Factor Analysis of Gazing Activities in Native and Second Language Conversations

Ichiro Umata[1], **Koki Ijuin**[2] and **Seiichi Yamamoto**[2]

[1]KDDI Research, Inc., Interaction Design Group, Tokyo, Japan
[2]Doshisha University, Department of Information Systems Design, Kyoto, Japan
`ic-umata@kddi-research.jp; euq1101@mail4.doshisha.ac.jp;`
`seyamamo@mail.doshisha.ac.jp`

Abstract

Gazing activities during utterances and silence were analyzed in a face-to-face three party conversation setting in a native language (L1) and in a second language (L2). The function of each utterance was categorized (Traum, 1994) so that gazes during utterances could be analyzed from the viewpoint of grounding in communication (Clark, 1996). The result of a factor analysis suggests that language difference is a dominative factor that affects gazing activities in communication.

1 Introduction

Gaze combines many functions in communication. Previous studies have observed that gaze helps coordinate turn-taking (Duncan, 1972) (Kendon, 1967), establish a given piece of information as part of common ground (Clark and Brennan, 1991) (Clark, 1996), and express intimacy (Mehrabian and Ferris, 1967), and the condition of conversational setup might change the relative importance of these functions (Kleinke, 1986).

In this study, we examine the effect of language difference on gazing activities in communication. The result of a factor analysis for gazing activities shows that language difference is a dominative factor that affects gazing activities in communication. The result suggests that multimodal communication support systems processing gaze information must take such effects of linguistic proficiency into consideration.

2 Data Collection

We analyzed data from the goal-oriented task in which the interlocutors collaboratively decided what to take with them on a trip to a deserted island or the mountains (for details, refer to (Yamamoto et al., 2015)). Each group had six-minute conversations on goal-oriented topics in both Japanese and English. The data contains multimodal conversations from 40 (20 goal-oriented in Japanese, and 20 goal-oriented in English) three-party conversations in L1 (Japanese) and in L2 (English) languages (Yamamoto et al., 2015). All participants were native-Japanese speakers whose second language was English. Three sets of NAC EMR-9 head-mounted eye trackers and headsets with microphones recorded their eye gazes and voices. The EUDICO Linguistic Annotator (ELAN) developed by the Max Planck Institute was used as a tool for gaze and utterance annotation. The utterances were annotated with Grounding Act tags established by Traum (Traum, 1994) for 20 groups of goal-oriented conversations (Umata et al., 2016).

3 Analysis I: Factor Analysis of Gazing Activities

We conducted a factor analysis for gazing activities of each communication channels in each group under the assumption that gazes are strongly affected by the language difference. Three participants (ex. A, B, C) formed a group, and we defined six communication channels in a group (i.e. A → B, A → C, B → A, B → C, C → A, C → B). Gazes during silence and during utterances with one of the four major Grounding Act tags (i.e., init, ack init, cont, ack) were subject to the analysis because there were very few occurrences of others (i.e., utterances with repair, reqRepair, reqAck, and cancel tags) (Umata et al. 2016). We define the indices of gazing activities via a communication channel between participant j and k during silence and utterances. The average of gaz-

ing ratio during silences (SILGR) is defined as follows:

Average of gazing ratios during silences (SILGR):

$$SILGR = \frac{\sum_{i=1}^{n} DSILG_{jk}(i)}{\sum_{i=1}^{n} S(i)}$$

Here, $DSILG_{jk}(i)$ is the duration when a participant j is looking at a participant k in the duration of the *i-th* silence $S(i)$. The average of speaker's gazing ratios (SGR) is defined as follows:

Average of speaker's gazing ratios (SGR):

$$SGR = \frac{\sum_{i=1}^{n} DSG_{jk}(i)}{\sum_{i=1}^{n} D_j(i)}$$

Here, $DSG_{jk}(i)$ is the total duration when the speaker j is gazing at a participant k in the duration of the *i-th* utterance by j. The average of listenerr's gazing ratios (LGR) is defined as follows:

Average of listener's gazing ratios (LGR):

$$LGR = \frac{\sum_{i=1}^{n} DLG_{jk}(i)}{\sum_{i=1}^{n} D_j(i)}$$

Here, $DLG_{jk}(i)$ is the total duration when the listener k is gazing at the speaker j in the duration of the *i-th* utterance by j.

We conducted factor analysis of gazes during silence, speakers' gazes (SGR) and listener's gazes (LGR) during utterances with four major grounding act tags in L1 and L2 conversations. A participant without a cont utterance in L1, one without an ackInit utterance in L2, and one without a cont utterance in L2 were excluded from the analysis. Factors were extracted by the principal factor method, and promax rotation was adopted. Five factors were extracted by giving consideration to the decay of the eigenvalues.

The factor structure of gazing activities shows that the language difference affects the gazing activities stronger than the utterance functions do. The first factor (FI) is characterized by high loading of the gaze during silence and the speaker's gazes in L1, and the second one (FII) is characterized by high loading of the gaze during silence and the speaker's gazes in L2. The third factor (FIII) is characterized by high loading of the listener's gazes other than during ack utterances in L1, and the fourth one (FIV) is characterized by high loading of the listener's gazes other than during ack

utterances in L2. The fifth factor (FV) is characterized by high loading of the listener's gazes during ack utterances both in L1 and L2. The factor correlations are high between FI and FII, and moderately high between FIII and FIV. FV also show moderately high correlation between FIII and FV.

4 Summary

We examine the effect of language difference on gazing activities in communication. The result of a factor analysis for gazing activities shows that language difference is a dominative factor that affects gazing activities in communication. The result suggests that multimodal communication support systems processing gaze information must take such effects of linguistic proficiency into consideration.

References

H. H. Clark and S. E. Brennan. 1991. Grounding in communication. In L. B. Resnik, J. M. Levine, and S. D. Teasley, editors, *Perspectives on socially shared cognition*, pages 503–512. APA Books.

H. H Clark. 1996. *Using language*. Cambridge University Press.

S. Duncan. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23:283–292.

A. Kendon. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63.

C. L. Kleinke. 1986. Gaze and eye contact: a research review. *Psychological Bulletin*, 100:78–100.

A. Mehrabian and S. R. Ferris. 1967. Decoding of inconsistent communications. *Journal of Personality and Social Psychology*, 6:109–114.

D. Traum. 1994. *A computational theory of grounding in natural language conversation*. Ph.D. thesis, University of Rochester.

I. Umata, K. Ijuin, M. Ishida, and S. Yamamoto. 2016. Quantitative analysis of gazes and grounding acts in l1 and l2 conversations. In *Proceedings of 10th edition of the Language Resources and Evaluation Conference (LREC 2016)*, pages 4249–4252.

S. Yamamoto, K. Taguchi, K. Ijuin, I. Umata, and M. Nishida. 2015. Multimodal corpus of multiparty conversations in l1 and l2 languages and findings obtained from it. *Language Resources and Evaluation*, 49:857–882.