

# Multimodal Analysis of “*well*” as a Discourse Marker in Conversation: A Pilot Study

G. Esfandiari Baiat<sup>1</sup>, M. Coler<sup>2</sup>, M. Pullen<sup>3</sup>, S. Tienkouw<sup>4</sup>, L. Hunyadi<sup>1</sup>

<sup>1</sup> Department of General and Applied Linguistics, University of Debrecen, Debrecen, Hungary

<sup>2</sup> Cognitive Systems Group, INCAS, The Netherlands

<sup>3</sup> Department of Linguistics and Web Technology, Phillips-Universität Marburg, Marburg, Germany

<sup>4</sup> National Electronics and Computer Technology Center (NECTEC), Ministry of Science and Technology, Thailand

**Abstract**— Discourse markers, a group of pragmatic elements, are used widely in our daily communication. Among these markers, *well* is found to be the most frequently used one in conversation. Although much research has been carried out on its theoretical bases such as its definition, functions and categorizations, few studies have taken a multimodal approach in analyzing this marker. The aim of the present study was to provide a multimodal description of the use of *well* as a discourse marker in spontaneous speech. Firstly, we wanted to investigate whether the use of this marker in conversation, regardless of its function, was accompanied with any kind of non-verbal behavior (posture shift, head shift, gaze and eyebrow movement) and secondly, we wanted to determine whether any of these non-verbal cues (if any) were specific to one pragmatic function of *well* and if they could be used as a signal to disambiguate the pragmatic function of *well*. The study of discourse markers is considered important due to the fact that it can contribute both to dialogue interpretation and generation resulting in a more natural dialogue modeling useful for human-infocommunications interface.

**Keywords**— Multimodal Communication, Discourse Structure, Discourse Markers, *Well*, Pragmatic Functions

## I. INTRODUCTION

In spoken interaction, speakers often use variety of signals and strategies, both verbal and non-verbal, to communicate successfully. The use of discourse markers (henceforth DMs) such as *but*, *and*, *so*, *well*, *you know*, *I mean* or *by the way* can be considered among such signals used to maintain coherence relations between the discourse segments in speech. As [1] states “the analysis of discourse markers is regarded as a part of the more general analysis of discourse coherence- how speakers and hearers jointly integrate forms, meanings, and actions to make overall sense out of what is said.” Consequently, in order to understand the structure of spoken discourse studying these coherent devices are considered crucial.

Understanding the features and functions of these markers enables us to achieve and construct a natural model of human-human communication resulting in a more natural dialogue modeling for human-machine communication. If the cognitive capabilities of a system can be augmented by allowing it to be more sensitive to these coherence devices, it will allow

humans to communicate with the system in a way that is closer to their own level of cognitive capabilities.

Among these markers, *well* is found to be the most frequently used one in conversation [1]. Although much research has been carried out on its definition, functions and categorizations, few studies have taken a multimodal approach in analyzing this marker [2]. Conversations, being more than simply a sequence of utterances strung together, are indeed considered as having a multimodal process [3]. There are some non-verbal communicative cues that lend naturalness to speech. These cues play an important role in human-human as well as in human-machine communication. They help to compensate for many hidden meanings not present on the surface of spoken language [4]. Thus, in order to fully understand, interpret and describe the structure of *well* used in spoken discourse, studying these non-verbal elements is considered to be crucial. Understanding the role of nonverbal behaviors in conveying conversation structure enables improvements in the naturalness of embodied dialogue systems, such as embodied conversational agents [5].

The goal of this study was twofold: firstly, to provide a multimodal description of the use of *well* as a discourse marker in spontaneous speech. We wanted to investigate whether the use of this marker in conversation, regardless of its function, was accompanied with any kind of non-verbal behavior (posture shift, head shift, gaze and eyebrow movement); secondly, we wanted to determine whether any of these non-verbal cues (if any) were specific to one function of *well* and if they could be used as a signal to disambiguate its pragmatic function.

In the following sections, considering the most prominent approaches in this field of study, a brief definition of DMs in general will be given followed by a specific focus on *well* and its most dominant pragmatic functions.

## II. RELEVANT TERMINOLOGIES

### A. Defining Discourse Markers

DMs have received a great deal of attention in the study of pragmatics for over twenty years. Although there is a general agreement among researchers on the primary function of discourse markers (i.e. discourse structuring), their view differs regarding inclusion and exclusion of particular

The present study is supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0002 project

linguistic items in the category of DMs [2]. Consequently, some definitions are more exclusive than others [6]. For example, [7] defines DMs as items that indicate “a sequential relationship between the current basic message and the previous discourse.” Based on this definition he includes *well* in turn-initial positions while excluding it when used as a pause marker in turn-internal positions. On the other hand, [8] takes an inclusive approach and defines DMs as “expressions that are used with the primary function of bringing to the listener’s attention a particular kind of attention between the discourse unit it introduces and the immediate discourse context.” Reference [8] includes items such as *frankly*, *obviously*, *stupidly* that [7] would label as pragmatic markers rather than DMs [6].

According to the literature, the most influential definition is given by [1] who concentrates on DMs’ role in “providing contextual coordinates for ongoing talk” and defines them as “sequentially dependent elements which bracket units of talk” for example, sentences or speech acts. Schiffin’s broad view of DMs also includes nonverbal devices (manual and facial gestures, posture changes, gaze direction, etc.) in addition to verbal markers [1]. Due to the multimodal nature of our investigation, we will employ Schiffin’s approach towards DMs throughout this study.

### III. PRAGMATIC FUNCTIONS OF “WELL” BASED ON PREVIOUS RESEARCH

*Well* as a DM is frequently used in conversation. Reference [1] believes the use of *well* in conversation is not based on any semantic meaning or grammatical statuses. Previous studies propose the following six functions as the major pragmatic functions of *well* as a DM [1] [6] [9] [11] [12].

#### A. *Well* as a Response Marker or as an Initiation Marker

According to [1], *well* is considered primarily a response marker “which anchors its user in an interaction when an upcoming contribution is not fully consonant with prior coherence options.” She also states when it is used outside of question-answer sequence, *well* locates a speaker as a respondent to one level of discourse. Additionally, reference [13] regards *well* as a turn initiator, often beginning a turn or initiating a new utterance.

#### B. *Well* as a Pre-Closing Device

On the other hand, at more global levels of conversational organization, [14] observe that *well* can also be used as a pre-closing device, offering its recipient a chance to reinstate an earlier or unexpanded topic, or to open another round of talk before conversational closure [1].

#### C. *Well* as a Delay Marker

Sometimes, in conversations we observe situations in which the speaker(s) are not immediately ready to give responses to another utterance, or are to some degree reluctant to tell the truth, or need some more time to organize what they want to say, they are not too sure about how to state what they

want to say. In these situations, a short silence might occur. However, more frequently, speakers may use some linguistic signals to hold the floor and to make the seemingly incoherent units of utterances into a coherent one; *well* is among such linguistic signals. The use of *well* in such context, is said to function as a delay marker [9].

#### D. *Well* as a Repair Marker

There are instances when speakers need to repair, correct or edit their utterance in the process of having a conversation. In these situations, the speaker usually gives the addressee(s) some hints that he/she is going to repair his/her own words. These hints reflect the speakers’ endeavor to make the seemingly semantic separated units of talk into coherent ones. *Well* is one of these hints. *Well* of this kind functions as a repair marker [9].

#### E. *Well* as a Frame Marker

In some other cases the speaker may change or shift the present topic of talk to another one for various reasons. In order to keep the coherence, he/she may use some signals to indicate that he/she is to change the current topic, and such signals can bring the two separated topics into a more coherent one to avoid an abrupt topic change. *Well* can be such a signal, and function as a frame marker [9].

#### F. *Well* as a Mitigation Marker

Sometimes *well* is used as a linguistic indicator of politeness (to save face), and thus can be regarded as a mitigation marker, which generally occurs at the beginning of a sentence. The use of *well* in this situation normally indicates denials, refusals, and objections to a given utterance [9].

### IV. DATA COLLECTION AND ANNOTATION

Videotaped short dialogues were used as the basis for this study. These dialogues were recorded as part of a workshop held in Digital Humanities Summer School in Leipzig, 2012. The focus of the workshop was on the multimodal nature of human-human communication.

There were six recordings available altogether each containing around seven minutes of informal dialogue between two participants. English was used as the communication language during the conversations. The participants were from different countries (China, Germany, Greece, the U.S, Hungary and Thailand) hence the group was both multicultural and bi/multilingual. This unique feature of our study group enabled us to observe the use of *well* in a wider context, by both native and non-native speakers of English varying in degree of fluency.

The recordings were subsequently transcribed and annotated using ELAN annotation tool [15] on different tiers, each demonstrating a specific non-verbal behavior of the speaker such as posture shift, head shift, gaze movement and eyebrow movement throughout the whole recording. The annotation was carried out based on two modalities (both audio and video). Figure 1 demonstrates the annotation scheme used in this study.

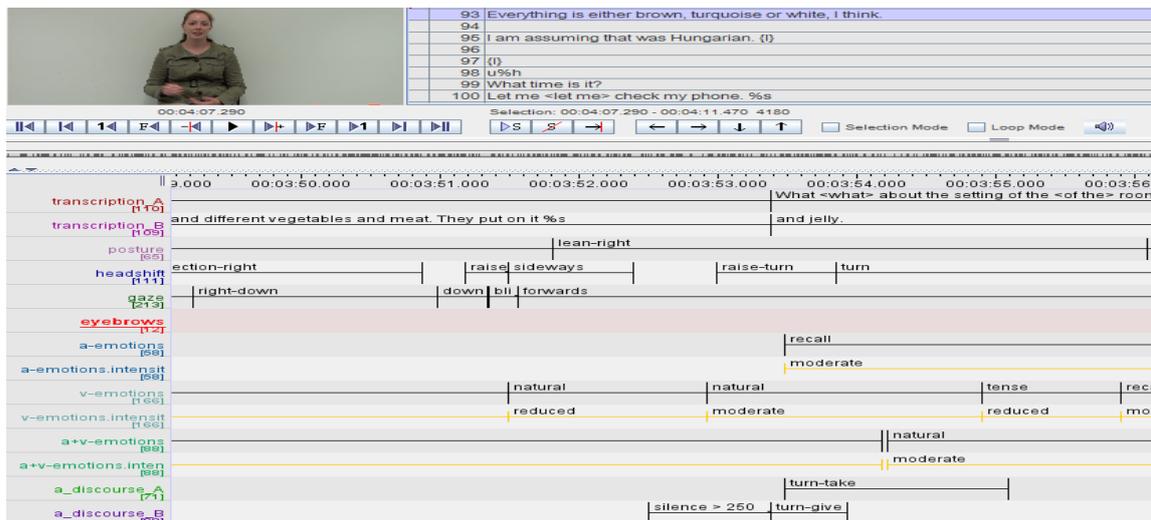


Fig. 1. The annotation scheme

*Well* was annotated on a word segment tier separated from the transcription. In addition, turn boundaries and also the emotional states of the speakers were annotated. The aim was to capture all possible emotional and non-verbal behaviors of the participants (both native and non-native speakers of English) associated with the use of DM *well*. The annotation scheme used in this study was borrowed from the annotation scheme used in HuComTech Corpus of Hungarian spontaneous speech [16]. Table 1 summarizes the structure of our annotation scheme.

TABLE I. THE STRUCTURE OF OUR ANNOTATION SCHEME

Levels of annotation	Annotated elements
Speech transcription	orthographic transcription of speech for both speakers
Word Segment	<i>well</i> was annotated on a separate tier
Emotions	happy, tense, sad, recall, surprise, neutral, other
Gaze movement	Averted: looking away from the speaker Eye-contact: looking directly at the speaker
Eyebrow movement	movement of the speaker's eyebrow using various directional labels
Head shifts	movement of the speaker's head using various directional labels
Discourse labels	turn take, turn give, turn keep and backchannels
Posture shifts	body shifts of the speaker using various directional labels

## V. OBSERVATIONS

Based on our annotation, 12 occurrences of *well* were observed throughout the recordings (42 minutes of conversation altogether) which are listed below:

- 1) A: *Why did you change from psychology to sociology?*  
B: *Because, uh, well, it is a long story.*
- 2) A: *What is the main point?*  
B: *Well, usually there is some regular ways in advertising investment.*

- 3) A: *Where is it?*  
B: *Uh, well, I'm not sure about how much you know about China. Well, it's a beautiful city.*
- 4) A: *For how long was your hometown a capital city?*  
B: *I think probably seven to eight hundred of years, uh, well, not that many, hundreds of years, I'm sure of that.*
- 5) A: *Could you tell me what you did first when you arrived at the hotel?*  
B: *Well, I remember I tried the internet connection.*
- 6) A: *Could you describe how it looks like?*  
B: *Well, it's self serving.*
- 7) A: *What do you find important to tell others?*  
B: *Well, like I said yesterday, I have arthritis.*
- 8) A: *How did he feel?*  
B: *Well, he said he was dumb, stupid and silly and it was a big mistake.*
- 9) A: *How did you find your seat there?*  
B: *Well, I just walked to somebody and sat down.*
- 10) A: *What do you find important about yourself?*  
B: *Well, um, I'm American.*
- 11) A: *Could you say a few words about your hometown?*  
B: *My hometown, well, I grow up in a small town of the east coast.*
- 12) A: *By the way, what technologies are you interested in?*  
B: *Well, I use computers more.*

Table 2 summarizes the 12 occurrences of *well* observed in our data plus their related pragmatic function(s) and also their accompanying non-verbal behavior.

TABLE II. OCCURRENCES OF *WELL*

DM <i>well</i>	Suggested Pragmatic Function	Posture Shift	Gaze Movement	Eyebrow Movement	Head Shift	Turn Management
(1)	Delay marker	None	Averted+Blink	One raised	None	Turn-internal
(2)	Initiation/Response marker	None	Averted+Blink	None	None	Turn-initial
(3)	Delay marker	None	Eye-contact	None	None	Turn-internal
(4)	Repair marker	None	Averted	Both raised	None	Turn-internal
(5)	Initiation/Response marker	None	Averted+blink	Both raised	None	Turn-initial
(6)	Delay/Response marker	None	Averted+blink	None	None	Turn-initial
(7)	Initiation/Response marker	None	Averted	None	None	Turn-initial
(8)	Initiation/Response marker	None	Eye-contact	None	None	Turn-initial
(9)	Initiation/Response marker	None	Averted	Both raised	None	Turn-initial
(10)	Delay/Response marker	None	Averted	None	None	Turn-initial
(11)	Delay marker	None	Averted	None	None	Turn-internal
(12)	Initiation/Response marker	None	Averted	None	None	Turn-initial

The literature enumerates several pragmatic functions for *well*, among which the use of it as an initiation/response marker was the most frequent one observed according to our data. *Well* functioning as a frame or mitigation marker was not observed in our data. This can be due to the short nature of our dialogues, not having enough time to change the topic.

Additionally, it was observed that the use of *well* as a discourse marker was mostly multifunctional. Hence, in most of the cases it was a problem not being able to assign each of the occurrences a single function, for example in one occurrence it functioned as a response marker (replying to a question) and an initiation marker (initiating an utterance or a turn) at the same time. Such cases suggest the following: the literature is only descriptive, enumerating the functions without regard to multifunctionality, without considering the chance that these functions are actually not discriminative enough by nature. This might suggest that a recategorization of these functions is needed. Some functions, like response function of *well*, may in principle be a subset of a broader pragmatic function of initiation. Such an approach could eventually yield an insight into the more complex nature of the functionality of these markers beyond mere description and solve the multifunctionality issue.

Considering the accompanying non-verbal behavior (the multimodal approach to this marker), neither posture shift nor head shift were observed, hence, these two non-verbal signals could be assumed unrelated to the use of *well* as a discourse marker. Gaze and eye-brow movements were mostly correlated with this marker. In almost all cases (except for two) averted eye gaze was present when using *well* regardless of its function. Only in two cases the speaker had direct eye-contact when uttering the marker. In four cases the averted gaze was immediately followed by a blink which might be indicative of the cognitive states of the speaker wanting to concentrate on what he/she wants to say and how he/she should structure their thought. Also, eyebrow movements were

observed accompanying the use of this marker which however were not indicative of a specific function of *well*.

Finally, regarding the role of this discourse marker in turn management, we can say that in most of the cases it appeared at turn-initial position which is not surprising because one of its functions is to be an initiation marker and also a response marker. Although, when it functioned as a repair marker or a delay marker (not in all cases) it appeared at turn-internal position.

However, due to the limitations of our data (the small size of the data and the limited number of occurrences of *well*) we cannot draw definite conclusions but only describe our observations and provide suggestions.

## VI. SUMMARY

Discourse markers being a group of pragmatic elements, are used widely in our daily communication. Reference [1] defines them as “sequentially dependent elements which bracket units of talk”. Among these markers, *well* is found to be the most frequently used one in conversation. According to the literature, this discourse marker displays six main pragmatic functions: being a response or an initiation marker, a pre-closing device, delay marker, repair marker, frame marker, mitigation marker. The aim of our study was to provide a multimodal description of the use of *well* as a discourse marker in spontaneous speech. In order to do so, we collected and annotated 42 minutes of conversation altogether. The conversations were annotated on various levels aiming at capturing different non-verbal behaviors. Based on our limited data, we observed that the use of *well* as a discourse marker is mostly multifunctional so it is not easy to attach one specific function to the use of it. In addition, we observed that the use of this marker was accompanied with non-verbal behaviors. Mostly gaze and eyebrow movements were observed, however, we could not disambiguate the pragmatic function of

*well* based on these non-verbal markers. In almost all cases, having an averted eye gaze was correlated with the use of this marker regardless of its pragmatic function.

Since these markers are frequently used in Human-Human spoken interaction, such multimodal investigations and theoretical results on the multifunctional nature of these devices can be beneficial in dialogue modeling and also in increasing the capability that an intelligent system can utilize for human-infocommunication interface. Due to the small amount of data and also variation in the degree of fluency of the participants in this pilot study, further work is suggested in order to be able to draw more general conclusions.

#### REFERENCES

- [1] D. Schiffrin, *Discourse Markers*, Cambridge: Cambridge University Press, 1987.
- [2] A. Abuczki, "A multimodal analysis of the sequential organization of verbal and nonverbal interaction," *Argumentum*, vol. 7, pp. 261–279, 2011.
- [3] L. Hunyadi, "Multimodal human-computer interaction technologies-theoretical modeling and application in speech processing," *Argumentum*, vol.7, pp. 240–260, 2011.
- [4] T. Wharton, *Pragmatics and Non-verbal Communication*, Cambridge: Cambridge University, 2009, pp. 1-25
- [5] Cassell J, Nakano I.Y, Bickmore W.T, Sidner L. C, Rich C, (2001) Non-verbal cues for discourse structure. In Proc of ACL:114-123. doi:10.3115/1073012.1073028
- [6] P. Furko, *The Pragmatic Marker - Discourse Marker Dichotomy Reconsidered-the Case of Well and Of course*, Debrecen: Kossuth Egyetemi Kiadó, 2007.
- [7] B. Fraser, "An approach to discourse markers," *Journal of Pragmatics*, vol. 14, pp. 383–395, 1990.
- [8] G. Redeker, "Discourse markers as attention cues at discourse transitions," *Approaches to Discourse Particles*, K. Fischer, Ed. Amsterdam: Elsevier, 2006
- [9] M. Li and Y. Xiao, "A comparative study on the use of the discourse marker *well* by Chinese learners of English and native English speakers," *International Journal of English Linguistics*, vol. 2 (5), pp.65-71, 2012.
- [10] L. Hunyadi, "Multimodal human-computer interaction technologies-theoretical modeling and application in speech processing," *Argumentum*, vol.7, pp. 240–260, 2011.
- [11] Y. Ran, "The pragmatic functions of the discourse marker *well*," *Foreign Languages*, vol. 151, pp. 58-64, 2003.
- [12] A. H. Jucker, "The discourse marker *well*: a relevant theoretical account," *Journal of Pragmatics*, vol.19, pp. 435–452, 1993.
- [13] H. Sacks, E.A. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn taking for conversation," *Language*, vol. 50(4), pp. 696–735, 1974.
- [14] E. A. Schegloff and H. Sacks, "Opening up closings," *Semiotica*, vol. 8(4), pp. 289-327, 1973.
- [15] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, H. Sloetjes, "ELAN: A professional framework for multimodality research," In Proc of LREC, Fifth International Conference on Language Resources and Evaluation, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, 2006.  
<http://tla.mpi.nl/tools/tla-tools/elan/>
- [16] K. Pápay, S. Szeghalmy, and I. Szekrényes, "HuComTech multimodal corpus annotation," *Argumentum*, vol. 7, pp. 330-347, 2011.
- [17] A. Kendon, *Gesture: Visible Action as Utterance*, Cambridge: Cambridge University Press, 2004.

