Research on Language and Computation (2006) 4:203–228
DOI 10.1007/s11168-006-9004-0

Alignment as the Basis for Successful Communication

MARTIN J. PICKERING[1] and SIMON GARROD[2]
[1]*Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK (E-mail: martin.pickering@ed.ac.uk);* [2]*Department of Psychology, University of Glasgow, Glasgow, UK.*

Abstract. Pickering and Garrod (2004) argued that alignment is the basis of successful communication in dialogue. In other words, successful communication goes hand-in-hand with the development of similar representations in the interlocutors. But what exactly does this mean? In this paper, we attempt to define alignment, contrasting alignment of situation models with alignment of linguistic representations. We then speculate on how these notions are related and why they lead to conversational success.

Key words: dialogue, alignment, situation model.

Although there has been a great deal of research into the psychological mechanisms underlying language use during the past 40 years or so, the focus has been almost entirely on the processing of isolated words, sentences and texts. Most of this research has investigated the way in which people successfully decode linguistic stimuli, turning them from sound into meaning, or encode non-linguistic ideas in sound. In response to this, Pickering and Garrod (2004) proposed an account of the mechanisms that interlocutors employ during dialogue. Their fundamental claim was that interlocutors align their linguistic representation during dialogue, with successful communication occurring when they become well aligned. Most of the paper developed a theory of alignment in which automatic processes play a central role and explicit modelling of one's interlocutor is secondary (see also Garrod and Pickering, 2004).

Our goal in this paper is to provide a more precise characterization of alignment. We first outline the interactive-alignment account, and then argue that alignment is achieved by what we call alignment of information states rather than information transfer. Next, we contrast the alignment of situation models with alignment of non-situational knowledge. The rest of the paper is concerned with the psychological mechanisms

underlying alignment, and focuses on the development and use of routines, the interpretation of alignment as a largely mechanistic process, the process of interactive repair, the decision not to align, and the impact of alignment on novel contributions to dialogue.

1. The Interactive-alignment Model of Dialogue

Pickering and Garrod (2004) argued that interlocutors align situation models during dialogue. We assume that a situation model is a multi-dimensional representation of the situation under discussion (Johnson-Laird, 1983; Sanford and Garrod, 1981; Zwaan and Radvansky, 1998). According to Zwaan and Radvansky, the key dimensions encoded in situation models are space, time, causality, intentionality, and reference to main individuals under discussion. Such models are assumed to capture what people are "thinking about" while they understand a text, and therefore are in some sense within working memory (they can be contrasted with linguistic representations on the one hand and general knowledge on the other). Successful dialogue occurs when interlocutors construct similar situation models to each other. In our terminology, their situation models become aligned. The main mechanism of alignment is automatic and primarily unconscious, and is the result of interlocutors tending to produce and interpret expressions in the same ways that their partners have just done. In other words, alignment of situation models is largely the result of alignment at other levels of representation.

In fact, interlocutors tend to produce words they have just heard, to assume that ambiguous words have meanings that they have recently given to those words, to use grammatical constructions they have recently used, and so on. For example, Garrod and Anderson (1987) found that participants in a cooperative maze game tended to describe their positions in the maze using similar words and interpretations for those words (e.g., *line* to mean horizontal row of nodes in the maze), and to employ similar description schemes (e.g., referring to positions by coordinates with the same origin). Similarly, Brennan and Clark (1996) had interlocutors describe cards to each other, and found each interlocutor tended to mirror the other's (often idiosyncratic) descriptions. Moreover, they often retained distinctions (e.g., specific details about the type of object involved) when these distinctions were no longer necessary for identification.

Likewise, Branigan et al. (2000) found that participants in an interactive card-sorting task tended to describe a card using the same grammatical structure (a dative) that their partner had just used when describing a card to them. Cleland and Pickering (2003) found similar results for the production of complex noun phrases (e.g., *the sheep that is red* vs. *the red sheep*). Haywood, Pickering, and Branigan (2005) found syntactic alignment even when choice of form affected the ambiguity of an instruction. Hartsuiker,

Pickering, and Veltkamp (2004) found alignment between languages, with bilingual participants tending to use an English passive more often if their partner had just used a Spanish passive than if she had just used a Spanish active (cf. Schoonbaert et al., in press). These findings suggest syntactic alignment in a range of different types of construction. Finally, Levelt and Kelter (1982) found that people tended to respond to a question with a lexically and syntactically congruent answer.

There is also evidence that interlocutors directly align at least one specific aspect of their situation models, namely spatial reference frames (e.g., Levinson, 2003). Schober (1993) had interlocutors describe the position of objects situated between them, so that *left* and *right* had opposite meanings, and found that they aligned on using their interlocutor's perspective (though cf. Schober, 1995). Watson, Pickering, and Branigan (2004) had interlocutors describe a dot next to a phone on its side as *the dot left of the phone* (i.e., from the speaker's perspective) or *the dot below the phone* (i.e., from the phone's perspective) and found that they aligned on the reference frame.

Recent corpus evidence indicates that syntactic alignment is an important factor in real conversations. Gries (2005) identified 3003 dative-alternations (e.g., *give a book to the boy* vs. *give the boy a book*) in a tagged and parsed corpus of spoken British English (the ICE-GB corpus). Even though the two forms of the alternation were of very similar frequency, he found that 63% of examples of the alternation were preceded by the same form of the alternation rather than by the other form. Notice that there were often many intervening sentences between these two utterances, though priming was stronger when the sentences were closer together. Although Gries investigated both within- and between-speaker effects, the analyses did not show any effect of speaker identity on priming. Gries also showed similar effects for the verb-particle construction (e.g., *gave the job up* vs. *gave up the job*). Szmrecsanyi (2005) found comparable effects using different constructions.

Pickering and Garrod (2004) proposed that syntactic and lexical alignment lead to alignment of the situation model, on the basis of evidence that alignment at one level of representation leads to more alignment at other levels of representation. For example, Branigan et al. (2000) found that priming was especially strong when prime and target descriptions employed the same verb. Hence, alignment at one level (here, the lexicon) leads to more alignment at another (here, syntax). Similarly, Cleland and Pickering (2003) found that alignment of noun phrase form (e.g., *the sheep that's red*) was enhanced after a noun phrase using the same noun (e.g., *the sheep that's red*) or a semantically related noun (e.g., *the goat that's red*), in comparison to a semantically unrelated noun (e.g., *the knife that's red*). Similar effects even occur between languages in bilinguals (Schoonbaert et al., in press).

This primitive priming mechanism requires no processing effort and entails no explicit negotiation between interlocutors. As part of this process, interlocutors do not model each others' mental states but simply align on each other's linguistic representations. However, interlocutors also employ a simple repair mechanism to rectify failures in alignment. When interlocutors determine that they cannot interpret new input in relation to their existing situation model, they then reformulate the utterance in a way that is consistent with the model. The following illustrates this with an example taken from a corpus of maze conversations (Garrod and Anderson, 1987).

1.

A: You are starting from the left, you're one along, one up?*(2 sec.)*
B: Two along: I'm not in the first box, I'm in the second box.

They can also seek clarification, for example by querying a word used by their interlocutor. If all else fails, they may explicitly model each other's mental states (or engage in explicit negotiation). In summary, Pickering and Garrod (2004) proposed that successful dialogue involves the alignment of situation models, and that this occurs via three processes: (1) an automatic mechanism of alignment involving priming at all levels of linguistic representation and percolation between these levels; (2) a mechanism that repairs alignment failure; (3) alignment via explicit "other modelling", which is used as a last resort.

Although alternative accounts assume that interlocutors build up a shared body of knowledge often called common ground, they do not assign a central role to priming. Clark (1996; Clark and Schaeffer, 1987) proposed that an interlocutor provides a contribution, which her partner either accepts or does not accept. The indication of acceptance can be an assent (e.g., *yeah*, *OK*), but can also be a new contribution (the assumption being that the partner would query if there were a problem, so the lack of query indicates acceptance). Following acceptance, the contribution enters common ground, so that both interlocutors assume that it is shared knowledge. Although Clark accepts that interlocutors tend to converge on the same expressions (Brennan and Clark, 1996), this convergence is not necessary for the construction of common ground. Interlocutors could build up extensive common ground without converging at any linguistic levels. Whereas we accept that conversation could in principle be successful without linguistic convergence, we argue that it would be greatly impoverished in practice.

2. Information Transfer Versus Informational Alignment

Standard psycholinguistic accounts assume an *information transfer* view of language use. Production involves *encoding* meaning into sound, and

comprehension involves *decoding* sound into meaning. Most psycholinguistics involves explicating the stages that occur in one or other of these processes (e.g., Fodor et al., 1974; Levelt, 1989). For example, the standard account of speech comprehension assumes that the listener hears a sequence of sounds, converts those sounds into lexical representations, then organizes those lexical representations into syntactic representations and extracts their meaning. This is a process of code-breaking in which the listener has no effect on the stimulus itself. Likewise, a speaker starts out with a message, converts that message into lexical, syntactic, and phonological representations, then articulates the sounds.

One view of dialogue simply imports these models into dialogue, so that it involves serial processes of encoding and decoding, with one interlocutor encoding while the other decodes. The problem with this approach stems from the fact that each interlocutor does not contribute autonomously to dialogue. Instead, dialogue should be regarded as a joint activity (e.g., Clark, 1996), akin to collaborative problem solving or ballroom dancing. We propose that it involves *informational alignment*, whereby interlocutors tend toward equivalent information states.

Let us first outline a non-interactive information transfer account and then contrast it with the informational alignment view. On this account, a speaker has a message that she wants to convey (roughly corresponding to a message-level representation in Levelt, 1989). This message is dependent on the situation model of the speaker at that point. To illustrate this, consider another extract from a maze game dialogue in which the two participants are attempting to describe the positions shown in the maze in Figure 1 (from Garrod and Anderson, 1987).

2.

1A: You know the extreme right, there's one box
2B: Yeah right, the extreme right it's sticking out like a sore thumb
3A: That's where I am
4B: It's like a right indicator?
5A: Yes, and where are you?
6B: Well I'm er- that right indicator you've got
7A: Yes
8B: The right indicator above that.

According to the information transfer account, *A* starts out with a model of the maze which represents certain configurations of positions and captures spatial relations between them (see Garrod and Anderson, 1987, for a detailed discussion of such models). He identifies his position with respect to this model (see the lower arrow in Figure 1) and then encodes the message into appropriate speech sounds. He has many choices about which

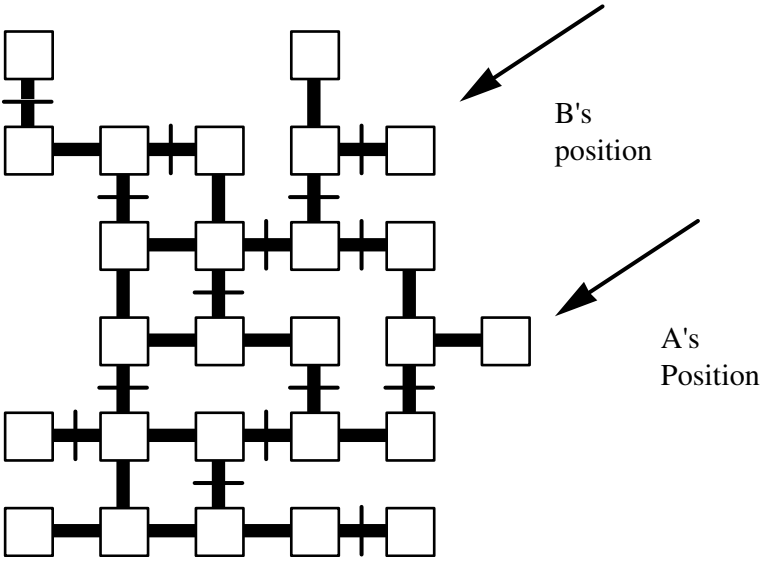

Figure 1. The maze configuration the *A* and *B* are talking about in the example dialogue in the text. The labelled arrows indicate *A*'s and *B*'s positions at the time of the dialogue.

expressions to use; such ambiguity is ubiquitous in language production. He may have high activation levels for the following: the lexical items *extreme* and *right*, a particular phonological realization of *right*, and the viewer-centred frame of reference in choosing right. The way that the speaker converts the message into sound is affected by these activation levels, so that they "direct" the message into a particular acoustic form, *You know the extreme right, there's one box*. On this account, these states of activation serve as the medium through which the process of information transfer moves, as it were.

Exactly the opposite process takes place in comprehension, where the listener converts the sounds into meaning. Again, there is considerable ambiguity in any sequence of speech sounds, so that the listener has to determine the phonemes, segment the signal into words, determine the appropriate meaning for homonyms, determine the grammatical structure of utterances, and determine the reference frame. In this case, the listener would need to determine what *box* refers to (a square configuration of positions on the maze or an individual position; see Figure 1) and whether *the extreme right* is relative to the speaker's perspective. The listener's levels of activation (e.g., for the different interpretations of *box*) would "direct" his process of information transfer from sound into meaning.

The problem with this information transfer account for dialogue is that interlocutors do not separate neatly into speaker and listener. There

are two reasons for this. First, dialogue does not involve long discrete contributions from each interlocutor in turn, but involves a constant process of feedback, which means that contributions result from joint actions in which both interlocutors play an active role. Second, the activation of knowledge like the words *extreme* or *right* does not tend to result from autonomous processes within the interlocutor, but can just as well be due to prior contributions by her partner.

To explain this, consider the first two utterances of (2). If *A* had uttered *far right*, it would be much more likely that *B* would have responded with *far right* (Brennan and Clark, 1996; Garrod and Anderson, 1987). But because *A* produced *extreme right*, *B* also used *extreme right*. In other words, the activation of *extreme* and *right* that led to *B*'s contribution was largely the result of *A*'s contribution. This means that the activation of knowledge that affects *B*'s contribution did not come from *B* on his own, but was at least partly due to *A*. In other words, the processes that led to *B*'s contribution are the result of joint action by *A* and *B*.

Additionally, at the turn 5 *A* interrupts *B* with 'yes', which serves as an indicator that *A* accepts *B*'s contribution. Had *A* said 'eh?' rather than 'yes', it would have indicated that *A* did not follow *B*'s utterance. Most likely, *B* would then reformulate, perhaps replacing 'that right indicator you've got' with an alternative description. This means that *B*'s full utterance is not the result of processes that are autonomous to *B*, but instead follows from *A*'s intentions as well as *B*'s. Given that dialogue involves a great deal of feedback (much of it non-linguistic, of course, and not therefore indicated in transcripts like (2)), it results from joint activity.

These problems with the information transfer approach undermine its main attraction, which is the idea that each choice by the speaker (e.g., use of *extreme* rather than *far*) can be seen as corresponding to information to be transferred to the listener, with each choice leading to a potentially quantifiable reduction of uncertainty in the listener. But, as we have shown such choices may have originated with the listener in the first place because of the joint nature of dialogue. So in practice it is impossible to establish what reduction in uncertainty, if any, has actually occurred in the listener.

The informational-alignment account assumes that dialogue is better characterized as a process of aligning information states than as a process of transferring information from speaker to listener. According to informational alignment, both interlocutors represent situation models that can contribute to the speaker's utterance, with the speaker contributing via the primary utterance and the addressee contributing via feedback. When interlocutors are well-aligned, their situation models are aligned with respect to the topic of discussion. Additionally, the history of the interaction (together with other things such as their common cultural heritage and past shared experiences on the one hand, and the shared visual

environment on the other) means that they have similar levels of activation of relevant knowledge, such as the activation of lexical items, grammatical rules, and so on. This means that the interlocutors are likely to produce their contributions in similar ways, not only discussing the same situation, but also using the same words, constructions, and so on. Hence, the dialogue is likely to be relatively fluent and misunderstanding will be uncommon.

3. The Role of Non-situational Knowledge in the Alignment of Situation Models

In order to understand the alignment of situation models, we now need to explicate the relationship between the way in which interlocutors develop their models and the way in which they activate other knowledge, for example knowledge about words and syntax. We then consider the process of routinization within this account.

3.1. Content of situation model

First, we need to clarify what is represented in a situation model and what is not. Traditionally, situation models have been discussed in the context of text comprehension (i.e., monologue). In such accounts, successful comprehension is taken to reflect the construction of a coherent situation model to represent the particular state of affairs described in the text (Zwaan and Radvansky, 1998). In relation to the dialogue extract in (2) above this would correspond to some kind of spatial model of the maze configuration representing different spatial entities in the maze and the relations between them. A portion of the model might be as represented in the top section of Figure 2. This portion identifies spatial entities in the maze (configurations RI_1 and RI_2), which will eventually be labelled as right indicators, and the spatial relations between these entities (the schematic picture of the maze is used to represent the mappings between the elements RI_1 and RI_2 in the model and the maze being talked about).

The specific contents of a person's situation model will of course depend on the way that he interprets the situation. Clearly, this is likely to change as the dialogue unfolds. In fact the entity RI_2 may only enter speaker *B*'s model as he formulates the description of his position given in Turn 6B of (2), *Well I'm er- that right indicator you've got*. Precisely how much information is represented in the model will differ according to how much he could reasonably infer on the basis of his general knowledge and the context. General knowledge does not of course form part of the situation model (see below).

Notice that the situation model is a model of the speaker's mental state, and does not automatically contain information about whether the

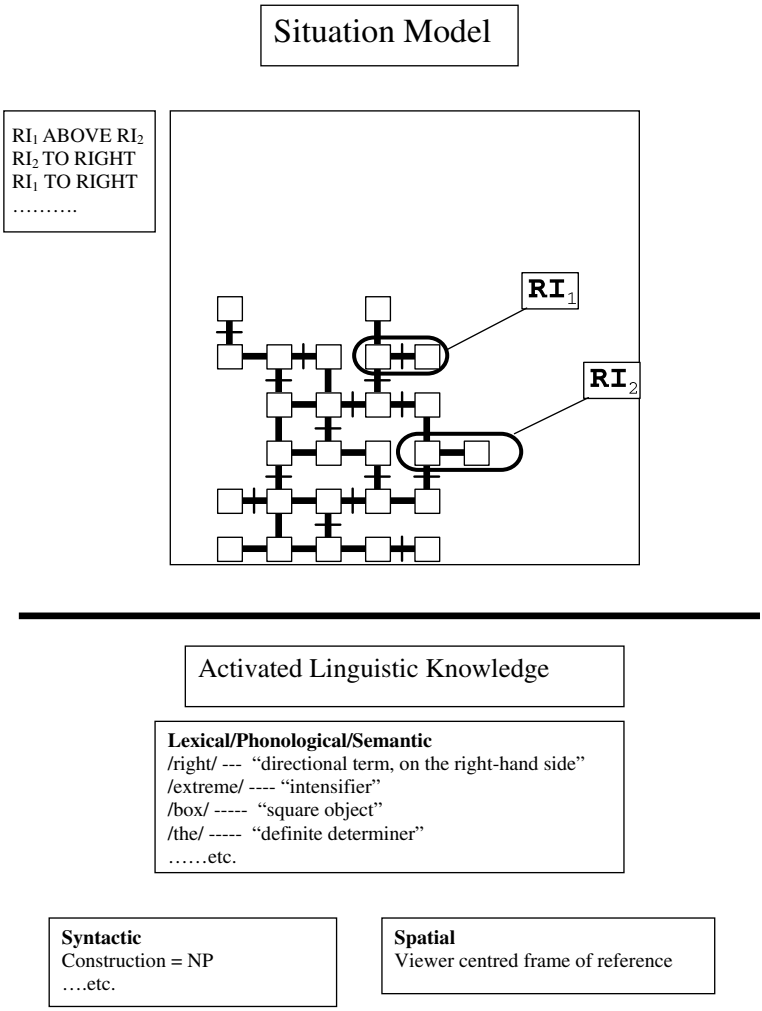

Figure 2. Schematic representation of the current situation model and activated linguistic knowledge when producing the utterance *You know the extreme right, there's one box* (Example 2, Utterance 1A).

information is also known to the interlocutor (or indeed whether it forms part of common ground; see Clark and Marshall, 1981). If the interlocutors are well-aligned, the speaker and the addressee's models are very similar, so a new contribution that is consistent with the speaker's model should be consistent with the addressee's model as well. By default, the speaker therefore assumes that any element in her model is also present in the addressee's model. Interlocutors are able to index information in their models to indicate that information is not shared, but such indexation incurs a processing cost, which sometimes leads to mistakenly

egocentric behaviour (e.g., Horton and Keysar, 1996). In general, such indexation is not necessary in successful conversation between equals, but may be needed in expert-novice interactions (e.g., Isaacs and Clark, 1987), when the addressee indicates failure to understand and the speaker does not immediately find some other way of conveying the same information (e.g., Clark and Wilkes-Gibbs, 1986), or in some complex multi-party interactions (e.g., when concealing meaning from overhearers; Clark and Schaefer, 1987). Notice that this account differs from Clark (1996), who assumes that speakers carefully track their addressees' mental states throughout conversations.

3.2. Non-situational knowledge and its activation

The situation model is comparatively limited in extent: it contains information relevant to the situation being described. The goal of dialogue is to construct aligned situation models, so that interlocutors have similar understandings of the situation. However, the interactive-alignment model proposes that interlocutors become aligned at many different linguistic levels, many of which are not part of the situation model itself. Additionally, the great majority of people's general knowledge does not form part of the situation model. Hence, knowledge that forms part of the situation model needs to be distinguished from two other forms of knowledge: linguistic knowledge and general knowledge (see Figure 3).

Let us first consider linguistic knowledge. A major part of this knowledge is the mental lexicon, containing entries such as those for *extreme, box,* and *right*. For example, it includes information about the meaning of the words, including the fact that *right* has more than one meaning (e.g., "correct" or a spatial direction), their grammatical category, their phonology, and so on. Importantly, these entries all have a level of activation, which can depend on "long term" factors such as their frequency, though in dialogue activation is heavily influenced by their use within that interchange. Other linguistic knowledge includes information about syntactic rules (e.g., the conditions of use of the passive), interpretation of thematic roles, or information about reference frames. Wholly new knowledge can be added, as in the development of routines (see below). Note that "linguistic" may be too narrow a term because interlocutors can align with respect to non-linguistic aspects of communication, such as their gestures. Additionally, information about reference frames may or may not be linguistic, depending on definition. We use "linguistic" to make a clear distinction from general knowledge on the one hand and knowledge about the situation model on the other.

To see how this linguistic knowledge is different from the situation model itself, consider again Turn 1A of dialogue extract (2). The listener

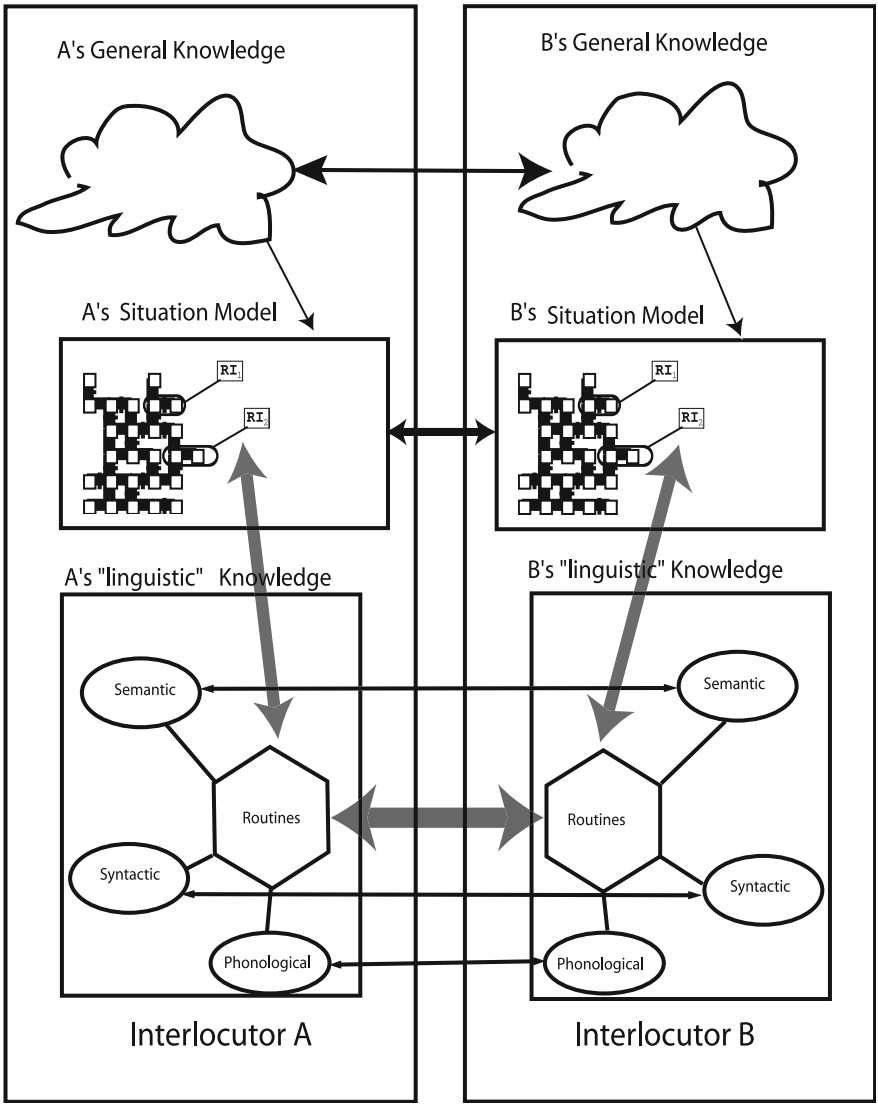

Figure 3. Schematic representation of the situation models and associated activated non-situational knowledge (linguistic and general knowledge) for two interlocutors *A* and *B* at the time of Example 2.

can construct a situation model that represents this information as discussed above. In addition, the process of comprehension activates linguistic information at many levels. For example, it activates words like *extreme*, *right*, and *box*, syntactic constructions, phonological information and so on (see Figure 2, bottom). Even though this linguistic information may be relevant for the listener's understanding of the utterance it would not form

part of the situation model. When the interlocutors move on to a different topic, information about the specific location (e.g., configuration RI_2) may leave the situation model, but some of the activation of *box* is likely to remain.

General knowledge is also separate from the situational knowledge. Most general knowledge is unrelated to the situation model, because it is irrelevant to the topic under discussion. However, some of it is closely linked to the knowledge that forms part of the situation model. A good example might be knowledge about the workings of cars that presumably led to the development of the term *right indicator* in (2). This knowledge is relevant to the situation model but is clearly not part of it. The distinction between general knowledge and situational knowledge is similar to the distinction between long-term memory and working memory, perhaps most obviously with respect to approaches that treat working memory as the activated component of long-term memory (e.g., Eriksson & Kintsch, 1995).

3.3. Linguistic and general knowledge, situation models, and alignment

Now consider the relationship between situation model and non-situational knowledge in the context of alignment. For this we turn to Figure 3, which represents linguistic knowledge as a set of nodes corresponding to different aspects of linguistic information (e.g., syntax). We ignore routines for the purposes of this section.

Consider first the "vertical" relation between the situation model and linguistic knowledge for each interlocutor. When *A* produces *You know the extreme right, there's one box* (Example 2, Turn 2B), she activates a situation model containing tokens corresponding to the entities, their spatial relationships, and so on (see Figure 3, middle). Additionally, she activates linguistic knowledge corresponding to the lexical entries for words like *extreme*, *right*, and *box*, information about phonology and reference frame (e.g., whether the box is to the right of the maze from the viewer's perspective), and so on (see Figure 3, bottom). When an addressee understands the same utterance, he also activates the same situation model and linguistic information. Hence, the vertical arrows are two-way, and it does not matter which interlocutor is the speaker and which the addressee (or whether, in fact, the utterance is the result of contributions by both interlocutors). This is a consequence of the assumption of parity of representations between comprehension and production (Pickering and Garrod, 2004): the same situation models and linguistic knowledge are activated (and employed) whether an utterance is comprehended or produced.

After *A* utters *You know the extreme right, there's one box* and *B* comprehends it, the interactive-alignment model predicts that they are likely to become well aligned (Pickering and Garrod, 2004). We can now see that interlocutors align with respect to their situation models and with respect to their linguistic knowledge. This duality of alignment is captured by the horizontal arrows (called *alignment channels* by Pickering and Garrod, 2004) in Figure 3, where the arrows between the situation models should be contrasted with the set of arrows between components of linguistic knowledge. Two interlocutors have aligned situation models to the extent that their situation models contain the same entities and relations. They have aligned linguistic knowledge to the extent that they have similar patterns of activation of linguistic knowledge. But in addition they are likely to align their general knowledge to at least some extent. If general knowledge about cars' indicators causes *A* to introduce information associated with right indicators into her situation model, and *B* constructs the same situation model as *A*, then *B* is also likely to think about cars' indicators "off-line", and it is therefore more likely that eventually *A* and *B* will come to share this aspect of general knowledge, even though it never formed part of their situation models. This motivates the alignment channel between *A* and *B*'s representations of general knowledge in Figure 3. However, successful conversations occur between interlocutors whose general knowledge is quite different in respects that are irrelevant to the conversation at hand, so alignment of general knowledge will be slow and partial at best. We briefly return to alignment of general knowledge in the final section.

3.4. Misalignment

In the "ideal" case, linguistic and situational alignment go hand-in-hand. However, interlocutors will sometimes become better aligned in one respect than in the other. For example, in turns 6 and 8 of (2), *B* described a new position (i.e., *The right indicator above that*). In this case, he may have retained roughly the same activation of *extreme* and *right*, a viewer-centred reference frame, and so on, but has now extended his situation model to include the new spatial entity RI_1 not yet present in the interlocutor's model. Such changes in the situation model are of course a normal part of conversation when new information is introduced. It is also possible that he might have realized that he has correctly understood *A*'s previous description, and hence activated the other meaning for *right*, but retained the previous situation model. In this case, some alignment of linguistic knowledge would be lost, without concomitant loss of alignment of situation model.

Such disassociations can occur in other ways. Consider the situation schematized in Figure 4a in which interlocutors *A* and *B* are both talking

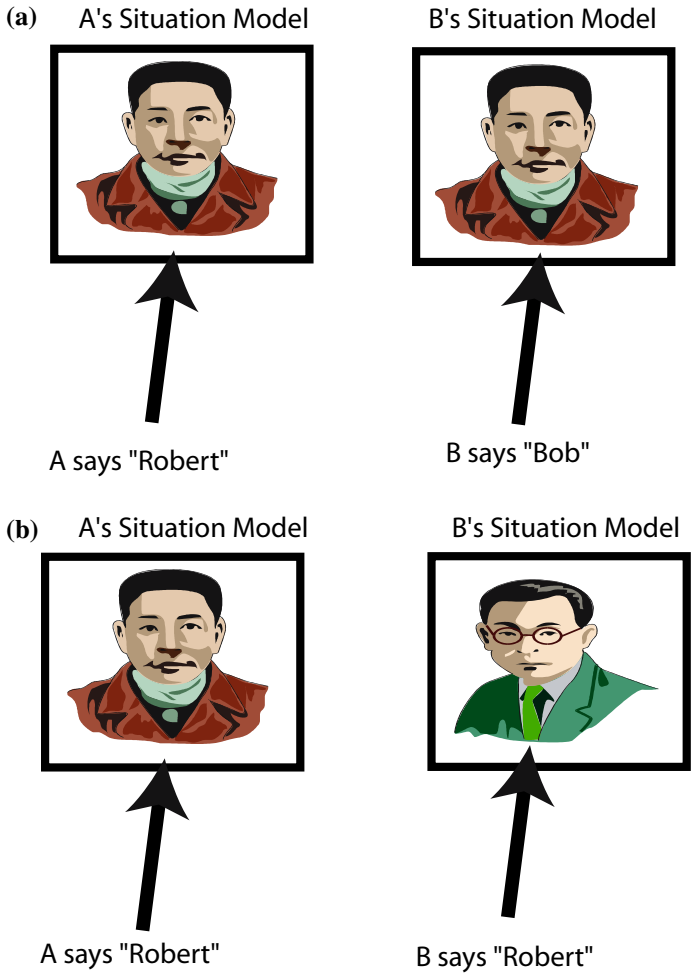

Figure 4. Different forms of misalignment. In (a) interlocutors A and B use different names for the same individual in their respective situation models. In (b) interlocutors A and B use the same name for different individuals in their respective situation models.

about the same person but *A* refers to him as *Robert* whereas *B* refers to him as *Bob*. Here *A*'s situation model is aligned with *B*'s situation model, but the mapping between names and interpretations is not aligned. Conversely, in Figure 4b both *A* and *B* are using the same name *Robert* but in this case not to refer to same man, and so the situation models are not aligned.

Of course, choice of description may affect the interlocutors' interpretations of what they are saying because the interpretation may be affected by what they might associate with the different forms of description and hence

what they might infer about the situation. We discuss such misalignments later but it is important to point out that the interactive-alignment model regards them as exceptional. So the conclusion is that alignment of linguistic knowledge and situation model are separate aspects of the overall process of alignment, but that they normally go in step with each other.

3.5. Routines and routinization

Figure 3 highlights the central role that routines have in our account. Within each interlocutor, they serve as the interface between specific components of linguistic knowledge such as syntax, semantics, and phonology on the one hand, and the situation model on the other. Additionally, they play a major role in the links between the interlocutors, in that they constitute an important channel of alignment.

A routine is an expression that is "fixed" to a relatively large extent. It occurs at a much higher frequency than the frequency of its components would lead us to expect (e.g., Aijmer, 1996). Stock phrases, some idioms, and some clichés are extreme forms of routines, if they have one immutable form. However, semi-productive expressions that contain some fixed elements and some elements that vary constitute less extreme routines. Most discussion of routines relates to expressions that are in some sense "permanent" within a language user or group (e.g., Aijmer, 1996; Kuiper, 1996; Nunberg, Sag, and Wasow, 1994). But routines may also be established during a particular dialogue and forgotten soon afterwards (e.g., Brennan and Clark, 1996; Garrod and Doherty, 1994; Malt and Sloman, 2003), in which case they have a temporary status. If one interlocutor starts to use an expression and gives it a particular meaning, the other will most likely follow suit; when the conversation ends, both interlocutors may abandon the expression and its special meaning. This observation justifies the assumption of an alignment channel between the *routines* boxes in Figure 3. Although we tend to think of routines as multi-word expressions, our account also allows routines to be individual words, as when interlocutors use a word in a novel way (e.g., in tangram description).

Various experimental studies show this process of routinization. For example, in (2) *B* says *It's like a right indicator* to refer to a box protruding from the maze on the right. *A* accepts this novel use of the expression, and they both then repeatedly used *right indicator* to refer both to this position and to other boxes protruding to the right. This suggests that both interlocutors develop this routine. However, there is no reason to assume that they will continue to use *right indicator* in this way after they have finished playing the maze game. Similar processes occur when interlocutors agree on a "shorthand" description of unfamiliar objects, as when referring to a geometric shape as *an ice skater* (e.g., Clark and Wilkes-Gibbs, 1986).

Pickering and Garrod (2005) interpreted routines in terms of Jackendoff's (2002) linguistic framework (see also Jackendoff, 1999). Jackendoff assumed a "rich" lexicon, which does not just contain individual words or morphemes, but also includes any form of complex expression that can be accessed directly, including idioms, constructions, and even whole speeches if they have been memorized. Lexical entries contain separate representations for syntax, semantics, and phonology, which are associated via "linking rules". For example, the "normal" interpretation of *right* is represented in Figure 5a, and that of *indicator* is represented in Figure 5b. These can be combined (during comprehension or production) to produce the representation for the complex expression given in Figure 5c, so that it has the standard compositional meaning "pointer on the right". However, this representation is not lexicalized. In contrast, the routinized interpretation of *right indicator* developed by the maze-game participants is represented in Figure 5d, so that it has an idiosyncratic meaning that is not compositional (as indicated by the link between its meaning and the N' node). Whereas Jackendoff (2002) focused on the description of a relatively fixed lexicon, the lexicon can change all the time during dialogue, with new expressions becoming routinized (and, presumably, other previously routinized expressions being "lost"). Evidence for this comes from experimental studies like Garrod and Doherty (1994) or Brennan and Clark (1996), where terms used in one interaction need not carry over to another interaction.

The potentially short-term nature of routinized expressions is closely related to our claim that the vertical links between linguistic knowledge and situation model (see Figure 3) involve the routines component of linguistic knowledge. Informally, we suggest that when comprehending an expression, people activate the syntax, semantics, and phonology of the expression, and integrate them into a "chunk" that can then be used to "enter" the meaning of the expression into the situation model. The integration is needed in order to compute an unified representation of the expression corresponding to one of Jackendoff's (complex) lexical entries. Similarly, when producing an expression, people construct a lexical entry on the basis of the situation model and the relevant syntax, semantics, and phonology. In fact, Jackendoff's (2002) account leads to no distinction between routines for complex expressions and routines corresponding to single words. Hence the "routines" box in fact incorporates the lexicon as indicated in Pickering and Garrod (2004, Figure 3).

We argue that this representation may be fairly temporary, though we assume that it retains some activation independent of the activation of the syntax, semantics, and phonology. This can be used to explain, for example, the tendency to repeat both words and complex expressions during dialogue (e.g., Garrod and Anderson, 1987; Brennan and Clark, 1996; Tannen, 1989), and the greater tendency for syntactic repetition in the

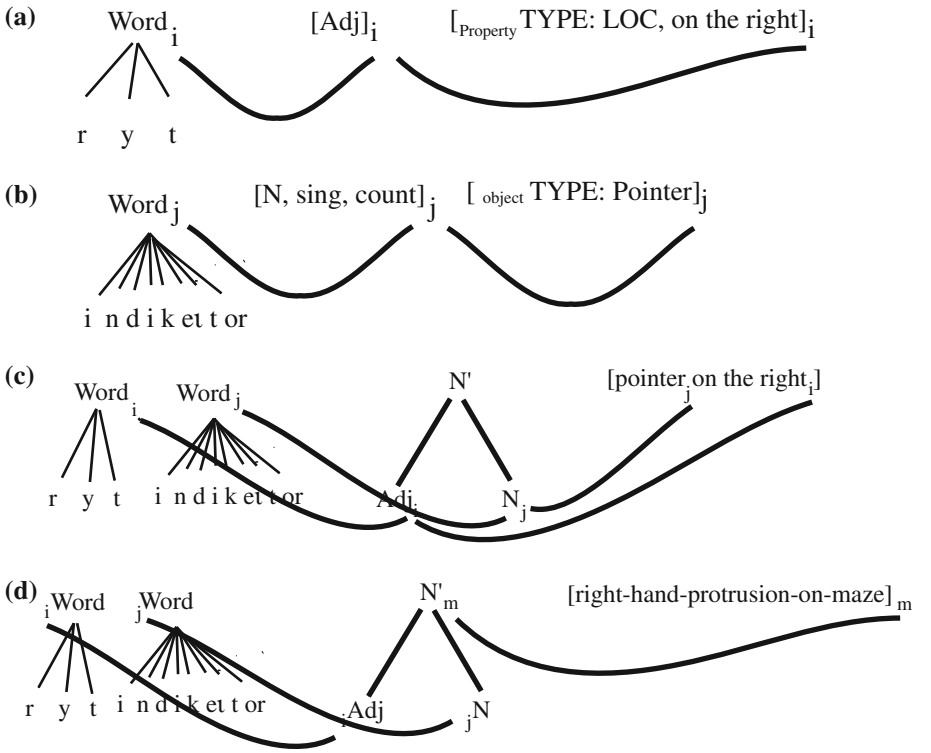

Figure 5. Lexicalisations in the spirit of Jackendoff (2002); (a) and (b) are lexicalisations of the words "right" and "indicator", (c) is the representation of the standard expression "right indicator", (d) is the representation of the routinized expression "right indicator" taken from Garrod and Anderson's (1987) maze-game corpus.

context of lexical or semantic repetition (Branigan et al., 2000; Cleland and Pickering, 2003; Schoonbaert et al., in press). Many of these temporary routines rapidly lose much of their activation. However, some expressions will become fixed for the length of the conversation (or beyond) (see Garrod and Doherty, 1994). On this account, routinization is a consequence of the way in which information is chunked in order that linguistic knowledge can be employed in the development and expression of the situation model.

4. The Psychological Processes Underlying Alignment

So far we have been discussing different aspects of the concept of alignment and how they are connected. We now turn to more empirical questions about how alignment takes place in practice, why interlocutors might decide not to align, and whether alignment can tell us anything about the content of novel contributions to dialogue.

4.1. Alignment as a mechanistic process

As we have argued, dialogue involves two interlocutors who have their own (individual) intentions about what they want to convey, but that the contributions follow from a combination of the intentions of speaker and addressee. This follows from the assumption that two interlocutors' contributions constitute a joint activity of dialogue, just as two dancers' movements constitute a joint activity of ballroom dancing.

How does joint activity relate to alignment? On one account, dialogue involves a process of negotiation, with alignment as the goal of that negotiation. This serves as part of what Clark (1996) refers to as the secondary channel, in which the interlocutors make contributions relating to the nature of the interaction itself (via collateral signals such as *um* to indicate difficulty).

For example, Brennan and Clark (1996) argue that when interlocutors employ the same expression to refer to a particular object (e.g., *pennyloafer* to refer to a picture of a particular kind of shoe), they enter into a tacit *conceptual pact* in which they agree to keep referring to the same object in the same way. Interlocutors align because certain "conversational moves" indicate acceptance of the pact. For instance, if *A* refers to a shoe as a *pennyloafer* 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 referential term, and (2) that both know (1). (In contrast, if *B* queried *A*'s expression, then the pact would not be formed.) Alignment is therefore the result of a process of negotiation that is specialized to dialogue and involves inference. In this respect, it suggests that dialogue is more complex than monologue, where the listener only needs to accept the speaker's contribution.

In contrast, the interactive-alignment account proposes that alignment is primitive. It is a form of imitation and drops out of the functional architecture of the system. People imitate each other in many different ways, from mouth movements in neonates (e.g., Meltzoff and Moore, 1977) to actions (e.g., Hommel et al., 2001) to aspects of social behaviour (e.g., Chartrand and Bargh, 1999). The neural basis for imitative behaviour seems to be linked to the neural basis for language use (Rizolatti and Arbib, 1998). In these accounts, imitation is an automatic, non-inferential process and is in some sense the default response. Generally, imitation does not appear to require any decision to act (see Hurley and Chater, 2005).

Imitation in language is similar. It occurs at many different linguistic levels, as we have discussed, and does not appear to be tied to awareness of the process of imitation. Whereas imitation of content words might sometimes be strategic, there is little evidence that people are aware of imitation of closed-class words or grammar (e.g., Branigan et al., 2000; Levelt and

Kelter, 1982). Even very young children perform syntactic imitation, as Brooks and Tomasello (1999) demonstrated for 2–3 year olds using passives. Perhaps most strikingly, it occurs very rapidly indeed. For example, Goldinger (1998) had participants shadow words, and found that the acoustic characteristics of the produced word tended to reflect those of the stimulus. Moreover, this imitation was greater when the words were produced immediately after the stimulus. Fowler et al., (2003) found that imitating a string of phonemes is almost as fast as making a simple reaction time judgment to the same stimulus. Fadiga, Craighero, Buccino, and Rizzolatti (2002) found that listeners activated appropriate muscles in the tongue while listening to speech (but not during non-speech). These findings make it highly likely that alignment is largely a mechanistic process.[1]

In Pickering and Garrod's (2004) terms, linguistic imitation and hence alignment occurs as a result of priming and parity of representation. For example, if the speaker utters *extreme right* the addressee activates the lexical items *extreme* and *right* during comprehension. This rise in activation does not decay rapidly, and the addressee is more likely to use *extreme right* subsequently (and is also more likely to understand *extreme right* rapidly.) This process of activation is therefore automatic and does not involve a conceptual pact between the interlocutors. Clearly, it is possible to select expressions in a more intentional way, in order to make a contribution most appropriate for interlocutors. For example, the tendency to design expressions that will be understood by a particular audience (e.g., experts vs. novices; Isaacs and Clark, 1987) may go against the tendency to imitate, though of course the expression used by one's interlocutor is a very good indication of what they are likely to understand.

A more complicated case occurs when one's interlocutor changes. Pickering and Garrod (2004) predict that the automatic tendency is for people to repeat expressions (for example), irrespective of whether they are responding to the same interlocutor or to another interlocutor, simply because the first use of the expression should cause it to be activated. However, it is of course sometimes the case that a new interlocutor is unlikely to understand a term that has specifically evolved in a conversation before that interlocutor arrived. In such cases, an alert speaker will go against her natural tendency to imitate and will employ a different term. In Pickering and Garrod's terms, any such tendency is a specific decision not to align, and involves processing effort (see Garrod and Pickering, in press, for a discussion of automaticity and effort in language processing). Brennan and Clark (1996) found such an effect of partner-specificity, with speakers using terms that would be understood by a new interlocutor. However, it is important to note that speakers often first employed the previous term to a new interlocutor, and only altered that term when the new

interlocutor provided feedback (p. 1491). This is a good example of the interactive repair mechanism, and occurs either because the speaker's use of the old term does not allow the new interlocutor to identify the right entity (in which case their situation models are not aligned) or because the old term is unnecessarily specific (in which case a change in term reduces processing effort).

The important point is that effects of partner specificity do not imply that interlocutors need employ complex reasoning whenever they produce an expression. Instead, they have a strong tendency to employ the form that they have just encountered (input-output coordination, in Garrod and Anderson's, 1987, terms). When resources allow, they modulate this tendency by paying some attention to their beliefs about what their interlocutor is likely to know. More regularly, they note when their interlocutor produces feedback (e.g., *eh*?) that indicates lack of alignment, and therefore reformulate in a way that may temporarily reduce alignment of non-situational knowledge in order to ensure alignment of situation models. We now sketch some implications of this mechanistic view of alignment for alignment and misalignment in dialogue.

4.2. Interactive repair and the externalising of inference

Consider again the situation depicted in Figure 4b where interlocutors *A* and *B* are both using the same name *Robert* to refer to different individuals in their respective situation models. How can this misalignment be remedied without *A* having to explicitly model what is in *B*'s mind? Pickering and Garrod (2004) proposed an interactive repair mechanism for such cases which only depends on interlocutors having access to their own situation models. As conversation proceeds, misalignment between *A*'s and *B*'s situation models will eventually lead one of them (e.g., *A*) to say something that conflicts with the other's knowledge of the individual represented in their model (i.e., *B*'s model). For example, imagine that *A*'s Robert has a sister, but *B*'s does not. At some point *A* says *I saw Robert yesterday with his sister*. So *B* updates her model to include Robert's sister and finds that her Robert does not have a sister. This then triggers her to repeat *Sister?* with a querying intonation. *A* can then reformulate his description (e.g., *I meant Robert Johnson*) and *B* can accept (e.g., *Ah, Robert JOHNSON*). This process normally continues until the models become realigned. Horton and Gerrig (2005) discussed a number of examples of this kind taken from corpora of telephone conversations. For example in (3), *A* introduces the referent with a pronoun *she* and then follows up later with the name *I-m-* (*I mean*) *Isabelle* as a check, presumably because *B* appears not to have understood.

3.

A: and um it- you know it's rea- it's it was really good and of course she teaches theology that was another thing
B: mm
A: I- m- I- Isabelle
B: oh that's great.

There are two important points about the interactive repair process. First, it does not necessarily require interlocutors to explicitly represent each other's situation models. In (3) *B*'s query and *A*'s response were formulated on the basis of what is in their own models at that point. In *B*'s case the formulation *mm* may have come from failing to resolve the pronoun *she* in *A*'s utterance in relation to what *B* has in her model. In *A*'s case it may have come from reformulating her description of the individual in her model. *A* may have realized that *B* has failed to understand (in other words, that *B* has not aligned with *A*), and that the difficulty relates to the reference to Isabelle with *she*. Without modelling *B*'s mental state, *A* can infer that her utterance may not pick out a unique individual, and so therefore *A* employs a simple strategy of being more specific. Obviously there are various ways to do this, but the most straightforward one is to mention Isabelle's name, because this is referentially fairly specific, is easy to access, and is comparatively quick to produce. Though *A* could tailor her reformulation to her specific knowledge about *B* (e.g., "the woman we both met at Godfrey's party"), this is not done here or in many other cases. In other words, the realignment process does not normally require either interlocutor to draw complex inferences about their partners' mental states. Second, the repair process is only available in dialogue. If *A* had produced the same statement in a letter to *B* (i.e., in monologue), then *B* could realise that *A*'s remark did not make sense and try to infer what woman *A* had in her mind. However, even if she could manage to guess correctly, she would have no way of verifying the inference. Only through the joint actions of dialogue is interactive repair possible. In effect, interactive repair externalizes the inference process.

4.3. Deciding not to align

The above evidence indicates that alignment tends to occur automatically. Hence the decision not to align requires additional effort. However interlocutors will sometimes choose not to align. Deciding not to align situation models (as in Figure 4b) may happen for complex reasons, as when a teacher cannot hope to convey the full complexity of a domain to a student. We shall ignore this case and instead focus on non-alignment of linguistic knowledge, as seen in Figure 4a.

Our underlying conceptualization of conversation is collaborative, in that we treat it as a "game of pure cooperation" (following Lewis, 1969), in which it is in both interlocutors' interest for it to succeed for both interlocutors. Even in an argument, the interlocutors need to interpret most of the expressions in the same way. However there are many cases in which one interlocutor attempts to impose her expression on the other interlocutor. An obvious case is when one interlocutor (say a teacher) corrects her partner's expression by an embedded repair, where the correction is embedded in the interlocutor's response (Jefferson, 1987). So long as the student now begins to use the teacher's term, alignment can be maintained at the linguistic (i.e., lexical) as well as situational level. The conversational game remains one of pure cooperation, but there is also a competitive game concerned with the choice of expressions (which the teacher tends to win).

In contrast, non-alignment with respect to linguistic knowledge does occur when interlocutors refer to the same entities but refuse to adopt the same terms (as in Figure 4a). Danet (1980) discussed the use of terms to refer to the result of pregnancy in an abortion trial. The prosecuting lawyer tended to employ terms like *baby* or *unborn child*, whereas the defendant tended to employ terms like *foetus*. It is clear that the different sides in the debate did not want to align on the same terms, because different terms would have had very different implications. For example, for the defendant to employ *baby* would have been taken to imply acceptance that the results of pregnancy were a person with the legal right not to be killed. To do this, the interlocutors have to model the implications of adopting a particular term for how they are likely to be perceived by the relevant audience, whether judge, jury, or the wider public who were exposed to the trial.

Without going into the full implications of this interesting situation, we note that the participants had a strong interest in not aligning on each others' terms. This aspect of the interaction is an adversarial conflict rather than part of a "game of pure collaboration", but they must also understand what each other is saying at the level of the situation model (so that the argument makes sense). In such cases, the effort needed to avoid alignment must have been considerable. In more mundane cases, interlocutors may on occasion try to impose their choice of expression on their interlocutor, when there is a conflict about who is right (e.g., name of a car), whose dialectical expression is to be employed (e.g., *lunch* vs. *dinner* to refer to the midday meal) or the political implications of a choice of expression (e.g., *benefit claimants* vs. *scroungers*). Attempting to impose one's term is effortful, and the "combatants" may hope that if they stick to their term, their interlocutor may give in and change theirs. What this suggests is that situational alignment in the absence of alignment of expressions and their interpretations is difficult to maintain.

A clear prediction is that interlocutors who use different terms to refer to the same things are not going to communicate so well. This is because they cannot take advantage of 'routinization', which depends upon the repeated use of the same expression to refer to the same thing(s) in the model. Not only does routinization facilitate the production and comprehension process, it also allows interlocutors to take advantage of a local principle of contrast (Clark, 1993). So long as an expression is a routine then alternative potential terms for that referent will be treated as introducing something new into the model. Being able to automatically identify new as opposed to old interpretations for any expression simplifies alignment. Notice that this contrast principle cannot work if speakers are already using different (non-routinized) terms to refer to the same thing.

4.4. Alignment and novel contributions to dialogue

In most (though by no means all) dialogues, one or both interlocutors have the goal of informing their partner about some state of affairs. The interactive-alignment account says relatively little about speakers' decisions about what to talk about, and instead focuses on how people discuss issues that they have decided to talk about. Predicting choice of topic seems to be as complex as predicting behaviour generally, and beyond the remit of a mechanistic psychology.

Having said that, Figure 3 does indicate a link between interlocutors' general knowledge, as discussed above. So encountering particular expressions would be likely to activate particular representations that would make it more likely for interlocutors to raise particular topics. Such aspects of people's behaviour can be affected by priming manipulations, as extensively demonstrated in the social-psychology literature (e.g., Greenwald and Banaji, 1995). In Danet's (1980) example, the choice of *baby* versus *foetus* might affect the representations that would be activated by the trial participants, which would in turn affect their subsequent contributions. For example, the prosecutor's use of *baby* might cause the defendant to be more like to modify his "pro-choice" stance in his subsequent contributions, and not merely to change his vocabulary. This might be because the word *baby* affects the addressee's situation model, or it might be because of a direct influence of the word onto the addressee's conception of the world.

Hence, if the interlocutors become well-aligned, making the same choices as each other at various linguistic levels, they may be more likely to focus on the same topics as each other. This can be seen as a weak "Whorfian" view, where choice of language shapes thought to some extent, and so people who speak like each other are thereby more likely to think like each other too. It remains to be seen whether such effects do occur and, if so, how strong they are.

5. Conclusions

This paper has characterized alignment within the framework of the interactive-alignment account of dialogue (Pickering and Garrod, 2004). It provides a perspective in which alignment of situation models and alignment of non-situational knowledge are treated as separate but complementary aspects of alignment. In addition, it proposes that dialogue should be interpreted in terms of informational alignment rather than information transfer. On the basis of these assumptions, it attempts to understand how alignment may come about during dialogue.

Notes

[1] These data suggest that alignment is closely related to prediction, with interlocutors constantly predicting what is likely to come next, at the level of expression, grammatical construction, or whatever. Many accounts of language comprehension assume that comprehension involves getting oneself into the right "position" in representational space, in effect predicting what is coming next to some extent (e.g., Tabor et al., 1997). This is compatible with some experimental evidence from monologue (e.g., Altmann and Kamide, 1999). On this account, alignment changes the predictions so that interlocutors will predict the same things as each other in various ways, with the listener predicting that the speaker will use words in the same way that the speaker has already done, not because of explicit modelling, but rather because the listener's previous utterances will have led to particular patterns of activation. Additionally, the parity between production and comprehension means that the listener's predictions about what the speaker will say will be the same as the listener's predictions of what he would say at that point if he were to interrupt.

References

Aijmer K. (1996) *Conversational Routines in English: Convention and Creativity.* Longman, London.

Altmann G. T. M., Kamide Y. (1999) Incremental Interpretation at Verbs: Restricting the Domain of Subsequent Reference. *Cognition*, 73, pp. 247–264.

Branigan H. P., Pickering M. J., Cleland A. A. (2000) Syntactic Coordination in Dialogue. *Cognition*, 75, pp. B13–B25.

Brennan S. E., Clark H. H. (1996) Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, pp. 1482–1493.

Brooks P., Tomasello M. (1999) Young Children Learn to Produce Passives with Nonce Verbs. *Developmental Psychology*, 35, pp. 29–44.

Chartrand T. L., Bargh J. A. (1999) The Chameleon Effect: The Perception-behavior Link and Social Interaction. *Journal of Personality and Social Psychology*, 76, 893–910.

Clark E. V. (1993) *The Lexicon in Acquisition*. Cambridge University Press, Cambridge.

Clark H. H. (1996) *Using Language*. Cambridge University Press, Cambridge.

Clark H. H., Marshall C. R. (1981) Definite Reference and Mutual Knowledge. In Joshi A., Webber B., Sag I. A. (eds.), *Elements of Discourse Understanding*. Cambridge University Press.

Clark H. H., Schaefer E. F. (1987) Concealing One's Meaning from Overhearers. *Journal of Memory and Language*, 26, pp. 209–225.

Clark H. H., Wilkes-Gibbs D. (1986) Referring as a Collaborative Process. *Cognition*, 22, pp. 1–39.

Cleland A. A., 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, pp. 214–230.

Danet B. (1980) 'Baby' or 'Fetus?' Language and the Construction of Reality in a Manslaughter Trial. *Semiotica*, 32, pp. 187–219.

Ericsson K. A., Kintsch W. (1995) Long-term Working Memory. *Psychological_ Review*, 102, pp. 211–245.

Fadiga L., Craighero G., Buccino G., Rizzolatti G. (2002) Speech Listening Specifically Modulates the Excitability of Tongue Muscles: A TMS study. *European Journal of Neuroscience*, 15, pp. 399–402.

Fodor J. A., Bever T. G., Garrett M. F. (1974) *The Psychology of Language*. McGraw Hill, New York.

Fowler C. A., Brown J., Sabadini L., Weihing J. (2003) Rapid Access to Speech Gestures in perception: Evidence from Choice and Simple Response Time Tasks. *Journal of Memory and Language*. 49, pp. 296–314.

Garrod S., Anderson A. (1987) Saying What You Mean in Dialogue: A Study in Conceptual and Semantic Co-ordination. *Cognition*, 27, pp. 181–218.

Garrod S., Doherty G. (1994) Conversation, Coordination and Convention: An Empirical Investigation of how Groups Establish Linguistic Conventions. *Cognition*, 53, pp. 181–215.

Garrod S., Pickering M. J. (2004) Why is Conversation So Easy? *Trends in Cognitive Sciences*, 8, pp. 8–11.

Garrod S., Pickering M. J. (in press) Automaticity in Language Production in Monologue and Dialogue. In Wheeldon L., Meyer A. (eds.), *Language and Executive Processes*. Psychology Press.

Goldinger S. A. (1998) Echoes of echoes? An Episodic Theory of Lexical Access. *Psychological Review*, 105, pp. 251–279.

Greenwald A. G., Banaji M. R. (1995) Implicit Social Cognition: Attitudes, Self Esteem, and Stereotypes. *Psychological Review*, 102, pp. 4–27.

Gries S. T. (2005) Syntactic Priming: A Corpus-based Approach. *Journal of Psycholinguistic Research*, 34, pp. 365–399.

Hartsuiker R. J., Pickering M. J., Veltkamp E. (2004) Is Syntax Separate or Shared Between Languages? Crosslinguistic Syntactic Priming in Spanish/English Bilinguals. *Psychological Science*, 15, pp. 409–414.

Haywood S. L., Pickering M. J., Branigan H. P. (2005) Do Speakers Avoid Ambiguities During Dialogue? *Psychological Science*, 16, pp. 362–366.

Hommel B., Müsseler J., Aschersleben G., Prinz W. (2001) The Theory of Event Coding (TEC): A Framework for Perception and Action Planning. *Behavioral and Brain Sciences*, 24, pp. 849–937.

Horton W. S., Keysar B. (1996) When Do Speakers Take into Account Common Ground? *Cognition*, 59, pp. 91–117.

Horton W. S., Gerrig R. J. (2005) Conversational Common Ground and Memory Processes in Language Production. *Discourse Processes*, 40, pp. 1–35

Hurley S., Chater N. (eds.) (2005) *Perspectives on Imitation*, 2 Vols. MIT Press, Cambridge, MA.

Isaacs E. A., Clark H. H. (1987) References in Conversations between Experts and Novices. *Journal of Experimental Psychology: General*, 116, pp. 26–37.

Jackendoff R. (1999) Parallel Constraint-based Generative Theories of Language. *Trends in Cognitive Sciences*, 3, pp. 393–400.

Jackendoff R. (2002) *Foundations of Language.* Oxford University Press, Oxford.

Jefferson G. (1987) On Exposed and Embedded Corrections in Conversation. In Button G., Lee J. R. E. (eds.), *Talk and Social Organisation*, pp. 86–100. Multilingual Matters, Clevedon.

Johnson-Laird P. N. (1983) *Mental models: Toward a Cognitive Science of Language, Inference and Consciousness.* Harvard University Press, Cambridge, MA.

Kuiper K. (1996) *Smooth Talkers: The Linguistic Performance of Auctioneers and Sportscasters.* Erlbaum, Mahwah, NJ.

Levelt W. J. M. (1989) *Speaking: From Intention to Articulation.* MIT Press, Cambridge, MA.

Levelt W. J. M., Kelter S. (1982) Surface Form and Memory in Question Answering. *Cognitive Psychology*, 14, pp. 78–106.

Levinson S. C. (2003) *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge University Press, Cambridge.

Lewis D. K. (1969) *Convention: A Philosophical Study*. Harvard University Press, Cambridge, MA.

Meltzoff A., Moore M.K. (1977) Imitation of Facial and Manual Gestures by Human Neonates. *Science*, 198, pp. 75–78.

Nunberg G., Sag I. A., Wasow T. (1994) Idioms. *Language*, 70, pp. 491–538.

Pickering M.J., Garrod S. (2004) Toward a Mechanistic Psychology of Dialogue. *Behavioral and Brain Sciences*, 27, pp. 169–225.

Pickering M.J., Garrod S. (2005) Establishing and Using Routines During Dialogue: Implications for Psychology and Linguistics. In Cutler A. (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones* pp. 85–101. Erlbaum, Mahwah, NJ.

Rizzolatti G., Arbib M. A., (1998) Language within Our Grasp. *Trends in Neurosciences*, 21, pp. 188–194.

Sanford A. J., Garrod S. C. (1981) *Understanding Written language*. Wiley, Chichester, UK.

Schober M. F. (1993). Spatial Perspective Taking in Conversation. *Cognition*, 47, pp. 1–24.

Schober M. F. (1995) Speakers, Addressees, and Frames of Reference: Whose Effort is Minimized in Conversations about Locations? *Discourse Processes*, 20, pp. 219–247.

Schoonbaert S., Hartsuiker R. J., Pickering M. J. (in press) The Representation of Lexical and Syntactic Information in bilinguals: Evidence from Syntactic Priming. *Journal of Memory and Language*.

Szmrecsanyi B. (2005) Creatures of Habit: A Corpus-Linguistic Analysis of Persistence in Spoken English, *Corpus Linguistics and Linguistic Theory*, 1, pp. 113–149.

Tannen D. (1989) *Talking voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge University Press, Cambridge.

Tabor W., Juliano C., Tanenhaus M. K. (1997) Parsing in a Dynamical System: An Attractor-based Account of the Interaction of Lexical and Structural Constraints in Sentence Processing. *Language and Cognitive Processes*, 12, pp. 211–271.

Watson M. E., Pickering, M. J., Branigan H. P. (2004) Alignment of Reference Frames in Dialogue. *Proceedings of the 26th Cognitive Science Meeting*. Erlbaum, Mahwah, NJ.

Zwaan R. A., Radvansky G. A. (1998) Situation Models in Language Comprehension and Memory. *Psychological Bulletin*, 123, pp. 162–185.