Entrainment, dominance and alliance in supreme court hearings

Štefan Beňuš [a,b,*], Agustín Gravano [c,d], Rivka Levitan [e], Sarah Ita Levitan [e], Laura Willson [e], Julia Hirschberg [e]

[a] Constantine the Philosopher University, Nitra, Slovakia
[b] Institute of Informatics, Slovak Academy of Sciences, Bratislava, Slovakia
[c] Departamento de Computación, FCEyN, Universidad de Buenos Aires, Argentina
[d] National Scientific and Technical Research Council (CONICET), Av Rivadavia 1917, C1033AAJ, Buenos Aires, Argentina
[e] Department of Computer Science, Columbia University, 450 Computer Science Building, 1214 Amsterdam Avenue, Mailcode: 0401, New York, NY 10027-7003, USA

ARTICLE INFO

ABSTRACT

A major goal of the Cognitive Infocommunication approach is to develop applications in which human and artificial cognitive systems are made to work more effectively. A critical step in this process is improving our understanding of human–human interaction so that it may be modeled more closely. Our work addresses this task by examining the role of *entrainment* – the propensity of conversational partners to behave like one another – in (1) the production of *conversational fillers* (*CFs*) and acoustic intensity; (2) patterns of turn-taking; and (3) Linguistic Style.

markers and how all of these relate to power relations, conflict, and voting behavior in a corpus of speech produced by justices and lawyers during oral arguments of the U.S. Supreme Court in the 2001 term. We examine several different measures of entrainment in justice–lawyer pairs to see whether or not they are related to justices' favorable or unfavorable votes for the lawyer's side. While two measures (a naive measure of similarity in CF rates and global similarity in CF phonetic realizations for the entire session) show no relationship, a third, which measures local entrainment in CFs in lawyer-justice pairs, does in fact identify a significant positive relationship between entrainment and justice votes. With respect to local entrainment in intensity, we found that lawyers do entrain more to justices than justices to lawyers, although there is no greater entrainment of female lawyers than of male lawyers. When we examine the relationship between entrainment in intensity and judicial voting, we find that, when justices voted for the petitioners, there is significant evidence of entrainment by both petitioners and respondents to justices. With respect to turn-taking behavior, we find that certain patterns of overlaps in turn exchanges between justices and lawyers are correlated with justices' voting behavior for four of the justices in our corpus. Finally, we find that there are lexical cues to divisiveness within the Court itself that can distinguish cases with close verdicts from cases with unanimous verdicts. We link these results to the possibility of building cognitive info-communication interfaces that exploit features of human–human entrainment for increasing effectiveness of human–machine interactions.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

One of the primary goals of the Cognitive Infocommunication approach [3] is to facilitate the development of "engineering applications in which artificial and/or natural cognitive systems are enabled to work together more effectively". One may approach this task by (1) improving understanding of the cognitive aspects of human–human interactions, (2) building formal models based on this understanding, and (3) implementing these models in human–machine systems to facilitate more natural and efficient interactions. In this paper, we report on a set of case studies aimed at the first step of this process in the area of speech **entrainment**, the tendency of interlocutors to become similar to each other in terms of their acoustic and prosodic production (e.g. [41,6]). We examine how several types of such entrainment between conversational partners in the judicial domain relate to cognitive and social aspects of communication and information transfer.

A better understanding of entrainment is important for a number of applications of human–machine communication that rely upon Spoken Dialogue Systems (SDS). Research has shown that

* Corresponding author at: Constantine the Philosopher University, Štefánikova 67, 94974, Nitra, Slovakia. Tel.: +421 911 275318.
E-mail address: sb513@nyu.edu (Š. Beňuš).

not only do humans perceive conversational partners who entrain to their speaking style as more socially **attractive** and **likeable** [62,12,1], more **competent** (Street 1984), and more **intimate** [19], but interactions with partners who unconsciously mimic them are seen as more **successful** [22,60,35]. Different dimensions of entrainment have been shown to be reliable predictors of **task success** [64,55,49] as well. It has also been shown that humans may consciously decrease their similarity to others in order to increase their social distance to the interlocutor [34] or to show a negative attitude toward the interlocutor [13].

Not only do humans entrain to other humans, but studies have shown that they also entrain to computer systems. Nass and colleagues showed that human subjects perceive systems that entrain to them to be more likeable and interactions with those partners to be more successful [52]. A number of studies have shown that subjects do adapt to machines as well as to human conversational partners [5,25], so the ability to mimic this tendency found in human–human conversation would appear to be important for human–machine conversation as well, if SDS are to be as natural and effective as human partners.

It is well known that in spoken interactions between humans, information flows in multiple channels. Hence, information is construed as including not only the propositional semantic content of utterances but also other aspects of communicative functions such as, for example, Jacobson's referential, aesthetic, emotive, conative, phatic, and metalingual functions [44]. Cognitive Infocommunication approaches the notion of *channel* in a more abstract way as "a combination of sensory substitution and sensorimotor extension to convey structured information via any number of sensory modalities" ([26]: 261). We contribute to the notion of channel integration in future Cognitive infocommunication applications by examining how structured paralinguistic information contained in the acoustic channel of the spoken modality links to coordination and social relations between humans and might enhance the success of their interactions. The larger goal of this research is to replicate such behavior in human–computer interactions.

1.1. Entrainment

Coordination is a basic feature of human interactions. Sometimes, coordination in movements is explicitly intentional, or dictated by rules of social contact (for example in dancing) while other times it can be unintentional and facilitated by affordances of visual or aural modalities. For example, seeing somebody rock in a chair makes another person's rocking unintentionally entrained to this visual rhythmical movement [65]. Hence, visual information can couple with the movements of people involved in interaction and result in (unintentional) coordination.

Support for the coordination view of human–human spoken interaction comes from the literature on entrainment. For example, conversational partners tend to become more similar to each other as they speak. This phenomenon, known in the literature as entrainment, alignment, coordination, adaptation, unconscious mimicry, or 'the Chameleon Effect' (below we will use the term 'entrainment' exclusively), occurs along many acoustic, prosodic, syntactic and lexical dimensions – as well as in social behavior such as turn-taking – in both human–human interactions [15,25,64,77,54,78,18] and human–computer interactions [16,72,5]. Evidence of entrainment has been demonstrated in vowel spectra [1]; fundamental frequency [2]; pronunciation [40,59]; intensity [53]; voice quality [67]; lexical and syntactic choice [14,64]; frequency and duration of pauses [45]; speaking rate [39,5]; response latency ([21]); utterance length [50]; turn-taking behavior [47], jokes and laughter [63]. It has been found in many cultures: Hungarian [46], Frisian and Dutch [36,81],

Hebrew [80], Taiwanese Mandarin [76], Japanese [79], Cantonese [32], and Thai [4].

1.2. Dominance

The notion of dominance is closely connected to entrainment between conversationalists. Thus, the amount of entrainment among interlocutors represents a potential window into the dynamics of the power relationship. The observed differences between speakers in their degree of entrainment to different individuals may indicate an asymmetrical power relationship. More specifically, if one interlocutor adjusts his or her behavior to that of a conversational partner more than the partner does, the former is likely to be perceived as playing a less dominant role than the latter. The understanding of asymmetrical distribution of power, and related aspects of dominance and status and how they are signaled through speech, has great potential for facilitating the quality of interactions between natural and artificial cognitive systems, since they represent a natural component of human–human interactions.

This view of an asymmetrical power relationship dynamically created or maintained through communicative interaction is closely related to the *dyadic power theory* of Dunbar and colleagues [29,20,30]. In this theory, dominance is seen as a combination of personal and contextual characteristics. The personal characteristics are the constant features of an individual which can be considered as personal traits that are independent of the situation with which the individual is faced. The contextual characteristics include the dominance or submissiveness of the individual's partner in the interaction. Here, we follow Poggi and D'Errico [61] and view dominance as a dynamic multidimensional communicative act by means of which one's interlocutor exerts power or influence over one or more conversational partners by displaying linguistic signals of dominance.

In Poggi & D'Errico's view, dominance is dynamic: it evolves over time. Interlocutors may begin an interaction with roughly equal power positions and finish with very different ones. Alternatively, the power relationship may be similar at the beginning and end of the conversation but may diverge in various dimensions during the conversation. In this sense, dominance is constantly being negotiated. Dominance is also multidimensional in that one conversational partner may be more dominant than their interlocutor in one dimension or aspect of the conversation while the roles might be reversed in other dimensions or aspects. Dominance is also relational, and is not assessed in absolute terms; rather the dominance of an interlocutor is only defined in relation to the dominance of their conversational partners.

Here we construe the ability or intention to influence a conversational partner as observable in the use of spoken language during the conversation. We hypothesize, following Giles et al. [33], that the degree of entrainment in speech will be asymmetric, and the less dominant speaker will entrain more to the more dominant speaker than vice versa. Assuming that social status can be principally linked to power and dominance, support for this hypothesis comes from studies of non-verbal behavior by Gregory and Webster [38] which showed that lower status partners entrained their voices to higher status partners.

1.3. Entrainment, dominance, and the judicial domain

Most studies investigating the relationship between entrainment and dominance have analyzed corpora collected in the laboratory or in situations where the 'stakes' were relatively low – that is, neither party was heavily invested in the outcome of the conversation. In such situations, the relationship between entrainment and dominance is hypothesized to form a social glue, indirectly facilitating successful outcomes in the low-stake tasks at hand