 145–156). Academic Press. https://doi.org/10.1016/B978-0-12-746302-5.50011-6

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Klein, J. (2009). Erklären-was, Erklären-wie, Erklären-warum. Typologie und Komplexität zentraler Akte der Welterschließung. In R. Vogt (Ed.), *Erklären. Gesprächsanalytische und fachdidaktische Perspektiven* (pp. 25–36). Stauffenburg-Verlag.

Kotthoff, H. (2009). Erklärende aktivitätstypen in alltags- und unterrichtskontexten. In J. Spreckels (Ed.), *Erklären im kontext. neue perspektiven aus der gesprächs- und unterrichtsforschung* (pp. 120–146). Schneider.

Lazarov, S. (2025). *The reflection of interactional monitoring in the dynamics of verbal and nonverbal forms of explaining* [Dissertation]. Universität Paderborn. https://doi.org/10.17619/UNIPB/1-2446

Lazarov, S., Biermeier, K., & Grimminger, A. (2024). Changes in the topical structure of explanations are related to explainees' multimodal behaviour. *Interaction Studies*, *25*(3), 257–280. https://doi.org/10.1075/is.23033.laz

Lazarov, S., Türk, O., Grimminger, A., Wagner, P., & Buschmeier, H. (2025). Annotation schemes project A02 "Monitoring the understanding of explanations". osf.io/j2dha

Lopez-Ozieblo, R. (2025). The dual functions of adaptors. *Languages*, *10*(9). https://doi.org/10.3390/languages10090231

Maricchiolo, F., Gnisci, A., & Bonaiuto, M. (2012). Coding hand gestures: A reliable taxonomy and a multi-media support. In A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann, & V. C. Müller (Eds.), *Cognitive Behavioural Systems* (pp. 405–416). Springer.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226514642.001.0001

Mueller, S. M., Martin, S., & Grunwald, M. (2019). Self-touch: Contact durations and point of touch of spontaneous facial self-touches differ depending on cognitive and emotional load. *PLOS ONE*, *14*(3), 1–21. https://doi.org/10.1371/journal.pone.0213677

Ricciardi, O., Maggi, P., & Nocera, F. D. (2019). Boredom makes me 'nervous': Fidgeting as a strategy for contrasting the lack of variety. *International Journal of Human Factors and Ergonomics*, *6*(3), 195–207. https://doi.org/10.1504/IJHFE.2019.104590

Rohlfing, K. J., Cimiano, P., Scharlau, I., Matzner, T., Buhl, H., Buschmeier, H., Grimminger, A., Hammer, B., Häb-Umbach, R., Horwath, I., Hüllermeier, E., Kern, F., Kopp, S., Thommes, K., Ngonga Ngomo, A.-C., Schulte, C., Wachsmuth, H., Wagner, P., & Wrede, B. (2021). Explanation as a social practice: Toward a conceptual framework for the social design of ai systems. *IEEE Transactions on Cognitive and Developmental Systems*, *13*, 717–728. https://doi.org/10.1109/TCDS.2020.3044366

Roscoe, R., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350. https://doi.org/10.1007/s11251-007-9034-5

RStudio Team. (2020). *Rstudio: Integrated development environment for R*. RStudio, PBC. Boston, MA. http://www.rstudio.com/

Sasaki, J., & Sasaki, G. (2014). *Deep sea adventure (tabletop game)*. Oink Games.

Sekine, K., & Hotta, H. (2025). The role of self-adaptors in lexical retrieval. *Languages*, *10*(9), 209. https://doi.org/10.3390/languages10090209

Seli, P., Carriere, J. S. A., Thomson, D. R., Cheyne, J. A., Martens, K. A. E., & Smilek, D. (2014). Mind wandering and performance: Why mind wandering occurs during easy and difficult tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 660–668. https://doi.org/10.1037/a0035841

Spille, J. L., Grunwald, M., Martin, S., & Mueller, S. M. (2022). The suppression of spontaneous face touch and resulting consequences on memory performance of high and low self-touching individuals. *Scientific Reports*, *12*(1), 8637. https://doi.org/10.1038/s41598-022-12044-4

Spille, J. L., Mueller, S. M., Martin, S., & Grunwald, M. (2022). Cognitive and emotional regulation processes of spontaneous facial self-touch are activated in the first milliseconds of touch: Replication of previous eeg findings and further insights. *Cognitive, Affective & Behavioral Neuroscience*, *22*(5), 984–1000. https://doi.org/10.3758/s13415-022-00983-4

Spille, J. L., Grunwald, M., Martin, S., & Mueller, S. M. (2021). Stop touching your face! a systematic review of triggers, characteristics, regulatory functions and neurophysiology of facial self touch. *Neuroscience & Biobehavioral Reviews*, *128*, 102–116. https://doi.org/10.1016/j.neubiorev.2021.05.030

Türk, O., Lazarov, S., Wang, Y., Buschmeier, H., Grimminger, A., & Wagner, P. (2024). Predictability of understanding in explanatory interactions based on multimodal cues. *Proceedings of the 26th International Conference on Multimodal Interaction*, 449–458. https://doi.org/10.1145/3678957.3685741

Türk, O., Wagner, P., Buschmeier, H., Grimminger, A., Wang, Y., & Lazarov, S. (2023). MUNDEX: A multimodal corpus for the study of the understanding of explanations. *Proceedings of the 1st International Multimodal Communication Symposium*, 63–64. http://nbn-resolving.de/urn:nbn:de:0070-pub-29805458

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings the 5th International Conference on Language Resources and Evaluation*, 1556–1559. https://aclanthology.org/L06-1082/