# The challenge of dialogue for speech and language processing

Simon Garrod & Anne Anderson
University of Glasgow

Human language processing and the nature of the language produced is dramatically affected by whether it is in the context of dialogue or monologue. Pickering and Garrod (in press, Garrod & Pickering, 2004) argue that during dialogue interlocutors can take advantage of a special interactive mechanism which supports the automatic alignment of linguistic representations at many levels. This mechanism is not available when producing or comprehending monologue. One of the key questions is what distinguishes monologue from dialogue with respect to the interactive alignment mechanism. The answer may seem obvious but studies of group communication suggest otherwise. For example, Fay, Garrod & Carletta (2000) present evidence that communication in large as opposed to small groups is closer to serial monologue than dialogue. Whereas speakers in small groups align most with those who speak before or after them, in large groups everyone only aligns with the dominant speaker. In other words the language processing in large groups is not governed by principles of interactive alignment. This project aims to define more precisely the distinction between monologue and dialogue language processing and its consequences on the language and speech produced.

This is both a scientifically interesting issue in itself and an issue of some practical significance for speech and language engineering. The quality of the speech and the nature of the language is quite different in dialogue from monologue as a consequence of interactive alignment. For instance, interactive alignment promotes the production and supports the interpretation of elliptical and phonetically attenuated speech. Hence dialogue presents a challenge for speech processing systems. However, interactive alignment also promotes the development of 'dialogue routines' or semi-fixed expressions during a conversation and automatic speech processing systems could take advantage of this special redundancy in the language/speech.

The project will also explore the role of verbal and visual feedback in alignment and in dialogue and small group interactions more generally. In dialogues due to the other cognitive demands on speakers we know that visual cues are used sparingly, yet they have considerable impacts on dialogue structure and speech quality (Anderson et al, 1997). We will investigate speakers' abilities to monitor their listeners, as the size of the group grows. The project will use a variety of research techniques including eye-tracking of spontaneous dialogues currently being deployed in the HCRC labs. This research addresses basic scientific question about multimodal interactions and has considerable implications for a range of speech technologies and the design of computational systems designed to support meetings. It builds on previous research we have conducted on Computer Supported Collaborative Work, notably on the support of 'virtual teams' (Anderson et al, in press).

The project would use both psychological and computational approaches to address the monologue-dialogue distinction and the role of verbal and visual feedback in these processes. It would also investigate language processing in both group discussion

contexts and more traditional two party dialogue contexts. Example sub-projects might include (1) Investigating aligned attention during dialogue by eye-tracking conversational partners in 'visual world' situations (see e.g., Brown-Schmidt, Campana, & Tanenhaus, in press), (2) Investigating coordinated gaze patterns and associated speech patterns in small and large groups, using the Edinburgh HCRC 'electronic meeting room', (3) Investigating processing differences between routinized and non-routinized language using the HCRC map task corpus (e.g., by examining the different rates of speech in the two cases and differences in intelligibility).
(4)Explorations of the uptake and impact of verbal and visual feedback in dialogue and small group settings on dialogue structure and speech quality.

Refs:

Anderson, A.H., Bard, E.G., Sotillo, C., Newlands, A. & Doherty-Sneddon, G. (1997) Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception and Psychophysics,* 59(4), 580-592.

Anderson, A.H., Mullin, J., MCEwan, R., Bal, J., Carletta, J., Grattan, E. & Brundell, P. (in press, 2004) Exploring why virtual team working is effective in the lab but difficult in the workplace. In H.Bromme, F. Hesse & H. Spada (Eds) *Barriers and biases in computer-mediated knowledge communication- and how they may be overcome*. Kluwer.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (in press). Real-time reference resolution in a referential communication task. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Processing world-situated language: Bridging the language as product and language as action traditions.* Cambridge, MA: MIT Press.

Fay, N., Garrod, S., & Carletta, J. (2000). Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological Science, 11*(6).

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*, 8-11.

Pickering, M. J., & Garrod, S. (in press). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*.