Online Journal of Applied Knowledge Management

A Publication of the International Institute for Applied Knowledge Management

Volume 6, Issue 1, 2018

Is human-human spoken interaction manageable? The emergence of the concept: 'Conversation Intelligence'

Vered Silber-Varod, The Open University of Israel, Israel, vereds@openu.ac.il

Abstract

Currently, via the mediation of audio mining technology and conversational user interfaces, and after years of constant improvements of Automatic Speech Recognition technology, conversation intelligence is an emerging concept, significant to the understanding of human-human communication in its most natural and primitive channel – our voice. This paper introduces the concept of Conversation Intelligence (CI), which is becoming crucial to the study of human-human speech interaction and communication management and is part of the field of speech analytics. CI is demonstrated on two established discourse terms – power relations and convergence. Finally, this paper highlights the importance of visualization for large-scale speech analytics.

Keywords: Conversation Intelligence (CI), speech communication, Interactive Communication Management (ICM), conversational systems, conversation intelligence platform, power relations, and conversation visualization.

Introduction

In an era when the written communication is used in instant messaging applications and social networks, speaking is turning from a simple yet vital skill in human-human interaction to a communication channel that is no more than "nice to have". On the other hand, well-established high-tech companies as well as start-ups are working hard to produce significant information from huge speech-based databases, such as YouTube.com. These trends do not skip the academia, where researchers are trying to crack the codes embedded in the speech signal, with all its complexity, to enable a better human-machine voice interface (For example, The Universal Cognitive User Interface (UCUI) (Huber & Jokisch, 2017). Nowadays goal is not only that individuals will *command* the computer (e.g., *'Find me a job please'*), but that they will *interact* with a conversational system (e.g., *'So, what do you say, should I apply to this position?'*). This system will not only be a robot that identifies the words the user says, but an entity that cooperates with the user, interfaces with her/his daily life, and is involved in daily dilemmas. Conversational systems can range from simple informal, bidirectional text or voice conversations such as an answer to *'What time is it?'* to more complex and lifesaving interactions such as generating a facial composite by collecting oral attestation from an eyewitness (Panetta, 2017), or a system for monitoring neurodegenerative disorders over remote healthcare applications (Noeth, Rudzicz, Christensen, Orozco-Arroyave, & Chinaei, 2016). The forecast is that intelligent conversational systems will be a key focus through 2020 (Panetta, 2017). Although on

the design level conversational system programmers develop human-machines interface in multiple modalities, and not only by voice (e.g., text, sight, sound, tactile, etc.), the present conceptual paper is focusing on the speech communication domain. The next section presents the concept of *conversation intelligence* (CI) and suggests a path to its emergence. In a following section, two natural spoken dialogues phenomena: *power relations* and *convergence*, are demonstrated in the framework of CI. At the end, the article is highlighting the importance of infographics for large-scale speech analytics, which foreshadows how conversational systems would have to identify not only the content of the oral conversation but also the varied nuances of speech.

Conversation Intelligence

Conversation intelligence is the processing of the massive amount of data that accumulate during talk in interaction – including content, vocal features, and even body gestures, and the analysis of this information in an integrative manner to achieve understanding of patterns, structures and insights of this multi-facets type of communication. In her book on *Conversational Intelligence*, anthropologist Glaser (2014) suggested the term *Conversational IQ (C-IQ)*, which deals with the interface between leadership and neurological research. According to Glaser (2014), a discourse which consists of high ratio of negative adjectives and judgmental perspective (for example, *'what X did is terrible'*) causes a particular chemistry in the brain that is responsible for bad feeling of the speaker himself, such as sadness and anxiety. Because these processes occur in interaction between speakers, the conversation is what triggers chemical and neurological processes in the brain. Choosing a certain vocabulary and intonation to express the message, often unconsciously, entails certain neural activity. In other words, the conversation leaves its footprints on the speakers, in the physical sense, and affects their brain. Hence, Glaser (2014) concluded that the practice of a certain rhetoric can lead to mutual trust and, consequently, to personal growth and the prosperity of the organization. Glaser (2014)'s research demonstrates how discourse analysis has penetrated the field of organizational culture, management, and leadership. For example, in the work of Jameson (2001) and Keyton (2011), it shows how an interdisciplinary point of view – communication studies on the one hand and neurological research on the other – suggested new insights. Similarly, Allwood (2008) claimed that communication management incorporates biological and psychological processes and mechanisms.

The business market identified the potential of C-IQ domain as well and adopted methods from artificial intelligence (AI) and Big Data analytics to process massive data of quantitative parameters of the conversations. This turned to be a CI technology or platform (Orlob, 2017a; 2017b). The aim of CI technology is to improve service and performance of marketing and sales personnel (e.g., Gong.io). For example, in a large-scale audio recording analysis, of hundreds of thousands of calls, Orlob (2017a) found that talking-listening ratios in marketing conversations are a parameter that can predict who are the star salespeople (for him, a good marketing person is someone who manages to close more deals in each call). In his blog, Orlob (2017b) claimed that for the duration of the conversation, the best marketers listen to the customer about 54% of the conversation, compared with 45% of the time they speak for themselves. Medium-sized

marketers listen to 32% of the conversation, while the poorer marketers listen 28% of the conversation. Such CI platforms give powerful conversation analytics to commercial companies about their marketing reps' performance. Other trends along the conversation, beyond the amount of silence/talk ratios, are, for example: A. Customer/rep turn-taking ratios, which refer to the alternations between speakers in a conversation. A *turn* is when a participant speaks, and the next turn is when the other participant 'grab the floor' and start to speak, thus, a conversation consists of one speaker at a time in alternating *turns*.); B. The timing of price mentions (which requires also an Automatic Speech Recognition technology). Such data analysis allows an efficient understanding of the quality of a call without even having to listen to the recording. The emergence of CI domain and the interest of the commercial market in its applications can be traced by the evolution of speech technologies according to Gartner (2016). Hype Cycles over the years. Figure 1 summarizes three technologies that are relevant to the issue of CI and that were mentioned over the years: Speech Recognition (Mobile, Desktop, or Cloud); Audio Mining/Speech Analytics; and Conversational User Interface that is not exclusively referring to speech.

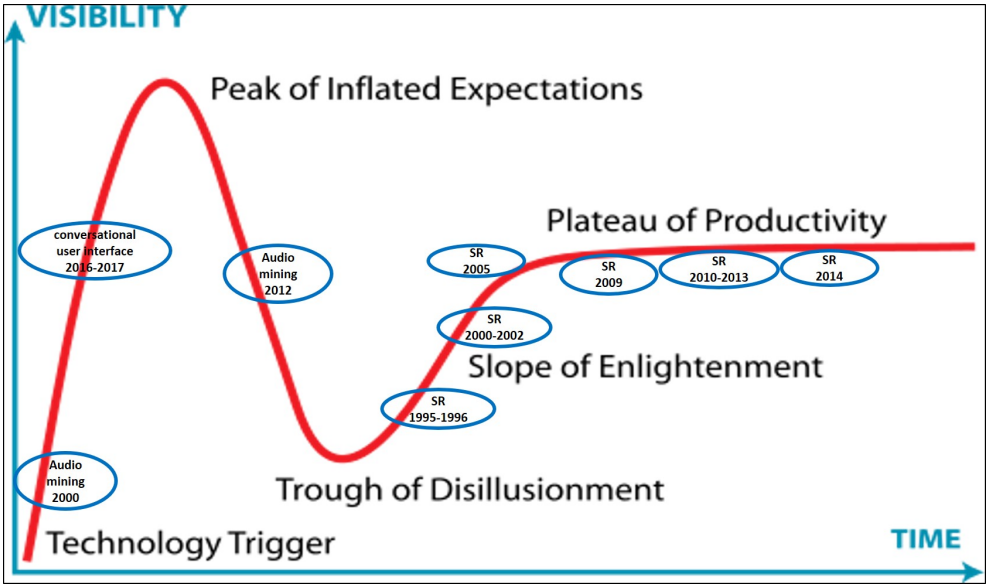

Figure 1. The evolution of conversational intelligence using the Gartner's Hype Cycle graph (Gartner, 2016): Speech recognition (SR) phases over the years (1995-2014) and the emergence of conversational user interface (2016-2017), via the mediation of audio mining phases (2002 and 2012)

Speech recognition technology, or Automatic Speech Recognition (ASR), is among the very few core technologies that the industry (and academia) has been working on for decades. It is mentioned in the first published Hype Cycle in 1995, climbing the Plateau of Productivity. In reality, ASR was far from mature in 1995, and only by 2014, with deep networks breakthroughs, it has reached human equivalent recognition (Mullany, 2016). In 2012, audio mining/speech analytics and text analytics appeared in the Hype Cycle, already after its Peak-of-Inflated-Expectations, and down to the Trough-of-Disillusionment stage, indicating that call centers, for example, have already began to store recordings of their customers and that the material is

analyzed to provide better understanding of these interactions. In 2016, speech technologies are not mentioned per se in the Hype Cylce, yet the term *conversation* is entering the technological scene, and one of the key technology trends is Conversational User Interface, which is climbing to the phase of Peak-of-Inflated-Expectations (Gartner, 2016). As to 2017, Panetta (2016) closely aligned the emergence of conversational systems with AI and machine learning, all are Gartner's 2017 Top 10 Strategic Technology Trends.

Power Relations

This section discusses the issue of power relations between speakers and how dominancy and subordination of speakers can be studied within the framework of CI. Understanding the diverse capabilities of the human voice is part of a growing research trend that tries to simulate synthetic speech to natural human speech as much as possible. Speech researchers strive to explore the nature of verbal and nonverbal symbols, and the situational behavior in varied social contexts, such as personal and organizational relationships (e.g., Silber-Varod, Lerner, & Jokisch, 2017; Silber-varod & Lerner, 2017), intercultural encounters, political and legal discussions, and the like. The focus on situated interaction is to trace the situations, contexts and human behavior that shape and create the interaction as it is. As speakers, we adjust the way we talk to the situation as well. For example, speakers tend to adjust the way they address a child compared to an adult; speakers adjust the level of formality when talking with friends compared to unknown service providers; speakers also tend to adjust the tone of speech to the hierarchical rank of their interlocutors in the organization, and so on.

In the field of Spoken Language Understanding (SLU) (Tur & De Mori, 2011), one of the challenges system developers are facing is how to model power relations in dialogues or multiparty conversations. Traditional research in the field of power as manifested in spoken language was mainly focused on the charismatic attributes of leaders and rhetoric studies (Boss, 1976). More recent studies explore the vocal parameters that are correlated with charismatic personality as been judged perceptually by subjects (Biadsy, Rosenberg, Carlson, Hirschberg, & Strangert, 2008). While it is plausible that signals within the spoken language could provide insight into the social background of the speakers, it has been challenging to extract useful acoustic signals to prove it. On the other hand, social power relationships had been exhaustively studied on written interactions, such as in on-line networks (Bramsen, Escobar-Molana, Patel, & Alonso, 2011; Guha, Kumar, Raghavan, & Tomkins, 2004), and more specifically in Twitter (Leskovec, Huttenlocher, & Kleinberg, 2010), and Wikipedia (Danescu-Niculescu-Mizil, Gamon, & Dumais, 2011; Danescu-Niculescu-Mizil, Lee, Pang, & Kleinberg, 2012). These contributions, which had been explored for decades via computational linguistics and natural language processing tools, had definitely contributed to CI and to power relations research, however speech is a communication code with unique symbolic characteristics and has its own challenges.

One of the attributes of power relations is its dynamics. During an interaction, interlocutors have to "recalculate the route" and find within milliseconds a new path to achieve their goals. This is achieved by Interactive Communication Management (ICM) mechanism (Allwood, 2001; 2008). ICM "concerns features of communication that support interaction, e.g. mechanisms for

management of turns, feedback, sequencing, rhythm and spatial coordination" (Allwood, 2008, p. 20). Dominancy can be the result of an official role of the speaker. For example, in a television panel, the host of the panel will be more dominant, in many senses, than the interviewee, even if the latter is the head of the state. Such a panel would expose power relations other than when the same leader is in a meeting with his ministers. Studies that attempted to automatically tag the dominant speaker found that the five most influential characteristics of conversation dominance were: the number of times the speaker took turns taking, the number of turns, the number of words spoken, and the number of questions asked (Rienks, & Heylen, 2005). The following dialogue extract (Example 1) can demonstrate how challenging it is to understand the power relations between two speakers. This extract is taken from a dialogue between two participants in a map guiding task, which was recorded in a lab. One of them, the *guide*, had the full information of a map; the other speaker, the *follower*, had only milestones that were not in accordance with the guide's full map information. The speakers could not see each other's map. In this sample text all punctuation symbols were deliberately omitted to give the reader a clue to how speech is conveyed and to emphasize the importance of intonation cues in speech. The full information of this sample dialogue, including hypothetical punctuation and the formal roles of the speakers that were assigned to them in this specific task, is presented in Example 2. To a certain extent, it is possible to rely on the content to understand who the guide is (supposedly dominant) and who is the follower (supposedly the subordinate speaker), since Example 1 provides the speakers' assignment to each utterance. However, it is important to recall that speaker's assignment is not a trivial technology and entails another speech signal processing technology – Speaker Recognition, a type of voice biometrics. On the structural level, assuming we do have the turn-taking information, we can count that both speakers had four turns. In terms of word count, speaker A spoke more words than speaker B (31 and 27, respectively. Note that this dialogue was originally not spoken in English and that this is an ad-hoc translation). Yet, can we tell with a high degree of certainty who is the dominant speaker and who is subordinate? This will remain an open question.

Example 1. An extract from a map guiding task dialogue (originally not spoken in English)

Speaker A: Oh it is really to the north
Speaker B: What do you mean really to the north it is approximately in the middle of the map
Speaker A: OK then
Speaker A: uhm
Speaker B: You see this you got ruins of a monastery in the middle of the map
Speaker A: Yes between ruins of the monas- to the lower … I got the Aviation museum
Speaker B: No there is no uhm
Speaker A: So I stop uhm eastern to the ruins of the monastery and much further to the south right
Speaker B: Uh you stop uhm right yes

Convergence in Dialogues

A meta-discourse on speech and conversation, like the present article, is usually biased toward observing the speaker(s), like the paragraph above that discusses the power relations between two speakers. But the exploration of the listening aspect is no less significant, since

communication relies on both sides – the addresser and the addressee (Gallois, Ogay, & Giles, 2005). Communication accommodation theory (Giles, 2007) has pointed out the significance of the addressee to any communication process and suggests that the addresser adjusts her/his message according to the addressee (this phenomenon is also known as *entrainment*, or *convergence*, between interlocutors). The phenomenon of convergence has been investigated in varied contexts, among which the context of intergroup relations that highlighted the importance of communication management to interpersonal as well as intergroup relationships (Gallois et al., 2005); and in the context of lifespan issues by Giles and Harwood (1997), who connected between the communication management model and successful aging. On the foundation of these models, and other related theories, is the understanding that accommodative processes and interactions are embedded within wider social forces. As much as the concept of discourse management seems well established, as part of the study on communication and relationship-maintenance strategies (Giles et al., 1991), it has recently begun to be studied within wider perspectives, such as neuroscience research, on the one hand, and marketing and sales industry, on the other. In Stephens, Silbert, and Hasson (2010) neuroscience research, they found the same phenomenon –speaker-listener neural coupling underlies successful communication. Stephens et al. (2010) showed that when the audience begins to hear a story, a lecture, or a speech, their brainwaves are in a different state, but a minute or two into listening, the brainwaves of all listeners are already responding in the same way. In brain wave research, this phenomenon is called *neural entrainment*. In speech and language research the phenomenon is known as *lexical convergence* and is expressed in varied ways, such as the words used by the speakers. The process occurs for example when a speaker produces a specific word or expression and her/his other interlocutors adopt the same new expression, or word coinage, into their own use. In the study of phonetics, the meaning of convergence is that speakers adopt each other vocal and nonverbal modalities, and it is called *phonetic convergence*. The studies on the phenomenon of convergence emphasize the indispensability of the synchronization between the participants in the conversation to create significant communication between them. The phenomenon is equally important for communication between speakers of different languages, coming from different cultures (Honey, Thomson, Lerner, & Hasson, 2012). Communication in this case is made possible by the linguistic code that transfers meanings, and not necessarily by the vocabulary or other surface representations. Speaker-listener relations were revealed also in a study that scanned speakers' brain in MRI and then compared them to the brain waves of listeners of the same speakers. Interestingly, the researchers found the same activity patterns in speakers and listeners' brains (Hasson, 2016). This finding suggests that speech and perception are similar processes. Moreover, Hasson (2016) found that the more similar the brain activity is between a speaker and its listener is, the better the understanding is. Again, it was found that convergence between conversation participants is vital for better communication.

Analysis by Visualization: Conversation Infographics

The preceding sections presented an overview of discourse phenomena and human-human interaction studies that contributed to the emergence of CI. However, how can large scale speech research be processed into a meaningful analysis? Even recent studies that use larger-scale corpora on language coordination between interlocutors, for example, the on-line domain

(Danescu-Niculescu-Mizil et al., 2011) or online collaborative communities, such as Wikipedia (Laniado, Tasso, Volkovich, & Kaltenbrunner, 2011), use manageable data, in the sense of standard text file formats. One of the ways to extract diverse data in speech recordings and to achieve an understanding of the structure of the dialogue is to produce a visualization of the dialogue in a manner that will enable the voices to be seen. The concept of CI includes this feature. CI platforms provide powerful conversation analytics by visualizing specific trends along the conversation (for example, in Orlob (2017a; 2017b)). CI with the mediation of Conversation infographics has the potential to represent efficiently power relations and convergence, and to save decision makers the time-consuming burden of listening to the recording.

Figure 2 demonstrates the potential of such visualization and a visualization that represents a comparison between four therapeutic sessions. In each session there is a different patient and a different therapist. Each of the four graphs that combine Figure 2 consists of information about the duration each speaker talked. The x-axis represents two-minute intervals, while the y-axis represents the speech duration (in minutes) of each speaker (two speakers in each dialogue). Sessions MS (top left), SG (top right), and to a lesser extent GB (bottom left) demonstrate a similar pattern in which until the 10th minute the patient (red line) speaks most of the time while the therapist utters short feedbacks.

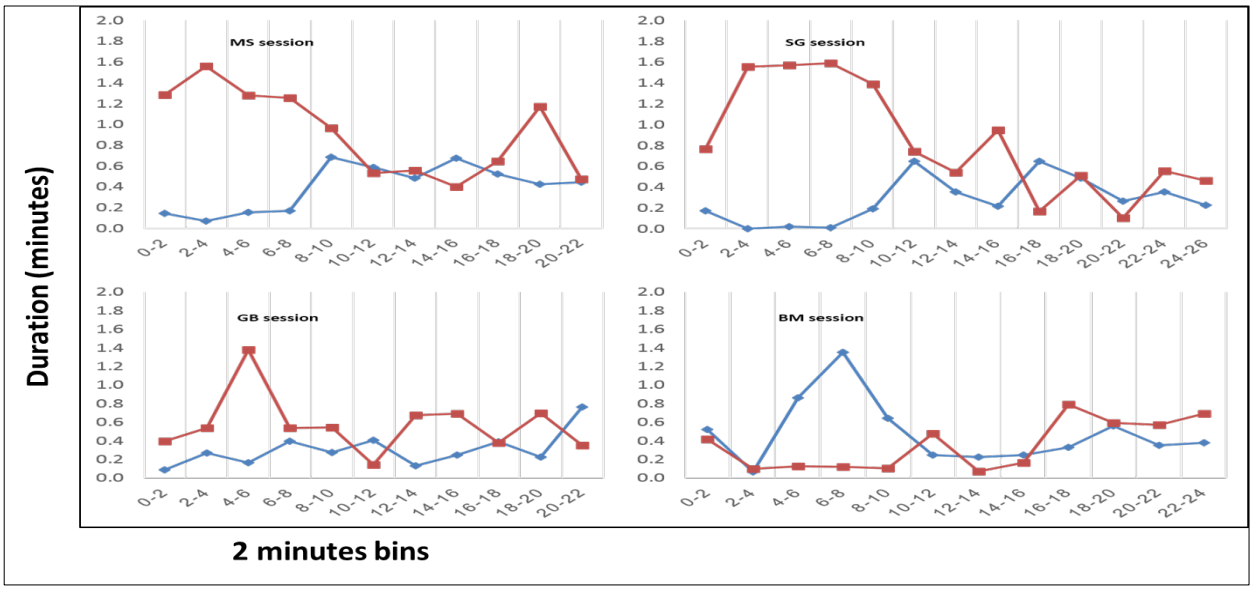

Figure 2. Durational characteristics of four therapeutic sessions between a client (red line) and a therapist (blue line) in intervals of two minutes bin.

The 10-12 minutes interval seems to represent a changing point in all of the four sessions, where the amount of patient's speech becomes closer to the amount of therapist's speech. In session BM (bottom right figure), the first part is characterized by a limited participation of the patient, while the therapist is relatively more active. The figures visualize the difference between session BM and the other three sessions and can be used for instructional purposes and for more inspections of the reasons for this anomaly. Moreover, the comparison of the three sessions to session BM

demonstrates that the role of the speaker does not necessarily entail the participation level, and therefore power relations cannot directly be derived from these durational parameters. In Lerner, Silber-Varod, Batista, and Moniz (2016), we examined the same sessions in a more complex way, using computational learning methods on acoustic parameters. We found that at the beginning of each session there were more differences in the acoustic variables between the two speakers than there were in the middle of the session while the least differences were found at the final part of the session. We concluded that during the dialogue a process of acoustic convergence occurred.

Previously, an extract of a dialogue was presented (Example 1), without critical information of the dialogue. Example 2 presents the same extract with more information, including the role of each speaker in the dialogue, the punctuation marks that to some extent represent the intonation in which the utterances were produced, and long silence intervals. While reading Example 2, readers can reflect on the above-mentioned challenges that have to be overcome by CI platforms. For example, the use of "I" by the follower compared to "you" by the guide; The amount of hesitations (*uhm* and *uh*) of each speaker; who utters more questions (Answer- the guide!), and so forth.

Example 2. An extract from a map guiding task dialogue with CI features such as: time consumption of spoken events (duration in seconds), speakers' role assignment; content (translation of text originally spoken in Hebrew), and intonation (as reflected in punctuation symbols).

Duration (seconds)	**Role**	**Translation and punctuations**
1.075	Follower:	Oh it is really to the north.
1.625	Guide:	What do you mean really to the north it is approximately in the middle of the map!
1.325		<Long pause>
0.825	Follower:	okay then,
1.45		<Long pause>
0.475	Follower:	uhm…
2.825	Guide:	You see this? you got ruins of a monastery in the middle of the map?
4.45	Follower:	Yes between ruins of the monas- to the lower … I got the Aviation museum.
0.625		<Short pause>
1.175	Guide:	No there is no uhm…
6.275	Follower:	So I stop uhm… eastern to the ruins of the monastery and much further to the south. Right?
2.225	Guide:	Uh you stop uhm… right. Yes.

Example 2 is minimal in its information regarding the vocal manipulations of the two speakers. It shows the duration of each line (whether a stretch of speech or a pause indication), the translation of the utterances and punctuation symbols that abstractly reflect the intonation. It is argued that even this information, when aggregated for millions of calls, can be useful for a company's CI analysis. For example, by measuring both objections of prospects (via text analysis) and the average length star-sellers pause after getting an objection, compared to their

peers, Gong. io set a "Patience Score" as a metric formula (Orlob, 2018). These findings and others are extracted from a spoken database that is managed similar to the above simple example: It aligns the utterances per each speaker, it recognizes the objection point from the automatic transcription and most important – it compares a single occurrence of "objection" interaction to thousands and millions of interactions and classifies them according to the salespersons proficiency. Figure 3 demonstrates the potential of such a CI analysis. The figure illustrates the distribution (in %) of four increments (Tags) in a single map guiding task dialogue: Guide, Follower, Pauses, and synchronized speech (Overlaps). The x-axis reflects the time-course of the dialogue (in 2-minutes bins). The peaks and valleys of each tag should be aligned with the transcription in order to model pivot points and junctures in the interaction (similar to objections in Orlob (2018)). For example, it can be hypothesized that during minutes 8-10 of the dialogue, the two speakers had to overcome a challenge, or to discuss a strategy, that will allow them to continue to accomplish the task.

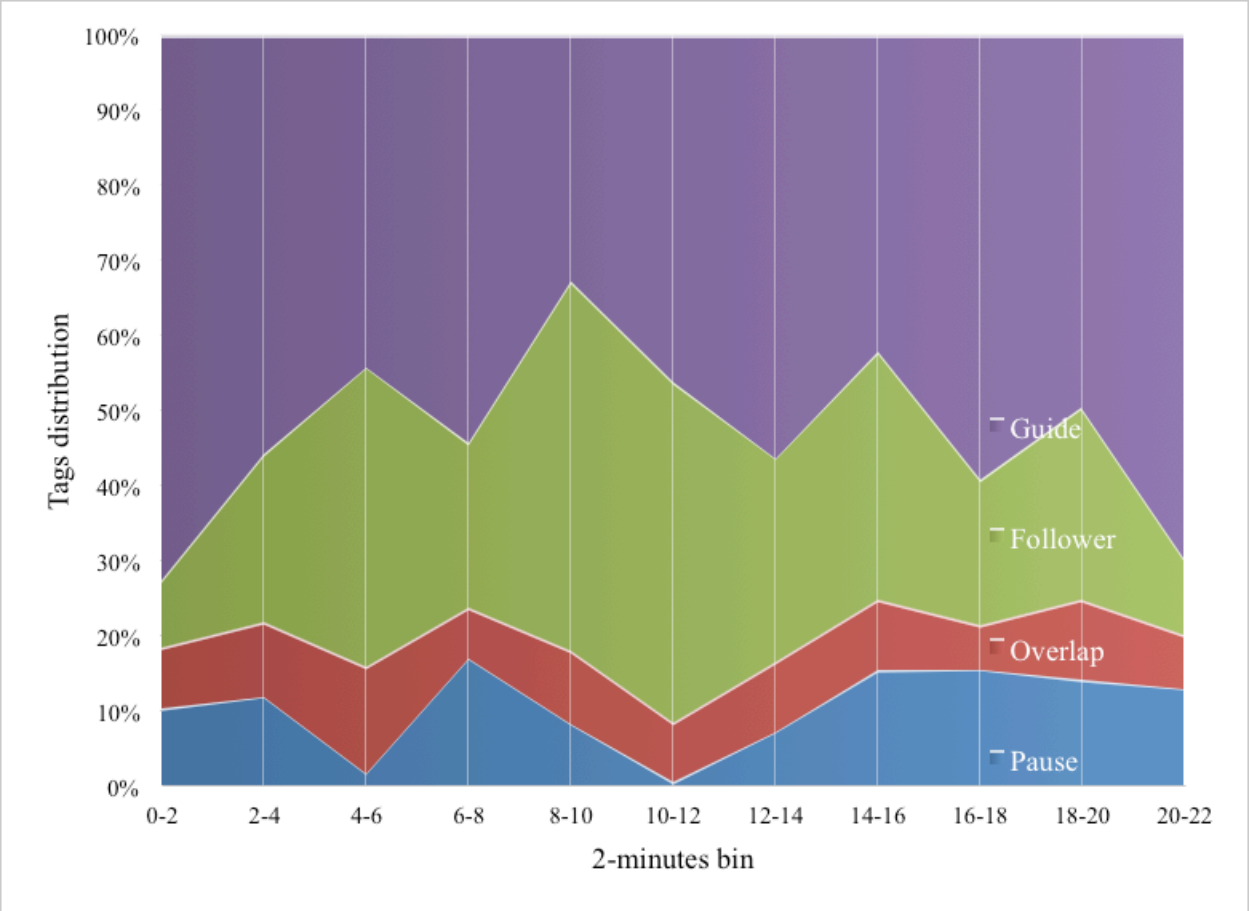

Figure 3. The distribution of four increments (Tags) in a single map task dialogue: guide, follower, pauses, and synchronized speech (Overlaps).

Discussion

This paper presented an emerging term, CI, which refers to the processing of massive amount of conversations. CI research is strongly related to spoken dialogue systems like Siri (Apple.com) or Cortana (Microsoft.com), and other personal assistant AI systems, which can understand natural language and carry out conversations with people. However, the emphasis of this paper is to show how CI is crucial to human-human conversation as well, and why managing through CI systems can contribute to knowledge management of business organizations, for example, by analyzing marketing calls. Moreover, similar to Music Search Engines (MSEs) (Inskip, Macfarlane, & Rafferty, 2012), CI can further contribute to knowledge management of conversational databases by adding attributes that are beyond the traditional classification paradigm, similar to recent Media Assets Management (MAM) design. For example, Institut National de l'Audiovisuel (INA) (http://www.ina.fr/) harvests audiovisual web content by law and it is an example of a MAM system that processes also the signal information of audio-visual collections and not only meta-data. The MAM system of INA includes many KM aspects: legal framework that allows it to collect web content, techniques of audio-visual web harvest, organization of the collections and making them available for research purposes. Microsoft Video indexer (http://vi.microsoft.com/) is another example for media AI technologies that extract CI insights from videos, such as: face recognition, speech/pause ratios, assigning the automatic transcription to a speaker (speaker recognition), speaker appearance ratios in the video, speech sentiments ratios: Neutral, negative, and positive (according to the textual input), and more.

The examples shown above demonstrate the features that are essential to the state-of-the-art dialogue management techniques (Traum, 2017). Apart from module of dialogue modelling – formal characterization of dialogue, evolving context and statistical algorithm to process decisions of how to contribute to dialogues, the system should also manage different technologies, such as: ASR, domain reasoner, syntactic parser, semantic decoder, and speech generator (i.e., Text-to-Speech (TTS)). Moreover, managing such a system means to be sensitive to the multiple roles that were formulated by ICM (Allwood, 2001) and are incorporated into AI system: firstly, information provider, service provider, tutor, and instruction-giver, but also collaborative partner, conversational partner, and advisor. In some cases, developers also wish to provide dialogue systems even with competitor and antagonist features (Traum, 2017).

Summary

Nowadays, via the mediation of audio mining technology and Conversational User Interfaces, and after years of constant improvements of ASR, CI is an emerging concept, and is becoming significant to the understanding of human-human communication in its most natural and primitive channel – our voice. CI platforms allow us to automatically and immediately understand the main intersections in dialogues and monologues as opposed to prolonged listening to the entire recording and subjective judgment of the nature of the interaction. The paper demonstrated the concept of CI on the analysis of power relations and convergence between speakers. Dominancy is not always what appears on the surface and is realized in a complex and sometimes even unexpected manner. Using CI tools, many human-human and

human-computer conversations can be compared, not only in terms of their content, but in terms of their behavior patterns and strategies in problem solving (Huber & Jokisch, 2017), as reflected in the voices of the speakers. Moreover, CI can further contribute to knowledge management of conversational databases by adding attributes that are beyond the traditional classification paradigm, given the vision of the CI concept to analyze structures and relations between speakers. Last, visualization of dialogues and conversations by infographics of large scale conversational data also solve ethical issues, where the information about the speakers, their voices, and sometimes even the content of their talk, is kept hidden, while other non-disclosure parameters are measured and exposed**.**

To conclude, this paper introduced the emerging concept of CI, which is part of the field of speech analytics, and is becoming crucial for manageable human-human speech interaction and communication management. The CI concept was demonstrated on natural spoken dialogues analysis and highlighted the importance of infographics for large-scale speech analytics. Moreover, the paper showed that CI is an interdisciplinary field, involving linguistics, engineering, neural sciences, and it is also motivated by business management and marketing interests.

References

Allwood, J. (2008). Dimensions of embodied Communication - towards a typology of embodied communication. In I. Wachsmuth, M. Lenzen, & G. Knoblich (Eds.), *Embodied communication in humans and machines* (pp. 1-24). Oxford, UK: Oxford University Press.

Allwood, J. (2001). The structure of dialog. In M. M. Taylor, D. G. Bouwhuis, & F. Néel (Eds.), *The structure of multimodal dialogue II* (pp. 3-24). Amsterdam, Netherlands: John Benjamins.

Bramsen, P., Escobar-Molano, M., Patel, A., & Alonso, R. (2011). Extracting social power relationships from natural language. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 773-782). Portland, Oregon: Association for Computational Linguistics.

Biadsy, F., Rosenberg, A., Carlson, R., Hirschberg, J., & Strangert, E. (2008). A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech. *Proceedings of the 4th conference on Speech Prosody*, Campinas, Brazil.

Boss, G. P. (1976). Essential attributes of the concept of charisma. *Southern Journal of Communication*, *41*(3), 300-313.

Danescu-Niculescu-Mizil, C., Gamon, M., & Dumais, S. (2011). Mark my words! Linguistic style accommodation in social media. *Proceedings of the 20th international conference on World Wide Web* (pp. 745-754). Hyderabad, India: Association for Computing Machinery.

Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of power: Language effects and power differences in social interaction. *Proceedings of the 21st*

international conference on World Wide Web (pp. 699-708). Hyderabad, India: Association for Computing Machinery.

Gartner, Inc. (2016). Gartner's 2016 hype cycle for emerging technologies identifies three key trends that organizations must track to gain competitive advantage. Press Release, Stamford, Connecticut, August 16, 2016. Available at: https://www.gartner.com/newsroom/id/3412017

Gallois, C., Ogay, T., & Giles, H. (2005). Communication Accommodation Theory: A look back and a look ahead. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 121-148). Thousand Oaks, CA: Sage.

Giles, H. (2007). *Communication accommodation theory*. In B. B. Whaley & W. Samter (Eds.), *Explaining communication: Contemporary theories and exemplars* (pp. 293-310). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Giles, H., & Harwood, J. (1997). Managing intergroup communication: Life span issues and consequences. *Trends in Linguistics Studies and Monographs*, *100*, 105-130.

Glaser, J. E. (2014). *Conversational intelligence: How great leaders build trust and get extraordinary results*. Brookline, MA: Bibliomotion, Inc.

Guha, R., Kumar, R., Raghavan, P., & Tomkins, A. (2004). Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web* (pp. 403-412). Hyderabad, India: Association for Computing Machinery.

Hasson, U. (2016). *This is your brain on communication*. TedTalk. Published on Jun 3, 2016. Available at: https://youtu.be/FDhlOovaGrI

Honey, C. J., Thomson, C. R., Lerner, Y., & Hasson, U. (2012) Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience*, *32*(44), 15277-15283.

Huber, M., & Jokisch, O. (2017). Cognitive data retrieval using a Wizard-of-Oz framework. *Proceedings of the 5th Knowledge Management Conference* (pp. 31-40), Novo Mesto, Slovenia: International Institute of Applied Knowledge Management.

Inskip, C., Macfarlane, A., & Rafferty, P. (2012). Towards the disintermediation of creative music search: Analysing queries to determine important facets. *International Journal on Digital Libraries*, *12*(2-3), 137-147.

Jameson, D. A. (2001). Narrative discourse and management action. *The Journal of Business Communication*, *38*(4), 476-511.

Keyton, J. (2011). *Communication and organizational culture: A key to understanding work experiences*. Thousand Oaks, CA: Sage Publications.

Laniado, D., Tasso, R., Volkovich, Y., & Kaltenbrunner, A. (2011). When the wikipedians talk: Network and tree structure of Wikipedia discussion pages. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (pp. 177-184), Menlo Park, California: AAAI Press.

Lerner, A., Silber-Varod, V., Batista, F., & Moniz, H. (2016). In search of the role's footprints in client-therapist dialogues. *Proceedings of Speech Prosody 2016* (pp. 400-404). Boston, MA, USA: International Speech Communication Association.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Predicting positive and negative links in online social networks. *Proceedings of the 19th international conference on World wide web* (pp. 641-650). Hyderabad, India: Association for Computing Machinery.

Mullany, M. (2016). 8 lessons from 20 years of hype cycles. LinkedIn blog. Published on December 7, 2016. Icon Ventures. Available at: https://www.linkedin.com/pulse/8-lessons-from-20-years-hype-cycles-michael-mullany

Noeth, E., Rudzicz, F., Christensen, H., Orozco-Arroyave, J. R., & Chinaei, H. (2016). *Remote monitoring of neurodegeneration through speech.* Third Frederick Jelinek Memorial Summer Workshop. Available at: https://www.clsp.jhu.edu/workshops/16-workshop/remote-monitoring-of-neurodegeneration-through-speech/

Orlob, C. (2018). *This is how successful salespeople handle objections, according to new data.* Published on March 12, 2018. Available at: https://www.gong.io/blog/handling-sales-objections/

Orlob, C. (2017a). [Infographic] *The science of winning sales conversations.* Published on April 10, 2017. Available at: https://www.gong.io/blog/winning-sales-conversations/

Orlob, C. (2017b). *This is what separates your star reps from the rest of the team.* Published on November 4, 2017. Available at: https://www.gong.io/blog/this-is-what-separates-your-star-reps-from-the-rest-of-the-team/

Panetta, K. (2017). *Top trends in the Gartner hype cycle for emerging technologies, 2017*. Published in August 15, 2017. Available at: https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/

Panetta, K. (2016). *Gartner's top 10 strategic technology trends for 2017*. Published in October 18, 2016. Available at: https://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017/

Rienks, R., & Heylen, D. (2005). Dominance detection in meetings using easily obtainable features. *Proceedings of the International Workshop on Machine Learning for Multimodal Interaction* (pp. 76-86). Berlin, Heidelberg: Springer.

Silber-Varod, V., & Lerner, A. (2017). Analysis of silences in unbalanced dialogues: The effect of genre and role. In R. Eklund & R. Rose (eds.). *Proceedings of the 8th Workshop on Disfluency in Spontaneous Speech*, (pp. 53–56). Stockholm, Sweden: Royal Institute of Technology.

Silber-Varod, V., Lerner, A., Jokisch, O. (2017). Automatic Speaker's Role Classification With a Bottom-up Acoustic Feature Selection. *Proceedings of the 2017 International Workshop on Grounding Language Understanding*, (pp. 52-56). Stockholm, Sweden: Royal Institute of Technology.

.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010) Speaker-listener neural coupling underlies successful communication. *Proceeding National Academy of Science, USA*, *107*(32) 14425-14430.

Traum, D. (2017). Computational approaches to dialogue. In E. Weigand (Ed.), *The routledge handbook of language and dialogue*, New York, NY: Routledge.

Tur, G., & De Mori, R. (2011). *Spoken language understanding: Systems for extracting semantic information from speech*. Hoboken, NJ: John Wiley & Sons.

Author's Biography

Vered Silber-Varod is currently a Research Fellow at the Open Media and Information Lab (OmiLab), The Open University of Israel. Formerly, being a Research Fellow at The Research Center for the Study of Innovation in Learning Technologies, The Open University of Israel. She holds a Ph.D. in Humanities from Tel Aviv University. Her research interests and publications focus on various aspects of linguistics, with expertise in speech prosody, acoustic phonetics, and speech and text analytics. Dr. Silber-Varod is a member of the International Speech Communication Association (ISCA). Personal site: https://www.openu.ac.il/en/personalsites/VeredSilberVarod.aspx