

A Corpus-based Study to Disambiguate Discourse Markers in Multimodal Context¹

Abuczki Ágnes

The present study aims at identifying the sequential and nonverbal features that typically characterize and best distinguish the different functions of the non-conceptual uses of three Hungarian lexical items (*mondjuk* (~'say'); *ugye* (~'obviously', 'isn't it?', 'right?'); *amúgy* (~'otherwise', 'by the way')) that can be considered multifunctional discourse markers (henceforth DMs). DMs are generally defined as “sequentially dependent elements which bracket units of talk” and “provide contextual coordinates for ongoing talk” that indicate for the hearer how an utterance is to be interpreted (Schiffrin 1987: 31). Previous research on discourse segmentation has found that DMs, particles, cue phrases and nonverbal cues all give indispensable clues to discourse structure interpretation (Grosz & Sidner 1986). Consequently, the disambiguation of the functions of DMs would be useful for the automatic segmentation and understanding of dialogues, too. It is argued in this study that a multimodal approach is indispensable in communication modeling in order to disambiguate the actual meaning of polysemous communicative signals such as DMs. The material of the study is comprised of 6 hours of spontaneous conversation (10 simulated job interviews and 10 informal conversations) from the Hungarian HuComTech corpus between a constant agent and 10 different young speakers (university students between 18-25 years of age). The HuComTech corpus is annotated at multiple multimodal levels in Praat (Boersma & Weenink 2007) for the audio material and in Quannot (Pápay, Szeghalmy & Szekrényes 2011), a custom designed environment for the video material. At the discourse level of its annotation, the transcribed dialogue is segmented into floor control types (turn-take, turn-keep, turn-give, backchannel). The video annotation of the corpus involves the labeling of facial expressions, gaze directions, eyebrow positions, head movements, handshape types, postures, deictic gestures and emblems. This subcorpus contains 208 tokens of *mondjuk* ('say'), 70 tokens of *ugye* ('obviously', 'isn't it?', 'right?'), and 33 tokens of *amúgy* ('otherwise'). After importing and uniting earlier (audio and video) annotations (Praat TextGrids and .qnt files) of the material in the ELAN 4.5.1 software (Brugman & Russel 2004), the discourse markers were segmented and functionally indexed using the same (ELAN) software. The corpus queries (e.g. Find overlapping labels, N-gram within annotations) in ELAN address the analyses of their contextual environment (lexical co-occurrences, presence or absence of surrounding silence), position in the utterance, prosodic features (duration, mean fundamental frequency, direction of pitch movement) and nonverbal-visual markers (the presence or absence of accompanying hand movements, gaze direction and affect displays of the speaker). The results of multimodal corpus queries and the statistical tests (Pearson's chi-square test, Crosstabs test, Fischer's exact test, independent samples t-test, paired t-test as well as box plot graphs) performed on them in SPSS 19.0 aim to identify the machine-detectable features of the different uses of DMs and distinguish the two most salient functional categories of each of the three DMs analyzed: 1.a. lexical search/approximation (as own speech management functions) versus 1.b. contrast/concession (as discourse-pragmatic relations between two segments) (expressed by *mondjuk*); 2.a. question as a directive act versus 2.b. explanation as a

¹ This research was supported by the **European Union** and the **State of Hungary, co-financed by the European Social Fund** in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'.

constative act (marked by *ugye*); and 3.a. topic change versus 3.b. topic elaboration/commenting (introduced by *amúgy*). The findings suggest that the defining properties distinguishing different functions are the duration of the DM and the simultaneous performance or cessation of manual gesticulation in all 3 DMs. In the case of *ugye*, gaze direction is also a distinguishing feature, while in the case of *mondjuk*, facial expressions of the speaker also help to disambiguate the actual function of the DM. Position has also been found to influence the actual function and the direction of pitch movement in the DM and its host unit. On the other hand, no relationship has been found either between preceding silence and the function of a DM or between the mean F0 and the function of a DM .

References

Boersma, P. & Weenink, D. (2007) *Praat: doing phonetics by computer 5.0.02*. University of Amsterdam: Institute of Phonetic Sciences, <http://www.praat.org>

Brugman, H. & Russel, A. (2004) Annotating multi-media/ multi-modal resources with elan. In: Lino, M., Xavier, M., Ferreire, F., Costa, R., Silva, R. (Eds.) *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC)*. Lisbon: Portugal. 2065–2068. (ELAN 4.5.1 software was used in the study. It can be downloaded at <http://tla.mpi.nl/tools/tla-tools/elan/download/>).

Grosz, B. & Sidner C. (1986) Attention, intentions, and the structure of discourse, *Computational Linguistics*, vol. 12, 175–204.

Pápay, K., Szeghalmy Sz. & Szekrényes I. (2011) HuComTech Multimodal Corpus Annotation. *Argumentum*, 7. Debrecen: Debreceni Egyetemi Kiadó. 330–347.

Schiffrin, D. (1987) *Discourse Markers*. Cambridge: Cambridge University Press.