Nonverbal Expressions of Turn Management in TV Interviews: A Cross-Cultural Study between Greek and Icelandic

Maria Koutsombogera[1], Sigrún María Ammendrup[2], Hannes Högni Vilhjálmsson[2], and Harris Papageorgiou[1]

[1] Institute for Language & Speech Processing, Artemidos 6&Epidavrou, 15125 Athens, Greece,
`{mkouts,xaris}@ilsp.gr`
[2] School of Computer Science, Reykjavik University, Menntavegur 1, 101 Reykjavík, Iceland
`{sma,hannes}@ru.is`

Abstract. In this paper we discuss a cross-cultural analysis of non-verbal expressions (gestures, facial expressions, body posture) that have a turn managing function in the flow of interaction. The study was carried out by analyzing and comparing the features of interest in two samples of institutional interaction, namely face-to-face political interviews, in Greek and Icelandic respectively. The non-verbal behavior of the participants in both interviews was annotated following the same annotation process. The attested turn management instances were compared in order to find similarities and differences in terms of frequency and modality preference.

Keywords: turn management, non-verbal expressions, institutional interaction.

1 Introduction

In this paper we study the multimodal behavior of interview participants in two different languages and cultures, Greek (GR) and Icelandic (IC). We focus on the non-verbal expression of turn taking, i.e. the non verbal expressions employed to regulate the flow of interaction, as attested in representative samples of TV interviews. Specifically, we explore if there are diversities -and to what extent- across two seemingly different cultures in the way participants manage the exchange of their turns non-verbally, in the domain of institutional interaction, namely political interviews.

Media talk in general echoes the tendencies in media discourse and thus may contain indicators of culturally driven relations not only between media and politics [1], but also between discourse practices and the overall multimodal interaction among the participants involved. The non verbal channel is inherently related to speech and discourse/conversational structures [2]; non-verbal features are incorporated in discourse practices and can highlight the speakers' behaviour as well as yield insights into culturally specific cues. So, the discourse representation used in the interactional functions across the two cultures may largely affect the production of the respective non-verbal behaviour. What we expect to find are variations in the semiotics of this behavior, namely the choice/ preference of specific facial, hand or torso movements as well as their frequency.

Several comparative studies have been conducted in order to explore differences in how non-verbal expressions are related to speech. Some examples involve the study of the gestural behavior of speakers belonging to different ethnic groups [3] social groups [4], native& non-native speakers of a specific language [5], and speakers whose language is of different structure and thus their non-verbal expressions fit aspects of the events that are linguistically expressed [6, 7].

The main points from the aforementioned studies are that differences may be due to the structure of the spoken language or to practices related to the interaction conduct per se. It is worth mentioning that, while the data studied in the literature so far pertain to the discourse types of either everyday interaction or narrations, our data, the political interviews, are a sample of institutional interaction. Even so, our expectations from the present study are that interview participants from different cultures make different non-verbal expressions.

However, according to the literature, the methodological outcome of conducting such comparisons is that the attention should be drawn not only to language differences but also to differences in culture or social classes within the culture and, finally, differences in discourse settings. Keeping in mind the above principles and setting aside the pure linguistic analysis, we will closely keep track of the interplay between cultural differences and discourse setting in order to discover which one of those principles is determinant over the non-verbal turn management behavior of the speakers.

2   Data Collection

In the current study, the comparison is focused on the non-verbal attestations of the communicative dimension of turn taking in political interviews between 2 different languages and cultures (GR-IC). In order to identify possible differences or similarities, we studied data taken from political TV interviews. Data in both languages share a common basis regarding (a) the setting of the interaction, namely live broadcast TV political interviews taking place in a studio and pertaining to an institutional type of interaction, (b) the identity of the interview participants; the hosts in both interviews are professional journalists, while the guests are government representatives, the Minister of Industry in the IC interview and the Minister of Public Order in the GR interview respectively, (c) the duration of the interviews, approx. 17 minutes each. The topic of discussion is different in each interview, namely power plants for aluminum companies in IC and police and public order in GR, both very current and controversial issues.

3   Annotation Process

The data we studied are TV political, face-to-face interviews multimodally annotated. In each interview, the audio signal is transcribed[1]. The subsequent video annotation[2] deals with the labeling of the non-verbal expressions (facial displays, hand gestures and torso movements) co-occurring with speech at two levels: (a) identification and

---

[1] http://trans.sourceforge.net/
[2] http://www.lat-mpi.eu/tools/elan/

boundary annotation on the time axis and, (b) assignment of turn management labels. The levels and labels used in the annotation scheme are mainly inspired by the MUMIN v3.3 coding scheme notation [8]. Both audio and visual signals as well as the annotations are perfectly synchronized; the overall set of annotation levels is distinguished by speaker, and all the annotation metadata are integrated into a single xml file.

Table 1. Annotation scheme used; for each feature of face, gesture and torso movements, a detailed set of labels is used to capture the exact form of the non-verbal expression employed

Facial display	Hand gesture	Body posture	Turn management
Gaze	Handedness	Torso	Turn take
Eyes	Trajectory		Turn accept
Eyebrows			Turn yield
Mouth			Turn offer
Lips			Turn complete
Head			Turn hold

4   Turn Taking in Political Interviews

The turn-taking mechanism enables the speakers to manage the smooth and appropriate exchange of speaking turns in face-to-face interaction. Turn taking organization has been investigated thoroughly in order to discover the detailed rules of conversational behavior [9, 10] and the constraints imposed by the context. Specifically, it has been examined how turns are introduced, the forms they may have and the devices by which they are expressed in terms of prosody, discourse strategies and non-verbal cues.

Turn taking is expressed actively through content (words), intonation, paralanguage, and non-verbal expressions [11]. We believe that the multimodal analysis of such behaviors provides significant information and accurate observations besides the study of speech only. Non-verbal expressions do not simply accompany speech but they are indicators of the degree of success of the speakers' intentions and projections and shed light on the strategies used for the accomplishments of the interaction. They are signals by which each participant indicates his/her state with regard to the speaking turn.

Turn taking in political interviews is largely affected by the situational and conversational constraints that are imposed by the institutional frame in which this genre pertains [12, 13].The interview can be regarded as an institutionalized discourse type, because it appears to be highly controlled and conventionalized, constrained by institutional role-distribution and turn pre-allocation and is less prone to spontaneous interventions. Consequently, the non-verbal behavior should follow the principles of this specific frame.

Situationally, the interview takes place in a particular setting, addressing an overhearing audience and complying with several talk-related restrictions, such as a predetermined agenda, time restrictions which lead to monitored turns, speaker selection restrictions, and turn-taking restrictions; the host is primarily responsible for selecting

the next speaker and for coordinating the turn-taking sequences; and finally discursive constraints such as talk-framing patterns.

5   Data Analysis

5.1   Non-verbal Turn Management: Frequency

The study of the annotated interviews, of approx. 17 minutes duration each, shows that speakers in both interviews frequently employ non-verbal expressions (NVEs) throughout their turns. Although the total number of attested NVEs is much bigger for the GR interview (982 instances) compared to the IC interview (511 instances), the NVEs related to turn management communicative function are comparable in both interviews (cf. Table 2).

The remaining NVEs are related to other communicative functions that do not form part of this study, such as feedback, emotions/ attitudes expression, content-related NVEs (e.g. iconic gestures) and so on. We therefore assume that the difference in the overall number of NVEs is attributed to instances which are not related to turn management and that GR speakers are more productive regarding this aspect.

Table 2. Distribution of attested non-verbal expressions; TM: related to turn management, other: not related to turn management, total: overall

	Interviewer				Interviewee			
Language	**GR**		**IC**		**GR**		**IC**	
		[Per Turn]		[Per Turn]		[Per Turn]		[Per Turn]
TM NVEs	147	1.27	112	1.22	137	1.18	139	1.51
other	193	1.66	42	0.45	505	4.35	218	2.37
total NVEs	340	2.93	154	1.67	642	5.53	357	3.88

In the Greek data the total number of turns amounts to 116 while in the Icelandic data to 92. Both interviews participants use non-verbal expressions to regulate their turns according to their interactional needs and goals. There is roughly the same rate of TM NVE production (i.e. NVEs/turn) by both GR and IC interviewers, namely 1.27 NVEs/turn versus 1.22 NVEs/turn respectively. However, the IC interviewee seems to be producing a few more TM NVEs per turn than the GR one with a rate of 1.51 and 1.18 respectively. Further experimentation and data are needed in order to explore whether this difference is due to cultural traits.

The majority of the turn management NVEs are employed by the GR interviewer to express turn offer (37.4%), by the IC interviewer to express turn take (60.7%), by the GR interviewee to express turn accept (46%) and by the IC interviewee to express turn hold (46%).

5.2   Non-verbal Turn Management: Preferred Modalities

The goal of non-verbal turn management may be achieved by single face, hand or torso movements, or through a combined use of distinct modalities, either in a synchronized

or a partly overlapping manner. In this study, all figures refer to unique instances of NVEs performed by a single modality.

Although speakers employ most of their expressive means throughout turn management, they show a certain affinity towards particular modalities, as depicted in Figure 1. Specifically, the GR interview participants clearly prefer the facial displays (50.3% for the interviewer and 73.7% for the interviewee), while in the IC interview the preference is shared between gestures for the interviewer (46.4%) and facial displays for the interviewee (58.3%).
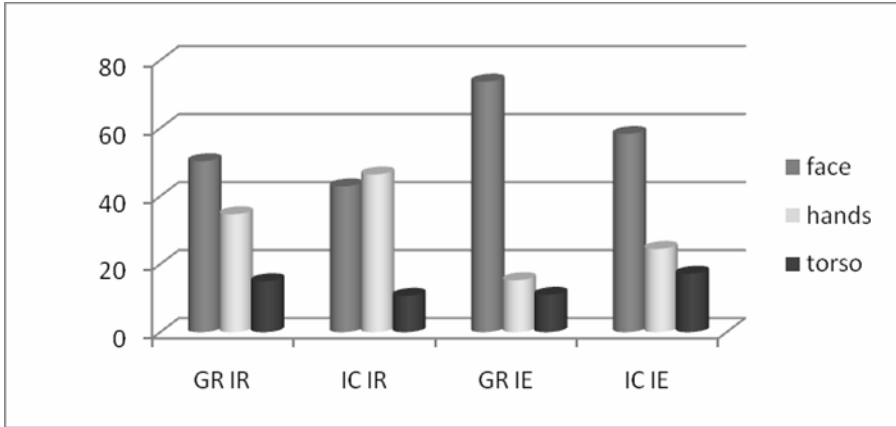

Fig. 1. Percentage distribution of non-verbal expressions of turn management in Greek (GR) and Icelandic (IC) interviewers (IR) and interviewees (IE) respectively

If we go into more detail at the modality subtypes, we attest certain variations regarding the choice of the most appropriate NVE according to the interview participants. Table 3 depicts the preferred modality subtypes. Usually, there is only one

Table 3. Preferred modality subtypes per turn management dimension by the interview participants

Turn Management dimension	GR Interview		IC Interview	
	Interviewer	Interviewee	Interviewer	Interviewee
Turn Take	Head Tilt	Eyebrows raise	Gaze Down	Gaze Down
Turn Accept	Smile	Head Tilt, Gaze Towards	Index Finger	Gaze Down
Turn Yield	Head Nod	Smile	Gaze Down, Gaze Towards	Eyebrows raise
Turn Offer	Head Tilt	—	Gaze Down, Gaze Towards	Eyebrows raise
Turn Complete	Smile, Gaze Towards	Head Nod, Smile	—	—
Turn Hold	—	Head Tilt, Gaze Down	Gaze Down, Gaze Towards, Gaze Sideways	Gaze Down

dominant subtype. In some cases though, some turn management dimensions were highly represented by more than one subtypes, as is the case of turn hold regarding the IC interviewer. Another interesting finding is the fact that the predominant NVE in the Icelandic data is gazing down. Again, more data are needed to interpret this NVE preference and its possible linking to culture-specific behaviour. Finally, the table cells with no entry in them denote that either there were no NVEs catering for the specific turn management dimension, or there was no significant lead of one subtype over another.

6 Discussion

As shown in Section 5, participants in both interviews share roughly the same rate of turn management NVEs. This was not expected, as it is a common belief that the further north you go in Europe, the more reserved and less overtly expressive people become. However, when we examined the most prominent NVEs for each turn management function (Table 3) we realized that there are some differences in the form of turn management between the two interviews. More data is required to substantiate whether the form observed here is characteristic of the culture, but it is interesting to see here how different modalities can serve the exact same function. Further qualitative analysis of non-verbal expression properties such as intensity, complexity and duration will also be the focus of our future work.

Thus, our initial expectations that there should be notable differences between multimodal behaviors among the two languages and cultures were not verified. Instead, non-verbal turn management seems to be coordinated to the discourse strategies that the interviewers use to portray the political situation and its representatives during an interview. The political interview is a strictly framed discourse setting, institutional & conventionalized, different from casual conversation where speakers can express themselves in a more spontaneous and unrestricted manner. Before drawing any general conclusions about the cultural differences regarding non-verbal behavior, it remains to be seen whether switching to another, less restricted interactional domain or setting (like casual conversation) would provide more evidence about culture-specific practices and conventions.

References

1. Lauerbach, G.: Discourse representation in political interviews: The construction of identities and relations through voicing and ventriloquizing. Journal of Pragmatics 38(2), 196–215 (2006)
2. McNeill, D.: Hand and Mind: What Gestures Reveal About Thought. University of Chicago Press, Chicago (1992)
3. Efron, D.: Gesture, Race and Culture. Mouton, The Hague (/1941/1972)
4. Kendon, A.: Geography of gesture. Semiotica 37(1/2), 129–163 (1981a)
5. Kendon, A.: Gesture: Visible Action as Utterance. Cambridge University Press, Cambridge (2004
6. McNeill, D., Duncan, S.: Growth points in thinking-for-speaking. In: McNeill (ed.) Language and Gesture, pp. 141–161. Cambridge University Press, Cambridge (2000)

7. Kita, S., Özyürek, A.: What Does Cross-Linguistic Variation in Semantic Coordination of Speech and Gesture Reveal? Evidence for an Interface Representation of Spatial Thinking and Speaking. Journal of Memory and Language 48, 16–32 (2003)
8. Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., Paggio, P.: The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena. Multimodal Corpora for Modeling Human Multimodal Behaviour. Journal on Language Resources and Evaluation 41(3-4), 273–287 (2007)
9. Sacks, H., Schegloff, E., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language 30, 696–735 (1974)
10. Ford, C.E., Thompson, S.A.: Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns. In: Ochs, E., Schegloff, E.A., Thompson, S.A. (eds.) Interaction and Grammar, pp. 134–184. Cambridge University Press, Cambridge (1996)
11. Duncan Jr., S.: Some Signals and Rules for Taking Speaking Turns in Conversations. Journal of Personality and Social Psychology 23(2), 283–292 (1972)
12. Heritage, J.: Conversation Analysis and Institutional Talk. In: Sanders, R., Fitch, K. (eds.) Handbook of Language and Social Interaction, pp. 103–146. Lawrence Erlbaum, New Jersey (2005)
13. Ilie, C.: Semi-institutional Discourse: The Case of Talk Shows. Journal of Pragmatics 33, 209–254 (2001)