

CHAPTER 26

Alignment in dialogue

Simon Garrod and Martin J. Pickering

26.1 Introduction

Almost without exception the other chapters in this handbook address the study of monologue, involving isolated language production or comprehension. Such work implicitly or explicitly adopts a perspective in which researchers attempt to understand the processes that occur when the speaker converts meaning into sound and the listener converts sound into meaning. But is such a perspective sufficient to explain dialogue?

In this chapter, we argue that psycholinguists need to think in a different way to understand processing in dialogue. According to this view, interlocutors do not use language to encode and decode messages, but rather as a means by which they can align their mental states, so that they come to have the same ideas about the topic under discussion (Pickering and Garrod, 2004). It would in principle be possible to decompose this task into discrete acts of comprehension and production, but in practice real dialogue involves constant overlaying of production and comprehension (as when the addressee provides feedback such as *yes*, *mm*, or a gesture during the speaker's contribution). It is much better to understand dialogue as a joint activity, like ballroom dancing or using a two-handed saw (Clark, 1996), and to assume that alignment follows from this inherently interactive process.

We simply assume that interlocutors seek to align their mental states, just as we assume that isolated speakers and listeners seek to encode and decode messages. This may sound overly optimistic, as people clearly do not always seek to agree with each other. But even in the most *irreconcilable* argument, interlocutors align on the topic that is being discussed, the referents of expressions, and so on. And of course people do

misunderstand each other, but isolated speakers also sometimes make speech errors, and isolated listeners sometimes fail to understand what they hear. Full alignment (in which interlocutors have identical mental states) may never occur, but interlocutors attempt to align just as isolated speakers and listeners attempt to encode and decode. So, on the assumption that the goal of dialogue is alignment, this chapter discusses ways in which interlocutors come to achieve this state.

Just like other complex cognitive processes, alignment in dialogue involves both automatic and strategic components. However, current theories of dialogue emphasize them to different extents. Clark (1996) argued that interlocutors use various strategies to accumulate what he terms "common ground," which refers to all the information that both interlocutors believe to be shared by themselves and their conversational partner. This is actually a stricter notion than alignment, which merely refers to the information that happens to be shared. Brown and Dell (1987) made the point that information accessible to the speaker will often tend to be accessible to the listener at that point as well. This notion of shared accessibility underlies Pickering and Garrod's (2004) implicit common ground (i.e. information that is automatically made accessible to both interlocutors as a result of the interaction). To the extent that interlocutors share common ground or implicit common ground, their mental states are aligned.

In contrast to Clark's (1996) account, other accounts assume that alignment is principally the result of automatic mechanisms such as priming between interlocutors (Pickering and Garrod, 2004) or associative processes underlying memory access (Horton and Gerrig, 2005a; 2005b). In this chapter, we assess various ways in

which interlocutors may become aligned, and consider whether the underlying mechanisms are largely strategic or automatic. In turn, we discuss alignment via beliefs about one's interlocutor, imitation, agreement between interlocutors, feedback, and physical co-presence.

26.2 Sources of alignment

26.2.1 Alignment via beliefs about one's interlocutor

Clark (1996) argued that common ground is partly determined on the basis of what the speakers already know or can reasonably infer about each other before the conversation begins (based on inferences about community membership or shared knowledge from previous conversations). For example, Isaacs and Clark (1987) had speakers describe pictures of New York landmarks either to other New Yorkers or to "out-of-towners," and found that they quickly established whether their listener was a New Yorker and modified the descriptions accordingly (e.g. referring to buildings by name for New Yorkers or in relation to their appearance for "out-of-towners"). Speakers similarly adjust their speech to adults versus children (e.g. Shatz and Gelman, 1973) and native vs. non-native speakers (Bortfeld and Brennan, 1997). This suggests that interlocutors explicitly infer the extent of common ground and modify their utterances accordingly.

No doubt speakers make some reasoned decisions about how to produce descriptions on the basis of what they believe their audience to know. But their choices may also reflect automatic resonance processes, of the kind exemplified in models by Logan (1988) and Hintzman (1986). Resonance could arise from the prior association of particular types of addressee with particular expressions, the activation of such types of addressee as a result of the interaction, and the subsequent activation of those expressions (Horton and Gerrig, 2005a). For example, if I know Bob is a psychologist, then I may automatically activate psychological terms when meeting Bob, and therefore circumvent the problem of deciding whether Bob is likely to understand every technical expression I use. Furthermore, Horton and Gerrig point out that information that is accessible to the speaker will also tend to be accessible to the listener because of resonance. When person A interacts with person B, information A associates with B will automatically become more accessible to A and

information B associates with A will automatically become more accessible to B. The net effect is to promote the accessibility of information that they have in common (e.g. from past encounters between A and B).

Horton and Gerrig (2005a) use examples of first mentions of characters during telephone conversations to show that speakers sometimes over- or under-estimate what their interlocutor knows about the character, but argue that interlocutors are often able to repair any misunderstandings easily. In experimental studies, language users do not always take common ground into account in producing or interpreting references. For example, Horton and Keysar (1996) found that speakers under time pressure did not produce descriptions that took advantage of what they knew about the listener's view of the relevant scene. In other words, the descriptions were formulated with respect to the speaker's current knowledge of the scene rather than with respect to common ground. Keysar et al. (1998) found that listeners initially considered objects to be potential referents for expressions even when they knew that the speaker was not attending to those objects, and Keysar et al. (2000) found that listeners initially looked at objects that they knew the speaker was not aware of.

In contrast, Hanna et al. (2003) reported a similar eye-tracking experiment in which they did find evidence that listeners preferred to look at referents that they knew to be shared with the speaker. In their experiment, however, the shared information not only was visually available to speaker and listener but also had already been established as such during prior conversation. For example, at the beginning of the experiment the instruction-follower might have a red triangle that only he could see and then be asked to put another red triangle onto the board that was visible to both participants. Subsequently he would be instructed: "Now put the blue triangle on the red one," referring to the recently mentioned and visually shared red triangle. At this point the instruction-follower was more likely to gaze first at the shared triangle, thus suggesting sensitivity to information in common ground. So addressees may make reference to common ground when the discrepancy between their knowledge and that of the speaker is made especially salient. However, Hanna et al.'s immediate effects of common ground on eye movements could be due to resonance rather than strategic inference (Horton and Gerrig, 2005a).

To represent common ground, the interlocutor needs to maintain a very complex situation

model that reflects both his own knowledge and the knowledge he assumes to be shared with his partner. To do this, he has to keep track of the knowledge state of his partner in a way which is separate from his own knowledge state. This is a very stringent requirement for routine communication, in part because he has to make sure that this model is constantly updated appropriately (Halpern and Moses, 1990). For example, each interlocutor would have to flag every piece of information as being in or out of common ground. Indeed, it is even more complicated when interlocutors need to differentiate common ground, their own personal knowledge, and knowledge they believe to be exclusive to their partner (e.g. in a spying scenario). Nevertheless, under certain circumstances interlocutors do engage in strategic inference relating to common ground, at least under conditions where the discrepancy between interlocutors' knowledge is very salient. Horton and Keysar (1996) found that speakers who were not under time pressure often take account of common ground in formulating their utterances. Keysar et al. (1998, 2000) argued that listeners eventually take account of common ground in comprehension, though they proposed that this occurs at a later monitoring stage, in a process that they called "perspective adjustment."

26.2.2 Alignment via imitation

People use beliefs about their interlocutors to start the process of alignment. However, we would expect the alignment process to continue after both interlocutors have made contributions, because these contributions provide valuable new information. The most straightforward way in which interlocutors could take advantage of these contributions would be to imitate aspects of them, with alignment of mental states following from the imitation of each others' behavior. Assuming there is a relatively regular relationship between behavior and mental states, people who behave in the same way will tend to develop similar mental states. In fact, there is a great deal of evidence for the ubiquity of imitation (for a review, see Hurley and Chater, 2005), and for the repetitive nature of all aspects of dialogue (e.g. Aijmer, 1996; Gries, 2005; Schenkein, 1980; Tannen, 1989).

Perhaps the clearest evidence comes from alignment of situation models themselves. Garrod and Anderson (1987) had pairs of participants play a cooperative maze game, in which they took turns to describe their positions to each other. There are many different ways to

describe one's position, but participants tended to align on the same description scheme. For example, if one player said *I'm two along, four up*, her partner tended to say *I'm one along, five up*; whereas if she said *I'm at B4*, her partner tended to say *I'm at A5*. These players aligned on a "path" or a "coordinate" description scheme, rather than specific words. They also aligned on the interpretation of these descriptions, for example treating the starting point as the bottom left corner of the maze. Schober (1993) had interlocutors describe the position of objects situated between them, so that *left* and *right* had opposite meanings, and found that they tended to align on the use of their interlocutor's perspective (though cf. Schober, 1995). Watson et al. (2004) found that participants were likely to align reference frames, which presumably constitute part of the situation model. A speaker can describe a picture of a dot to the side of a chair that is itself on its side as *the dot to the left of the chair* or *the dot above the chair*. The former involves a relative reference frame (i.e. from an observer's point of view), whereas the latter involves an intrinsic reference frame (i.e. from the chair's point of view). Watson et al. found that participants tended to use a description employing the same reference frame as their confederate interlocutor, and ruled out explanations in terms of lexical priming.

Interlocutors also tend to repeat each other in terms of their grammatical choices. Levelt and Kelter (1982) telephoned (Dutch) shopkeepers and found that they tended to say *At five o'clock* in response to *At what time do you close?*, but tended to say *five o'clock* in response to *What time do you close?* This demonstrates repetition of either syntax or function words. Branigan et al. (2000) had participants take turns to describe and match picture cards, and found that they tended to use the form of utterance just used by their interlocutor. For example, they tended to use a "prepositional object" form like *the pirate giving the book to the swimmer* following another prepositional object sentence, but a "double object" form like *the pirate giving the swimmer the book* following another double object sentence. Such repetition of grammatical form is similar to the syntactic priming that takes place during isolated production (Bock, 1986; Pickering and Branigan, 1998). Similar alignment also occurs with other constructions (Cleland and Pickering, 2003; Haywood et al., 2005), and even between languages in bilinguals (Hartsuiker et al., 2004).

There is also a great deal of evidence for alignment in word choice (sometimes called *lexical entrainment*). For example, Garrod and Anderson (1987) found that interlocutors used the same expressions and, moreover, that they tended to interpret words in the same way (e.g. *line* to mean horizontal row of nodes in the maze) (see also Garrod and Clark, 1993). Brennan and Clark (1996) had directors describe a set of cards depicting common objects to matchers so that they could reconstruct the directors' array. One set of trials contained multiple objects from the same category. Directors and matchers settled on subordinate terms to refer to the objects (e.g. *penny loafer*), because basic-level terms (e.g. *shoe*) would not discriminate between these objects. A subsequent set of trials included one object from each category, so basic-level terms would now be sufficient. However, participants often continued to use the subordinate terms. Results such as these can be interpreted in terms of imitation, though they assume an explanation based on tacit agreement (see below).

Interlocutors also tend to align on such dimensions as accent and speech rate (e.g. Giles et al., 1992). There is also good evidence for "contagion" effects, whereby people tend to imitate each other's facial expressions (e.g. Bavelas et al., 1986) and other indicators of their emotional states (e.g. Hatfield et al., 1994). Not surprisingly, these external indications of alignment lead to alignment of mental states. These data are also compatible with behavioral mimicry (the chameleon effect), whereby people imitate each other's behaviors (Chartrand and Bargh, 1999).

Such imitation appears to be largely automatic, with interlocutors being almost entirely unaware that it has taken place (Pickering and Garrod, 2004). At phonological and acoustic levels it can occur very rapidly and seems largely resource-free (Fowler et al., 2003), with listeners activating appropriate muscles in the tongue while listening to speech but not during non-speech (Fadiga et al., 2002; see also Watkins and Paus, 2004). Additionally, speakers who shadow spoken words produce more faithful imitations of the original when they speak more rapidly (Goldinger, 1998).

Another interesting point is that alignment at one level can lead to greater alignment at other levels. Thus, if interlocutors start referring to entities using the same term, they are likely to accept the implications of that term and therefore conceptualize the entity in the same way. For example, Danet (1980) examined lawyers'

use of terms like *foetus* vs. *unborn child* to refer to the same entity in an abortion trial. In this case, opposing lawyers were careful not to align, presumably because using the opposition's term would at least give the impression of accepting their conceptualization. Clearer evidence comes from experimental studies in which lexical repetition or semantic similarity enhances syntactic repetition between interlocutors. Thus, Branigan et al. (2000) found that such repetition was much stronger (occurring almost eighty percent of the time in their experiment) when prime and target utterances used the same verb, and Cleland and Pickering (2003) found similar results when prime and target used semantically related nouns. Pickering and Garrod (2004) suggest that such interactions between levels explain why alignment at different levels leads to alignment of mental states.

26.2.3 Alignment via agreements between interlocutors

There is something of a tension between the two previous sections, with the former emphasizing decisions to mold one's contributions so that they accord with beliefs about one's interlocutor, and the latter emphasizing automatic processes of imitation. We have already suggested one response: that some apparently strategic decisions may in fact reflect automatic resonance processes (Horton and Gerrig, 2005a; 2005b). However, it is also possible that some apparent imitation is in fact due to tacit agreements between interlocutors about how to refer to entities. Although explicit negotiation about reference (e.g. "Shall we call this X for the purposes of this discussion?") is very rare in most interactions (e.g. Garrod and Anderson, 1987), some researchers argue that interlocutors routinely make so-called "conceptual pacts" when they align on referring expressions (Brennan and Clark, 1996; Metzinger and Brennan, 2003).

As we have noted, Brennan and Clark (1996) found that interlocutors aligned on descriptions of cards depicting objects, even when this required use of a more specific term than was necessary for discrimination. But rather than appeal to imitation, they argued that such interlocutors enter into a conceptual pact in which they (tacitly) agree to keep referring to the same object in the same way. The pact comes about because certain "conversational moves" indicate acceptance of the pact. For instance, if A refers to a shoe as a *penny loafer* and B does not query this use but rather responds to A's instruction, then both A and B assume (1) that B has accepted

this term, and (2) that both know (1). (In contrast, if B queried A's expression, then the pact would not be formed.) On this account, this aspect of alignment is therefore the result of a process of negotiation (grounding) which is specialized to dialogue and involves inference.

Brennan and Clark (1996) used partner-specificity effects to support the conceptual pact explanation of lexical entrainment. They reasoned that conceptual pacts only hold for their originators, so changing the interlocutor should lead speakers to change their descriptions. This did occur to some extent. However, it is important to note that their speakers often first employed the previous term to a new interlocutor, and altered it when the new interlocutor provided explicit feedback to indicate that they had not understood the original description (p. 1491). So it is possible that partner-specificity effects at least partly reflect tendencies for speakers to use the term that they employed with their previous partner first, and only to use a different term when their new partner displays lack of understanding. Since then there have been two major studies on partner-specificity effects which come to rather different conclusions.

Barr and Keysar (2002) recorded eye movements to objects in a display as listeners interpreted repeated references to those objects. Listeners fixated objects more quickly when they were referred to repeatedly, but it did not matter whether the repetition came from the same speaker or a new speaker (who was not present during the initial reference). They concluded that partner-specific effects did not occur, and that only the improved availability of the repeated reference facilitated comprehension. However, Metzger and Brennan (2003) recorded listeners' eye movements in a similar task, but crucially included novel descriptions of previously mentioned objects. For example, an object was repeatedly called *the silver pipe* but then called *the shiny cylinder*. When the new expression was used by the original speaker, addressees searched for a long time before fixating the original object; but when it was used by a new speaker, addressees were not disrupted. They argued that these partner-specific effects were the result of conceptual pacts between interlocutors.

There are therefore empirical issues about the extent to which interlocutors respect partner-specificity in production and comprehension of expressions. But in addition, partner-specificity can be explained without appealing to conceptual pacts. In Horton and Gerrig's (2005b)

terms, people may associate particular expressions with particular interlocutors (or types of interlocutor), so that they are more likely to use those terms when subsequently interacting with them again. For example, Pickering and Garrod (2004; 2005) argue that repeated references to the same objects with the same expressions become routinized and represented as such in memory. It is quite possible that such routines become associated with the person with whom they were established, and that this leads to partner-specificity effects during comprehension and production.

26.2.4 Alignment via feedback

Apart from repetition, there are other aspects of interlocutors' contributions that lead to alignment. Addressees are constantly providing feedback to the speakers that indicate understanding, disagreement, emotional reactions, and so on. When discussing Brennan and Clark's (1996) findings on how speakers changed their references following a change of partner, we pointed out that this was facilitated by explicit feedback from the new partner. This raises the question of whether feedback promotes alignment. Kraut et al. (1982) controlled the amount of feedback that listeners could give to a speaker describing a complicated movie plot. They found a range of evidence to indicate that feedback increases alignment. For example, listeners free to give any form of verbal feedback ("active" listeners) were more likely subsequently to provide descriptions of the movie characters that were similar to those of the speaker than were listeners allowed to give only limited or no verbal feedback ("restricted" listeners). Active listeners' subsequent summaries of the movie plot also showed better understanding of the movie than did those from the restricted listeners. Similarly, Bavelas et al. (2000) found that narrators who received normal feedback from their addressees told better stories than narrators who received highly reduced feedback. Presumably this meant that narrator and addressee ended up with more aligned understanding of the story in the normal feedback condition than in the reduced feedback condition. Using a tangram description task, Schober and Clark (1989) showed that listeners who actively participated in the task by being able to give feedback were better at identifying the tangrams than those who could only overhear the conversation. Presumably the speakers were tailoring their contributions to the specific listener on the basis of the feedback

provided by that listener. Again, this points to the role of feedback in promoting alignment between conversational partners.

Such studies show the effectiveness of feedback, but provide little evidence about how it works, and whether different forms of feedback differ in their effects. For this we need to turn to theoretical analyses of fragments of conversation. According to Clark (1996), addressees respond to contributions by providing signals that the contribution has been grounded; in other words, that the contribution has become part of common ground (well enough for current purpose). Such signals are often (though need not be) feedback, such as *yeah*, *OK*, or a nod, which tend to indicate acceptance and understanding, or *mm*, which may merely indicate understanding. Beyond the prior establishment of common ground, Clark (1996) proposed a specific mechanism that increases the common ground as a consequence of the history of the interaction. According to Clark, the grounding mechanism works by establishing closure on dialogue contributions. In Clark and Schaefer's (1987) terms, a contribution consists of a presentation phase, in which the speaker produces an utterance directed at the addressee, and an acceptance phase, in which the addressee signals whether he has understood the utterance. Acceptance may be signalled through assertions of understanding (*uh-huh*, *I see*, etc.), presupposition of understanding (e.g. making an appropriate next turn), or by display or exemplification (e.g. giving an appropriate answer to a question or offering a paraphrase of what was said in the presentation phase). Crucially in Clark's account, closure and hence updating of common ground requires positive evidence from the interlocutor in the acceptance phase. So, the grounding account contrasts with the interactive alignment account discussed earlier, in which there is no requirement for positive evidence of acceptance.

In an alternative analysis consistent with Pickering and Garrod (2004), feedback such as *yeah*, *OK*, or a nod occurs when the addressee believes he has made a coherent update to his mental state. So long as the speaker has conveyed her message appropriately, such feedback will occur when the interlocutors are aligned, without any need for the addressee to represent the fact that they are aligned. On this account, such feedback serves to promote implicit common ground rather than full common ground (i.e. an explicit model of what is jointly available to the interlocutors). Similarly, feedback such as *eh?* or *pardon?*, or clarification requests such as repetition of part of the speaker's utterance

(Ginzberg, 2004), occurs when the addressee believes he has not made a coherent update to his mental state. This feedback will therefore tend to occur when the interlocutors are not aligned, but again without the addressee needing to represent that they are not aligned. At this point, the speaker can seek to clarify her utterance using interactive repair. The speaker's first attempt to repair can employ a simple strategy of being more explicit about the part of her utterance that the addressee has queried. Although this may require some decision about whether the query refers to the last constituent or the whole utterance (for instance), there is no need to model what she believes the addressee does not know. Finally, the addressee can provide corrective feedback (e.g. replacing one of the speaker's words) when he realizes some aspect of the utterance is incompatible with his background knowledge, and the speaker can incorporate this correction, again without modeling the interlocutor.

26.2.5 Alignment via physical co-presence

A final source of alignment comes from interlocutors' shared physical environment. Schiffer (1972) argued that communicators establish mutual knowledge (similar to common ground) on the basis of physical co-presence, with something which is manifestly present to both interlocutors forming part of mutual knowledge. Clark and Marshall (1981) treated physical co-presence as one of the three sources of common ground (along with community membership and linguistic co-presence). When interlocutors are discussing an aspect of their immediate environment (e.g. objects in view or clearly audible sounds), the environment itself can assist alignment. It can clearly have direct consequences for their shared mental states (e.g. both interlocutors focus on an object that suddenly appears). But it can also serve as an additional communicative resource during dialogue, because interlocutors can combine linguistic expressions with gestures and demonstrations to align their focus of attention.

Bangerter (2004) found that interlocutors substitute pointing gestures (together with deictics such as *this*) for verbal descriptions when referents are visible to both parties and close at hand. He argued that pointing forms part of a composite signal combining gesture with speech to focus joint attention (and hence alignment). Clark and Krych (2004) manipulated whether a director could see a builder's workspace (and hence whether the space was shared) during a

Lego construction task. When the workspace was visible, the director used many deictic expressions (e.g. *here, there, this, that, like this*) and timed speech to fit in with the builder's sequence of actions. She also provided many demonstrations (deliberately positioning objects in ways to elicit responses from the builder). It appeared that the builder's actions were treated as continuous feedback by the director, and directly affected both the content and timing of the speech. Not surprisingly, having access to the shared workspace led to substantially better performance in the collaborative task (see also Kraut et al., 2003). Hence physical co-presence supported alignment.

Again, one can question the extent to which using these devices requires interlocutors to construct models of each other's mental states. On one account interlocutors may reason: If my partner is pointing at something, I can assume that he is attending to it; and if he sees me looking at the object, he knows that I am also attending to it; so it is in common ground. However, the situation may be much simpler than this, because there is evidence that viewers automatically attend to where another is looking or pointing (e.g. Langton and Bruce, 2000), and this directly increases spatial acuity in that region (Schuller and Rossion, 2001). This means that, even if interlocutors take an egocentric perspective, their attention will tend to be aligned with respect to the source of the pointing.

26.3 Discussion

We have argued that dialogue involves the alignment of mental states, so that a particular interaction is successful to the extent that such alignment occurs. Most psychological research on dialogue can be interpreted as investigating the different ways in which such alignment comes about, and we identified five different strands within this research. Throughout this review, we have emphasized that alignment can be due to strategic or automatic factors, with particular pieces of research drawing attention to one or other of these.

Some accounts focus on strategic factors, in which interlocutors construct their contributions and interpret their partners' contributions on the basis of extensive modeling of their partners' mental states and decisions about what contributions would be most useful for them. This is very much the perspective taken by Clark (1992; 1996), and his proposals can be taken as representative of this approach. On this account,

speakers largely construct their utterances for the benefit of their addressees. Most aspects of their speech are intentional (e.g. even disfluencies are seen as deliberate signals that the speaker is having problems; Clark and Fox Tree, 2002). Conceptual pacts are a good example.

By contrast, other accounts treat interlocutors as largely egocentric but operating in a processing environment "designed" to promote alignment. For example, Pickering and Garrod (2004) argue that interactive alignment automatically promotes an implicit common ground. As we have seen in this chapter, there are many potential sources for extending implicit common ground, such as interactive alignment through imitation, resonance, and joint attention through physical co-presence. We suspect that dialogue processing involves a balance between dependence on implicit common ground and dependence on full common ground. To the extent that alignment through imitation and physical co-presence is automatic, relying on implicit common ground is relatively cost-free. More strategic processes of alignment that require the speaker to take the perspective of the addressee would only be used when circumstances allow. Teasing apart the contributions of implicit common ground and full common ground to alignment seems to be a major goal for future research in dialogue.

Finally, we note that our discussion has been largely limited to two-person (dyadic) exchanges considered as isolated events. However, much real dialogue is multi-party, with different people playing different roles (e.g. addressee, side-participant, overhearer) and with the roles often changing throughout the conversation. These different roles greatly affect alignment, for example with addressees aligning with the speaker better than overhearers (Schober and Clark, 1989). Similarly, alignment proceeds differently in small groups than large groups, with interlocutors aligning more with the previous speaker in small groups but with the dominant speaker in large groups (Fay et al., 2000). Additionally, dialogue has long-term effects on a community, with dyads who form part of a community (in that they have previously interacted with the same people) aligning better than dyads who do not form part of a community (Garrod and Doherty, 1994). Alignment between dyads appears to lead to the establishment of community conventions in both experimental studies (Garrod and Doherty, 1994) and simulations (Barr, 2004), and may form part of the explanation for diachronic patterns of language change.

References

- Aijmer, K. (1996) *Conversational Routines in English: Convention and Creativity*. Longman, London.
- Bangerter, A. (2004) Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, 15: 415–19.
- Barr, D. J. (2004) Establishing conventional communication systems: is common knowledge necessary? *Cognitive Science*, 28: 937–62.
- Barr, D. J., and Keysar, B. (2002) Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46: 391–418.
- Bavelas, J. B., Black, A., Lemery, C. R., and Mullett, J. (1986) "I show how you feel": motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, 50: 322–9.
- Bavelas, J.B., Coates, L., and Johnson, T. (2000) Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79: 941–52.
- Bock, J. K. (1986) Syntactic persistence in language production. *Cognitive Psychology*, 18: 355–87.
- Bortfeld, J. K., and Brennan, S. E. (1997) Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, 23: 119–47.
- Branigan, H. P., Pickering, M. J., and Cleland, A. A. (2000) Syntactic coordination in dialogue. *Cognition*, 75: B13–B25.
- Brennan, S. E., and Clark, H. H. (1996) Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22: 1482–93.
- Brown, P. M., and Dell, G. S. (1987) Adapting production to comprehension: the explicit mention of instruments. *Cognitive Psychology*, 19: 441–72.
- Chartrand, T. L., and Bargh, J. A. (1999) The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76: 893–910.
- Clark, H.H. (1992) *Arenas of Language Use*. University of Chicago Press, Chicago.
- Clark, H. H. (1996) *Using language*. Cambridge University Press, Cambridge.
- Clark, H. H., and Fox Tree, J.E. (2002) Using *uh* and *um* in spontaneous speaking. *Cognition*, 84: 73–111.
- Clark, H. H., and Krych, M. A. (2004) Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50: 62–81.
- Clark, H. H., and Marshall, C. R. (1981) Definite reference and mutual knowledge. In A. K. Joshi, I. A. Sag, and B. L. Webber (eds), *Elements of Discourse Understanding*, pp. 10–63. Cambridge University Press, Cambridge.
- Clark, H. H., and Schaefer, E. F. (1987) Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2: 19–41.
- Cleland, S., and Pickering, M. J. (2003) The use of lexical and syntactic information in language production: evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 49: 214–30.
- Danet, B. (1980) "Baby" or "foetus"? Language and the construction of reality in a manslaughter trial. *Semiotica*, 32: 187–219.
- Fadiga, L., Craighero, L., Buccino, G., and Rizzolati, G. (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15: 399–402.
- Fay, N., Garrod, S., and Carletta, J. (2000) Group discussion as interactive dialogue or as serial monologue: the influence of group size. *Psychological Science*, 11: 481–86.
- Fowler, C. A., Brown, J. M., Sabadini, L., and Weihing, J. (2003) Rapid access to speech gestures in perception: evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49: 396–413.
- Garrod, S., and Anderson, A. (1987) Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, 27: 181–218.
- Garrod, S., and Clark, A. (1993) The development of dialogue coordination skills in school children. *Language and Cognitive Processes*, 8: 101–126.
- Garrod, S., and Doherty, G. (1994) Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, 53: 181–215.
- Giles, H., Coupland, N., and Coupland, J. (1992) Accommodation theory: communication, context and consequences. In H. Giles, J. Coupland, and N. Coupland (eds), *Contexts of Accommodation*, pp. 1–68. Cambridge University Press, Cambridge.
- Ginzburg, J. (2004) Intrinsic misalignment in dialogue: why there is no unique context in a conversation. (Commentary on Pickering and Garrod's *Toward a Mechanistic Psychology of Dialogue*). *Behavioral and Brain Sciences*, 27: 197–9.
- Glucksberg, S., Krauss, R. M., and Weisberg, R. (1966) Referential communication in nursery school children: method and some preliminary findings. *Journal of Experimental Child Psychology*, 3: 333–42.
- Goldinger, S. D. (1998) Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105: 251–79.
- Gries, S. T. (2005) Syntactic priming: a corpus-based approach. *Journal of Psycholinguistic Research*, 34: 365–99.
- Halpern, Y., and Moses, Y. (1990) Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37: 549–87.
- Hanna, J. E., Tanenhaus, M. K., and Trueswell, J. C. (2003) The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49: 43–61.
- Hartsuiker, R.J., Pickering, M.J., and Veltkamp, E. (2004) Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15: 409–14.
- Hatfield, E., Cacioppo, J., and Rapson, R.L. (1994) *Emotional Contagion*. Cambridge University Press, New York.
- Haywood, S. L., Pickering, M. J., and Branigan, H. P. (2005) Do speakers avoid ambiguities during dialogue? *Psychological Science*, 16: 362–6.
- Hintzman, D. C. (1986) Schema abstraction in a multiple-trace memory model. *Psychological Review*, 93: 411–28.

- Horton, W. S., and Gerrig, R. J. (2005a) Conversational common ground and memory processes in language production. *Discourse Processes*, 40: 1–35.
- Horton, W. S., and Gerrig, R. J. (2005b) The impact of memory demands on audience design during language production. *Cognition*, 96: 127–42.
- Horton, W. S., and Keysar, B. (1996) When do speakers take into account common ground? *Cognition*, 59: 91–117.
- Hurley, S., and Chater, N. (eds) (2005) *Perspectives on Imitation*. MIT Press, Cambridge, MA.
- Isaacs, E. A., and Clark, H. H. (1987) References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116: 26–37.
- Keysar, B., Barr, D. J., Balin, J. A., and Brauner, J. S. (2000) Taking perspectives in conversation: the role of mutual knowledge in comprehension. *Psychological Science*, 11: 32–7.
- Keysar, B., Barr, D. J., Balin, J. A., and Paek, T. S. (1998) Definite reference and mutual knowledge: process models of common ground in comprehension. *Journal of Memory and Language*, 39: 1–20.
- Kraut, R. E., Fussell, S. R., and Siegel, J. (2003) Visual information as a conversational resource in collaborative physical tasks. *Human–Computer Interaction*, 18: 13–49.
- Kraut, R. E., Lewis, S. H., and Swezey, L. W. (1982) Listener responsiveness and the coordination of conversations. *Journal of Personality and Social Psychology*, 43: 718–31.
- Langton, S.R., and Bruce, V. (2000) You must see the point: automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, 26: 747–57.
- Levelt, W. J. M., and Kelter, S. (1982) Surface form and memory in question answering. *Cognitive Psychology*, 14: 78–106.
- Logan, G. D. (1988) Toward an instance theory of automatization. *Psychological Review*, 95: 492–527.
- Metzing, C., and Brennan, S. E. (2003) When conceptual pacts are broken: partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49: 201–13.
- Pickering, M. J., and Branigan, H. P. (1998) The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, 39: 633–51.
- Pickering, M. J., and Garrod, S. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27: 169–225.
- Pickering, M. J., and Garrod, S. (2005) Establishing and using routines during dialogue: Implications for psychology and linguistics. In A. Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, pp. 85–102. Erlbaum, Mahwah, NJ.
- Schlenker, J. (1980) A taxonomy for repeating action sequences in natural conversation. In B. Butterworth (ed.), *Language Production*, vol. 1: pp. 21–47. Academic Press, London.
- Schiffers, S. R. (1972) *Meaning*. Oxford University Press, Oxford.
- Schober, M. F. (1993) Spatial perspective-taking in conversation. *Cognition*, 47: 1–24.
- Schober, M. F., and Brennan, S. E. (2003) Processes of interactive spoken discourse: the role of the partner. In A. C. Graesser, M. A. Gernsbacher, and S. R. Goldman (eds), *Handbook of Discourse Processes*, pp. 123–64. Erlbaum, Hillsdale, NJ.
- Schober, M. F., and Clark, H. H. (1989) Understanding by addressees and over-hearers. *Cognitive Psychology*, 21: 211–32.
- Schuller, A. M., and Rossion, B. (2001) Spatial attention triggered by eye gaze increases and speeds up early visual acuity. *NeuroReport*, 12: 381–86.
- Shatz, M., and Gelman, R. (1973) The development of communication skills: modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, 38.
- Tannen, D. (1989) *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge University Press, Cambridge.
- Watkins, K., and Paus, T. (2004) Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience*, 16: 978–87.
- Watson, M., Pickering, M. J., and Branigan, H. P. (2004) Alignment of reference frames in dialogue. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Erlbaum, Mahwah, NJ.