Title	Reference frame selection and representation in dialogue and monologue
Author	Watson, Matthew E.
Qualification	PhD
Year	2006

Reference frame selection and representation in dialogue and monologue

Matthew E. Watson

PhD
University of Edinburgh
2006

Table of Contents

Figures and Tables

Declaration

This thesis has been composed by myself, and the research presented herein is my own. No portion of the work has been submitted for any other degree or professional qualification.

Matthew Watson

Acknowledgements

I would like to thank my two supervisors, Martin Pickering and Holly Branigan, who have been nothing short of excellent and from whose supervision I have benefited greatly. I am also grateful to Laura Carlson for allowing me to work with her in Notre Dame for three months, during which time I developed the work presented in Chapter 5. I would also like to thank Aaron Ashley and Pat Hill from the spatial cognition lab at Notre Dame. Special mentions go to Manon Jones who has played confederate for me as well as supporting me a great deal. Many thanks are due to my confederates Laura Tully, Clare Huxley, Ciara Catchpole, and Claudine Raffray. Finally I would like to thank my partner for her support and patience during my write up and my parents who have also been supportive.

Abstract

This thesis is an investigation into the use and representation of reference frames. A reference frame is an axial co-ordinate system that is represented by the setting of at least four parameters: origin, orientation, direction and scale (Logan & Sadler, 1996). The reference frame parses space into different regions around an object (*reference object*) so that the location of other objects (*figure objects*) can be described in relation to the reference object. There are at least three different types of reference frame recognised in the literature (e.g. Levinson, 2003; Logan & Sadler, 1996; Miller & Johnson-Laird, 1976), the absolute reference frame, based upon directional features of the environment (e.g. gravity), the intrinsic reference frame, based upon the intrinsic direction of the reference object and the relative reference frame, based upon the intrinsic sides of a viewer.

This thesis is split into three separate sections each investigating how it is that reference frames are employed to allow us to describe the location of objects. The first section investigates the use of reference frames in dialogue. There has been little work investigating how interlocutors communicate object location to each other. Research investigating the use of non-spatial language in dialogue has shown that interlocutors come to speak in a similar fashion to one another (e.g. Branigan, Pickering & Cleland, 2000; Garrod & Anderson, 1987). Pickering and Garrod (2004) have argued that this is evidence that interlocutors align representations during dialogue. The first section investigates whether or not interlocutors align reference frames. If a speaker uses an intrinsic reference frame is an addressee more likely to then use an intrinsic reference frame? The results show that interlocutors do align reference frames and this is interpreted in terms of the interactive alignment model of dialogue (Pickering & Garrod, 2004).

The second part of the thesis uses a dialogue paradigm, and the fact that interlocutors are more likely to use the same reference frame as just used by a partner than an alternative, to investigate different taxonomies of reference frames. The experiments investigate the different predictions made by Levinson's (1996, 2003) taxonomy of reference frames and the Traditional taxonomy of reference frames (Miller & Johnson-Laird 1976). The results support the Traditional taxonomy of reference frames over Levinson's taxonomy, but future research is suggested to help separate

the contribution of different aspects of reference frames to the alignment effect in dialogue.

The final part of the thesis is an investigation into how the distance parameter is set. These experiments show that a functional interaction between the figure and the reference object leads to the distance parameter being set as shorter than if the two objects are not interacting. In addition the results show that apprehending a spatial relationship between a figure and reference object requires the use of attention. In addition the distance parameter is not set according to the reference frame selected, but is set according to other contextual factors such as the interaction between the figure and the reference object (Carlson & Covey, 2005).

This thesis shows that interlocutors align reference frames during dialogue. It also shows that spatial language can be investigated using dialogue paradigms. The results suggest that future research into spatial language should use dialogue methods as this allows a more naturalistic investigation than traditional monologue contexts.

Chapter 1: Introduction

1.0. Chapter Overview

This chapter presents a brief overview of the questions that will be addressed in this thesis. It describes the background and rationale for the Experiments presented.

1.1. The importance of Spatial Language

One of the features of language which humans excel at is the ability to describe the relationship between two or more entities. In particular, humans are able to describe the spatial relationship between two objects, for example, the location of an entity in relation to a second entity within a landscape. This is important because in addition to being linguistic, humans are also, primarily, a species reliant on vision. Spatial language allows humans to describe the whole layout of a scene to a non-observer and indicate not only which objects are present in the scene, but also the relationships between those objects in 3-dimensional space. The importance of spatial language is highlighted further in non-spatial domains. For abstract concepts a spatial metaphor is often used, for example, in social relations (e.g. upper class, moving up the social ladder), the concept of time is also dependent upon the language used to describe space (Boroditsky, 2000). Spatial language requires a mapping from a visual or spatial representation onto a conceptual representation, and then onto a linguistic representation. Similarly, in order to understand spatial language requires constructing a spatial representation from a linguistic representation. The process is a mapping from a co-ordinate spatial relationship in a spatial representation, which encodes precise metric relationships between objects, to categorical spatial relationships in a conceptual representation, which are coarse and imprecise. Categorical and co-ordinate encodings of spatial relationships have been shown to dissociate in patients (Kosslyn, Koenig, Barrett, Cave, Tang, & Gabrieli, 1989) and are represented differently in neural network models (Kosslyn, Chabris, Marsolek, & Koenig, 1993). The mapping between a spatial representation and linguistic representation is not a trivial affair: Several steps are required in order to achieve a correct description or understand a description correctly. These include; spatial indexing, reference frame adjustment, spatial template alignment and computing the goodness of fit (Logan & Sadler, 1996). It is the goal of spatial language research to elucidate the details of these steps in computation and describe fully how people are able to talk about object location.

1.2. Perspective Taking

The primary purpose of language is communication with others and the natural setting for it is a dialogue context, one in which there are multiple turns with multiple partners. When a person describes the location of an object it is probably to communicate the location of the object to a second person, rather than a statement to himself. As such it is important to explain how it is that people are able to communicate and understand one another and how language is produced and understood as a communicative tool. Much work on language has been carried out in a monologue context (e.g. Frazier, & Rayner, 1982, Glaser & Glaser, 1989; Levelt, Schriefers, Norberg, Meyer, Pechmann, & Havinga, 1991; Schriefers, Meyer & Levelt, 1990, Pickering & Traxler, 1998), investigating how people produce and comprehend language on their own in isolated contexts. This is partly because it is much easier to conduct and control variables in monologue than in dialogue and partly because the studies looked to remove the effects of context and other such extraneous variables so as to get a view as to how language operated by default (cf. Bock, 1996). However, this gives an unbalanced explanation of language because language rarely occurs without context. Much of this thesis is devoted to an investigation of spatial language in a dialogue context with an aim to discover how it is that we describe the location of objects to each other.

Speakers have a choice of how to describe the location of an object; they can choose to describe it from their own personal perspective or they can choose to describe it from an external perspective, such as an addressee's perspective, the perspective of a third person, or in terms of the intrinsic axes of another object. The choice of how to describe the location of an object is dependent upon the use of a reference frame. A reference frame is a representation that parses space around an object into regions corresponding to spatial terms (e.g. *above*). When speakers select a perspective to describe a scene from they are also selecting a reference frame. If speakers describe the location of an object using their own perspective (e.g. *the ball is to my left of the car*), they are using a *relative* reference frame. If speakers describe the location of an object in terms of the directional axes of an object (e.g. *the ball is in front of the car*; at the car's headlights), they are using an *intrinsic* reference frame. This presents a problem for addressees because they have to decide which reference frame the speaker was using in order to correctly interpret the description. Chapter 3 presents a series of experiments that investigate whether interlocutors use the same reference frames as each other. This would enable them to co-ordinate the choice of perspective and it would provide a mechanism for how ambiguity may be avoided. In these experiments a confederate described the location of a dot (figure object) in relation to a second object (reference object) using either an intrinsic or a relative reference frame. The participants then had to choose which of two pictures matched the confederate's description. Following this the participants described the location of a dot in relation to a reference object. The investigation was concerned with whether or not the participants used the same reference frame as they had just heard the confederate use. The results showed that interlocutors do align reference frames during dialogue.

1.3. Reference Frame Representation

Reference frames are axial co-ordinate systems that define an origin, orientation, direction and distance (Logan & Sadler, 1996). However, there has been some disagreement about how to divide reference frames up into different types. One dispute focuses upon whether or not sentences such as *the ball is in front of me* use the same reference frame as sentences such as *the ball is to the left of the car* (from my perspective), in which case both use a *deictic* (egocentric) reference frame (Miller & Johnson-Laird, 1976), or whether or not sentences such as *the ball is in front of me* use the same reference frame as sentences of the form *the ball is in front of the car* (at the car's headlights), in which case both use an *intrinsic* reference frame (Levinson, 1996, 2003). In Chapter 4 a series of experiments is presented that used a confederate priming paradigm to investigate the representation of reference frames. In these experiments a confederate described the location of an object using a sentence of the form *the X is in front of us*. This was so the participants could choose which of two pictures matched the description. Participants then had to describe the location of an object which could either be described using a *relative/deictic* reference frame or an *intrinsic* reference frame. The results of the experiments in Chapter 3 showed that interlocutors align reference frames and so participants should be more likely to use the reference frame to describe the object's location which is used in the sentence *the X is in front of us*. Participants were more likely to use a relative/deictic reference frame after hearing the sentence *the X is in front of us*, suggesting that this sentence also uses a relative/deictic reference frame.

1.4. The Distance Parameter

The distance parameter defines the distance between two objects in a spatial relationship. Very little work has investigated how this parameter is set. However, it is known that the size of the objects affects the setting of this parameter (Carlson & Covey, 2005). For example *the planet is near the Earth* sets the distance parameter as larger than *the mouse is near the cheese*. In Chapter 5 three experiments are presented that investigated the setting of the distance parameter when the figure and reference objects are functionally related (e.g. a *light bulb* and a *lamp*) compared to when the figure and reference object were not functionally related (e.g. a *light bulb* and a *pear*) and whether or not the distance parameter is set independently of the other parameters that define a reference frame. In Experiment 8 participants were asked to imagine two objects (which could be functionally related or unrelated) in a spatial relationship (*above* or *below*), and then report the distance apart that they imagined the two objects to be. The results suggest that the distance parameter is set as smaller when the two objects are functionally related. Experiments 9 and 10 used a visual search paradigm to investigate the use of attention in apprehending a spatial relationship between two letters. Participants had to decide whether or not two letters in a given spatial relationship, at a given distance apart were present amongst a number of distracters. The results of these experiments show that using distance to apprehend a spatial relationship does require the use of attention and that the distance parameter is set independently of the reference frame used.

Chapter 2: Literature Review

2.0. Chapter Overview

In this chapter I give a review of the literature from two separate areas. Firstly, I review the research on spatial language and in particular focus upon the use of projective spatial terms (e.g. *above*) and the reference frames that are required to use projective terms. I describe the different theories of reference frames and the different types of reference frames. Then also discuss research that has investigated the factors that influence the selection of reference frames for production and comprehension and the different features of reference frame representation. Secondly, I review research on dialogue. In this section, I discuss theories of common ground before moving on to describe the interactive alignment model of dialogue (Pickering & Garrod, 2004) and evidence for alignment effects between interlocutors. Finally, I discuss issues of partner specific effects in comprehension and production. In a final section I discuss the small amount of work that has been conducted investigating the use of spatial language in dialogue.

2.1. Spatial Language

Spatial language is any word or sentence that describes the location or path of an object. In English, much of spatial language is encoded by the preposition word class, although some spatial relationships are encoded by verbs, for example, the verb *hanging* denotes a certain spatial relationship between objects. Prepositions are a closed class lexical category because there are relatively few of them, (approximately 100) and it is not possible to generate any new additions to the class. This is in contrast to nouns, which form an open class of words. There are at least 10,000 nouns in the average English speaking adult's vocabulary (Landau & Jackendoff, 1993). It is also possible to create new nouns.

There has been a large volume of research on how it is that concrete nouns are understood. At a coarse level of analysis the phonology and syntax of a word are mapped onto a conceptual representation which gives the word its meaning (e.g. Dell, 1989; Levelt, Roelofs and Meyer, 1999), although other theories argue for only two levels where the syntax and semantics are included in the same level of representation (Caramazza, 1997). Taken together the three representations form the word's entry in the mental lexicon (Jackendoff, 2002). With regards to prepositions a phonological representation and syntactic representation are also mapped onto a level of representation which forms the meaning of the word, but for this word class some form of spatial representation is required in addition to a conceptual representation.

Much research suggests that there is a representation that encodes the location of objects which can receive input both perceptually and linguistically (Bryant, 1997, Hayward & Tarr, 1995; Jackendoff, 1996), although there is also work suggesting that there are separate linguistic and perceptual representations of space (Crawford, Regier, & Huttenlocher, 2000; Huttenlocher et al, 2004). According to Landau & Jackendoff (1993, see also Jackendoff 1987, 1996, 2002) the spatial representation is an amodal representation, rather than being specifically associated with the visual modality. The reason for positing an amodal representational system is that it allows for spatial information to be received from the other perceptual modalities, for example the haptic system when we feel the shape of an object, or the auditory system when we hear the sound of an object. Similarly, it also allows for a spatial representation to be built up from linguistic information, when we read or hear a discourse that describes the location of objects we are also able to build up a spatial representation of that scene.

The richness of information which is passed onto the spatial representation differs depending upon which modality perceives the information. The visual system provides the most accurate and analogous information in terms of correspondence to the real world, although it is possible to trick the system as in the case of visual illusions. The auditory system can provide information about direction, distance and possibly also the environment one is in, but can provide little information regarding the relative positions or absolute distance of objects. Spatial information from a linguistic input is mostly less rich than that from the visual modality. Language encodes spatial relations between objects in a categorical fashion whereas vision can encode space in a metric nature with precise relationships between objects conveyed (Kosslyn et al, 1989). For example, an object can be described as to the left of a second object, but this does not provide exact relationships; is the object exactly to the left or slightly in front of and to the left? Similarly, distance is not easily conveyed accurately in spatial language, in English things are described as near, far and possibly next to (other languages divide distances into different segments; Bowerman, 1996).

2.1.1. Figure and Reference objects

The location of an object can be specified by describing its location in relation to a second object or place. The object that is located is termed the *figure object* (trajector), the object that it is related to is termed the *reference object* (also the landmark, ground or relatum). The location of a figure object is described by defining a search domain somewhere around the reference object. The description of the location of an object is mostly given in order that it may be understood by a second person. Experiments investigating reference frame selection show which reference frames people prefer to use and which factors influence the selection process. The idea of the locative description is that it locates the figure object, the location of which is presumably unknown to the addressee, in relation to the reference object, the location of which is known. This is reflected in the asymmetry between figure and reference objects. Typically figure objects are smaller, less mobile, and less salient than reference objects (Talmy, 1983). For example it is acceptable to say *the bicycle is next to the house*, but it sounds unusual to say *the house is next to the bicycle*. The idea of the locative utterance is to narrow down the search domain to a specified region around the reference object.

Initial investigations into how space was divided up into different regions assumed that there would be a universal method of doing this common to all languages (Miller & Johnson-Laird, 1976). However, it is now clear that this is not the case. For example, English has the prepositions *in* and *on*, which broadly describe a containment relationship and a contact with support relationship respectively. Dutch, however, divides contact with support into two different categories, *op* meaning support on a horizontal surface (e.g. a cup on a table), and *aan* meaning support on a vertical surface (e.g. a picture on a wall) (Bowerman, 1996). Korean uses no relational terms analogous to *in* and *on*, but instead uses a division which cross cuts both terms and divides space into *kkita* tight fitting (e.g. a ring on a finger and a peg in the ground) and *nehta* loose fitting (e.g. a hoop on a pole and a marble in a bowl) (Munnich, Landau & Dosher, 2001). At 9 months children are sensitive to all distinctions in how space can be parsed, but by the age of 14 months the child is only sensitive the distinctions made in their native tongue (McDonough, Choi & Mandler 2003).

2.2. Spatial language taxonomy

There are several different types of spatial language that need to be analytically distinguished (Levinson, 2003). Firstly, there are the spatial terms which Garnham (1989) describes as '*basic*' or '*deixis*'. These types of spatial language refer to descriptions which specify a region in relation to the speech act, for example, *The ball is here,* or *It's over there.*

Second, there are *topological* descriptions. These are descriptions which describe some continuity between the figure object and the reference object. For example, the figure object is contained within the reference object or is supported by the reference object. In English this type of spatial description includes the prepositions *in, on, between, next to,* amongst others. Levinson (2003) also argues for a separate class of toponymic descriptions that describe the location of a figure object in terms of a named place where it is situated (e.g. *he is in America).* The class of spatial descriptions described above, Levinson (1996) terms as *coincident* because there is no angular specification given to define the location of the figure object.

The remaining class of spatial descriptions are the co-ordinate descriptions. These descriptions specify the location of the figure object by using a co-ordinate system to define where the figure object is in relation to the reference object. This class of spatial descriptions are projective descriptions because the figure object is located in a region of space that is projected out from the reference object. In English, this includes spatial terms such as *above, below, front,* and *behind.*

2.3. Reference frame taxonomy

A reference frame is an axial co-ordinate system that has a number of parameters (Logan & Sadler, 1996): *origin*, *orientation*, *direction,* and *distance* (originally Logan & Sadler (1996) used the term scale, but this implies a parameter with discrete intervals and so Carlson (2003) has replaced it with the term *distance*). There has been much discussion and debate over how to define the different types of reference frame that are employed in spatial language. The majority of researchers identify three types of reference frame: *absolute, intrinsic* and a *viewer* or *egocentric* reference frame (e.g. Logan & Sadler 1996, Carlson-Radvansky & Jiang 1998, Levinson 2003). However, Jackendoff (1996) suggests the possibility of there being eight different reference frames, although he groups these into two broad groups, *intrinsic* and *environmental*, that are analogous to the intrinsic and absolute reference frames suggested in the former tripartite classification systems (Jackendoff includes viewer centred reference frames as one type of environmental reference frame).

2.3.1. The Traditional taxonomy of reference frames

The three types of reference frames traditionally identified in psycholinguistics are the *absolute* (environment centred) reference frame, *intrinsic* (object centred) reference frame, and the *deictic* (egocentric) reference frame (e.g. Miller & Johnson-Laird 1976, Carlson-Radvansky & Logan 1994, Carlson-Radvansky & Jiang, 1998). I shall term this the *Traditional* taxonomy of reference frames.

The *absolute* reference frame uses fixed features of the environment, such as the points of the compass or gravity (or to directions such as downwind or inland in some cultures ;Levinson, 2003). In Figure 1, the dot can be described as *above the chair*, using an absolute reference frame. The absolute reference frame maintains spatial relations independent of the viewpoint of an observer and the reference object's orientation (i.e., which way the chair is facing).

When using the *intrinsic* reference frame the position of the figure in relation to the reference object is interpreted with respect to the actual orientation of the reference object. For example, the intrinsic use of *above* is (roughly) nearer to the top of the object than to any other part of it. As a chair has a top, the dot in Figure 1 can be described as *above the chair* using this reference frame. The terms used to refer to the position of the figure can be based on many different characteristics of the reference object. For example, front can be defined in terms of the position of the face (in people) or direction of movement (for electric trains), or by the direction of the street (for a house). In contrast, a stationary symmetrical reference object (e.g., a ball) cannot be used intrinsically. The determination of the left-right axis based upon the direction of the front-back and top-bottom axis. There are two different ways this can be done (Herskovitz, 1986, Miller & Johnson-Laird, 1976). The speaker can either imagine themselves inside the reference object. For example, if the reference object is a chair they imagine themselves sitting in the chair. If this is the case then the chair inherits the left-right dimensions of the person by analogy, so that a chair facing a person will have reversed left and right axes, just as a person facing a person does. In contrast, a speaker can imagine him or herself outside of the reference object, for example facing a desk, in which case the intrinsic right side of the desk is closest to the speaker's right hand. Notice that rotating the reference object can affect the appropriate description using this frame, but is independent of the viewpoint of an observer.

The *deictic* reference frame is an egocentric reference frame (it can sometimes also be centred on an addressee; Levelt, 1989). The origin of the reference frame is positioned upon the speaker and the directions of the axes are determined based upon the directional axes of the speaker. If you imagine yourself as the person in Figure 1 then the dot would be described as *above the chair*.

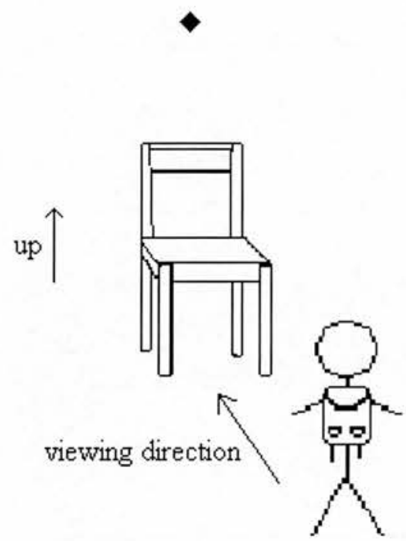


Figure 1: The vertical axis of all three reference frames is aligned; therefore the dot can be described as *above* the chair according to all reference frames.

In Figure 1, all three of the reference frames are conflated and so using any of them results in describing the dot as above the chair. However, there are occasions when the reference frames are dissociated and allow different descriptions of the same scene. In Figure 2 the same figure and reference object are used as in Figure 1, but this time all three reference frames have been dissociated. Now the dot is *above* the chair according to an absolute reference frame because the dot is higher in the plane of gravity than the chair. The dot is *left* of the chair according to a deictic reference frame because the dot is in the region to the left of the chair according to the person's axes (again assuming the position of the person in the scene). The dot is *below* the chair according to an intrinsic reference frame because the dot is located in the region that extends from the bottom of the chair.

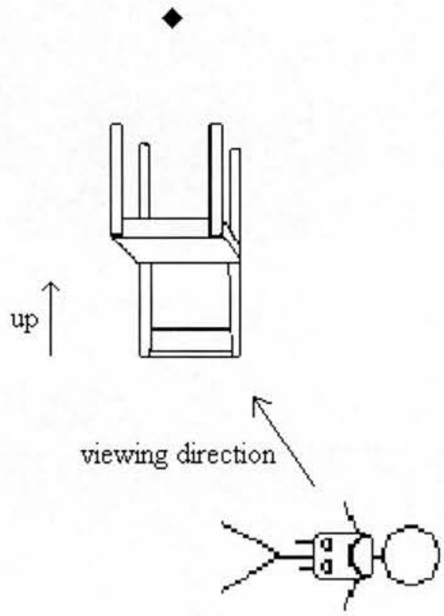


Figure 2. All three reference frames are dissociated: The dot is above the chair (absolute), below the chair (intrinsic) and left of the chair (deictic).

The Traditional taxonomy is not without problems, which in part has led to the classification of reference frames proposed by Levinson (1996, 2003). In particular it is not clear what the deictic reference frame applies to. Does it solely apply when the origin of the reference frame is the speaker of a spatial description? In which case phrases such as *the ball is in front of you* use an intrinsic reference frame. Does it apply when the origin of the reference frame is placed upon any person involved in the conversation? In which case, phrases that use an addressee centred perspective (e.g. *the ball is in front of you*), do use a deictic reference frame. Or does it apply when the origin of the reference frame is situated upon any person? In which case phrases such as *the ball is in front of John* use a deictic reference frame. The Traditional taxonomy of reference frames affords a special status to egocentric reference frames, giving them a reference frame category of their own (the deictic reference frame). However, it remains an open question as to whether or not addressee centred or third person centred reference frames should also be described as using a deictic reference frame.

2.3.2. Levinson's Taxonomy of Reference Frames

The Traditional taxonomic system of reference frames has been used to produce many fruitful results in spatial language and has often been adopted by researchers in psycholinguistics (e.g. Carlson-Radvansky & Irwin, 1993, Carlson-Radvansky & Jiang, 1998). Levinson (1996, 2003), however, disputes the Traditional taxonomy and suggests an alternative taxonomic system that cross-cuts the boundaries imposed by the Traditional taxonomy. Levinson (2003) maintains a tripartite taxonomy, but proposes a system where what would, according to the Traditional system, be categorised as a deictic reference frame may be categorised as an intrinsic reference frame. In opposition to the above system he terms his different reference frames as *absolute*, *intrinsic* and *relative*. Levinson's use of the absolute reference frame is the same as that of the traditional taxonomy, the difference lies in the use of intrinsic and deictic reference frames. The distinction between the two systems is exemplified in the following sentences:

1) The ball is to the left of the car (from the speaker's perspective).

2) The ball is to the left of me (from the speaker's perspective).

3) The ball is to the left of you (from the addressee's perspective).

4) The ball is to your left of the car.

5) The ball is in front of the car (at the car's headlights).

The traditional account groups together sentences 1 and 2 as using a *deictic* reference frame because in both sentences the directional axes of the reference frame are defined by the ego. Therefore sentences 3, 4 and 5 are classified as using an intrinsic reference frame because the directional axes of the reference frame are defined by an external object. Levinson argues that this grouping is incorrect and that 'clearly' sentences 2, 3 and 5 belong together and sentences 1 and 4 belong together. According to Levinson these groupings are conceptually the same because they have the same argument structure. Levinson therefore argues that sentences 2, 3, and 5 use an intrinsic reference frame, whilst sentences 1 and 4 use a relative reference frame. The use of the term relative is because the location of the figure object is relative to the viewpoint of an observer. The viewpoint can be the ego, an addressee or even a third person. The key conceptual difference between the reference frames is the number of arguments in the locative description; intrinsic reference frames form *binary* relations where there is a figure and reference object with the figure object being located in relation to some part of the reference object. Relative reference frames form *ternary* arguments where there is a figure object, a reference object and a viewpoint which the origin of the reference frame is situated upon and which determines the direction of the co-ordinate system.

Focusing on the argument structure of the spatial relations is a fundamentally different way of categorising reference frames from that used in the traditional taxonomy. The Traditional taxonomy is based upon the location of the origin and how the directions of the axes of the reference frame are determined. If the origin is situated upon, and the direction determined by, the ego then it is a deictic reference frame; if it is situated upon a second person or an object then it is an intrinsic reference frame. According to Levinson (2003) linguistic reference frames are defined with respect to the argument structure of the linguistic expression. If the expression is a binary argument then it is an intrinsic reference frame (other factors contribute to it being absolute reference frame); if the expression is a ternary argument then it is a relative reference frame.

There are also other reasons why Levinson argues for a taxonomy based upon argument structure over that of the traditional origin based distinction. Reference frames have certain logical properties (Levelt, 1984, 1996) and only when they are grouped according to Levinson's definition do those of the same type possess the same logical properties. The absolute and intrinsic reference frames both define binary relationships, but these are distinguished by the spatial inferences that each allow. The absolute reference frame supports converseness of relations. For example, if a ball is north of a chair then the chair is south of the ball. The absolute reference frame also supports transitivity so that if the ball is north of the chair and the table is north of the ball then the table is north of the chair. The intrinsic reference frame supports neither of these and allows few spatial inferences to be drawn. In Figure 3 the television is in front of the binoculars according to the binoculars' intrinsic reference frame, but the binoculars are left of the television according to the television's intrinsic reference frame, and this violates conversness. The intrinsic reference frame also fails to support inferences of transitivity. In Figure 3 the telephone is in front of the television and the television is in front of the binoculars, but the telephone is not in front of the binoculars. If the viewpoint upon which a relative reference frame is based remains constant then the relative reference frame does support inferences of transitivity and conversness.

These logical commonalities, Levinson argues, add further weight to his claim for adopting his reference frame definitions. If the traditional taxonomy is adopted then reference frames that are of the same type can possess different logical properties. For example, both sentences *The ball is in front of me* and T*he ball is in front of the tree* are deictic according to the traditional theory (they are intrinsic and relative respectively according to Levinson); however, only the second possesses the logical property of transitivity in that the tree is necessarily also behind the ball.

Levinson notes other properties of reference frames that are specific to each of his classifications, namely the properties each shows when part of or the entire scene is rotated. The distinguishing feature is whether or not the description of the scene remains the same when the *viewpoint*, the *reference object* or the *whole scene* are rotated.

For the intrinsic reference frame if the viewpoint is rotated around the scene then the description of the objects in the scene remains the same. If the ball is in front of the chair intrinsically then this is the case regardless of the angle it is viewed from because the viewer is irrelevant. If, however, the reference object is rotated then the description of the scene does not remain the same. The ball was in front of the chair, but rotating the chair means the ball is now in a region projecting from another side of the chair (unless the rotation is 360°). Finally, if the entire scene (both figure and reference object) is rotated, as in Figure 4, using an intrinsic reference frame results in the same description of the scene pre and post rotation. In Figure 4 the ball is above the chair, the entire scene is then rotated through 180°, but the spatial relationship remains the same.

For the absolute reference frame if the viewpoint is rotated around the scene then the description of the scene does not change, again, because the viewer is irrelevant. A ball is still north of a chair no matter which angle it is viewed from. Similarly, if the reference object is rotated, the description using an absolute reference frame also remains the same. The ball is north of the chair regardless of the chair's orientation. Finally, if the entire scene is rotated then the description according to an absolute reference frame does change. If the ball is north of the chair and both the chair and the ball are rotated together then the ball is no longer north of the chair. In Figure 4, the ball can be described as above the chair according to the direction of gravity, but the rotation of the scene results in the ball now being below the chair.

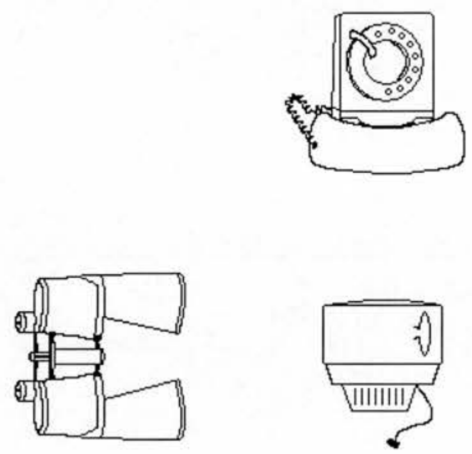


Figure 3. The intrinsic reference frame does not support the logical properties of transitivity and conversness.

For the relative reference frame, if the viewpoint is rotated around the scene the description of the scene changes. If the ball is in front of the chair from a viewer's perspective, then rotating that person around the chair through 180° results in the ball now being described as behind the chair using a relative reference frame. If the reference object is rotated then the relationship between the two objects remains the same according to the relative reference frame because the orientation of the reference object is irrelevant when using the relative reference frame. If the entire scene is rotated then the description according to a relative reference frame does not remain the same. Doing this has the same effect as rotating the viewpoint around the scene.

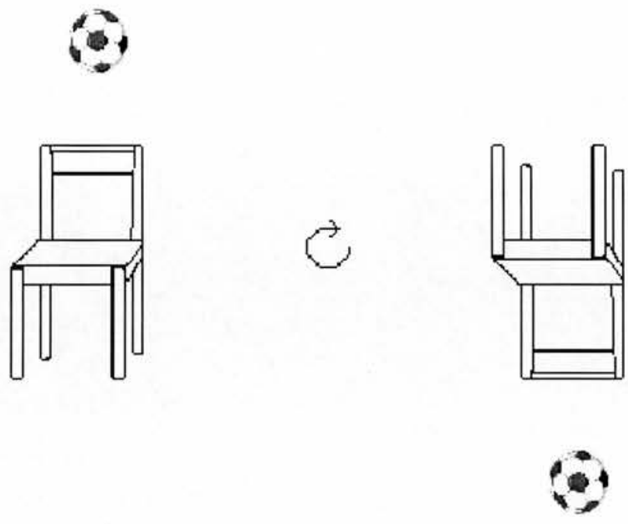


Figure 4: According to the intrinsic reference frame the ball remains above the chair when the whole scene is rotated.

According to Levinson (2003) these three reference frames are the only three reference frames that are available for use. It is not necessary, however, that a language is able to express object location using all three reference frames. In English it is possible to use all three reference frames to describe the location of an object. Other languages rely on one or two of these reference frames, although it does seem that a language needs to have an in intrinsic reference frame if it is also to have a relative reference frame (Levinson, 2003). Each of the three reference frames is incommensurable with each other. That is if one of the reference frames is used to describe a spatial relationship it is impossible to convert the description to another reference frame without additional knowledge. If a speaker says that the tree was north of the car then it is not possible for the addressee to work out the relationship of the tree to the car using an intrinsic reference frame because there is no description of the car's orientation. The tree could be behind the car according to an intrinsic reference frame, or it could be in front of the car according to an intrinsic reference frame; the tree being north of the car does not preclude either of these scenarios. Similarly, there is no viewpoint given when the tree is described as north of the car so the addressee cannot know the relationship between the tree and the car from the speaker's viewpoint (i.e. using a relative reference frame).

2.3.3. Distinctions between Levinson's Taxonomy and the Traditional Taxonomy

The distinction between Levinson's and the Traditional classification system rests upon which feature of a reference frame is used to classify it. For the Traditional theory it is how the axes of the reference frame are labelled and the position of the origin. If a reference frame's origin is situated upon the speaker or addressee, and the axes of the reference frame are labelled according to the directional axes of then it is a deictic reference frame. If the axes are labelled using the intrinsic sides of an object (almost always the reference object) then it is an intrinsic reference frame. If the axes are labelled using a fixed point of the environment then it is an absolute reference frame. For Levinson the logical argument structure is the important feature of a reference frame for its classification and as a consequence this theory is more complex and sophisticated than the traditional approach.

For the most part the distinction between the two definitions has not mattered for research in the field of spatial language. This is because the investigations have focused upon expressions which are intrinsic according to both systems or relative/deictic according to both systems. The distinction between the two theories, however, may be a false one. Levinson (2003) clearly describes his system of absolute, relative and intrinsic as being *linguistic reference frames*, and in the quote below he describes them as a linguistic phenomenon:

> As far as we know, and according to a suitably catholic construal, there are exactly three frames of reference grammaticalized or lexicalized in language (often, lexemes are ambiguous over two of these frames of reference, sometimes expressions will combine two frames, but often each frame will have distinct linguistic expressions associated with it). Each of these frames of reference encompasses a whole family of related but distinct semantic systems. (Levinson, 2003 pp38-39)

If we treat Levinson's (2003) taxonomy as being a taxonomy of *linguistic reference* frames then there is not necessarily any conflict with the Traditional taxonomy, except where the Traditional taxonomy is used to classify reference frames based upon the linguistic evidence. The two are compatible if the linguistic reference frames of Levinson refer specifically to the linguistic representations (e.g. a grammatical representation that encodes argument structure and thematic roles;

Jackendoff, 2002). If this is the case then the Traditional system of reference frames can be thought of as occurring in a representation that is more spatial in nature and encodes different spatial parameters of a reference frame (e.g. the direction, orientation, origin and distance; Logan & Sadler, 1996).

2.3.4. Jackendoff's Reference Frame Taxonomy

In addition to the two systems of classification described above, Jackendoff (1996) suggests the situation may be even more complex and argues for eight different reference frames. Jackendoff (1996) describes his theory as a theory of reference frames, but rather it seems to be a theory describing the different processes of labelling the axes of reference frames (i.e. deciding which axis is the front-back axis, which is the top-bottom axis and which is the left-right axis). Achieving the labelling of the axes of a reference frame is non-trivial and there are many potential solutions (e.g. Jackendoff, 1996, van der Zee & Eshuis, 2003), but it remains an empirical question whether or not labelling the axes is a separate process from determining the reference frame applied to a scene. If the two are separate processes the order in which they occur is a question for future research.

The eight reference frames suggested by Jackendoff fall into two separate categories; four intrinsic reference frames and four environmental ones. The four intrinsic reference frames that Jackendoff (1996) describes are: *the geometric frame, the motion frame, the canonical orientation frame* and *the canonical encounter frame*.

In the *geometric* reference frame the direction of the axes are determined based upon the geometry of the reference object. For example, a symmetrical object can be divided along its axis of symmetry in order to form a top-bottom axis.

The *motion frame* argues that for symmetrical objects (e.g. an electric train), the object's direction of motion will determine which end is labelled the front.

The *canonical orientation* frame of reference determines the object's top-bottom axis with reference to its normal orientation. For example, the canonical orientation of a car is with its wheels on the ground and this therefore defines the bottom of the car.

Finally, the *canonical encounter* frame of reference, determines the intrinsic parts of an object according to the way they are usually encountered by a person. For example, the front of a house is the side that faces the street.

It is important to note that there is no preferred or default system that operates on all objects; often one of these systems can be seen to override another. For example, the front of a crab is defined by the location of its eyes and mouth, rather than the crab's direction of motion. In this case the geometry frame is overriding the motion frame. Similarly, the top of a light bulb is the region at the top of the bulb. Therefore, when the light bulb is in a ceiling light fitting it is upside down. This is an example of the canonical encounter frame (because people pick light bulbs up with the intrinsic top of the bulb upright) overriding the canonical orientation frame.

The four environmental frames, Jackendoff (1996), describes are: *the gravitational frame, the contextual frame, the geographical frame,* and *the observer frame.*

The *gravitational frame* defines the top-bottom axis in terms of the direction of gravity.

With the *contextual frame* an object inherits the directional axes of a second object. For example, if a circle is in the centre of a long narrow room it may inherit as its length axis the room's length axis.

The *geographical* frame is a horizontal absolute reference system such as the cardinal directions.

Finally, the *observer* frame maps the front of an object as the side facing the observer, although in other languages (e.g. Hausa) the opposite side is labelled as the front. Jackendoff suggests that this latter situation may constitute a separate reference frame.

2.3.5. Garnham's Taxonomy

To further complicate the issue, Garnham (1989) has suggested yet a different way of looking at spatial relations. He also divides relational terms into three classes, but in a markedly different way to either the traditional account or Levinson (1996, 2003).

He begins by arguing for a basic level of spatial term. This is basic in the sense that semantically all other spatial relations are elaborations of this class of relations. It includes sentences like *the ball is right of me*. The next level is the deictic meaning of relations, where two objects are related to each other using an egocentric co-ordinate system. The traditional deictic class of reference frames subsumes both of these meanings of spatial relations, whereas Levinson's taxonomy would term Garnham's (1989) deictic meaning as a relative reference frame and his basic meaning as intrinsic. Garnham's (1989) final sense of spatial relations is the intrinsic which occurs when a speaker takes the perspective or axes of another person or object in order to locate a figure object. Garnham ignores environmental reference frames and frames his discussion of the issue in terms of how reference frames would operate in space where there is no fixed environmental bearing.

2.3.6. Reference frame Taxonomy Conclusion

It is evident that no clear and agreed upon definition of reference frames has been achieved in the literature. It is the case that there is a universal acceptance of the plane of gravity as a provider of the up-down axis, and this is reflected in the various definitions. Most of them also acknowledge that there is the potential for a fixed reference frame based upon salient features of the environment in the horizontal plane. Also none deny that an object-centred intrinsic reference frame can be used. Jackendoff (1996) divides this into four further reference frames based upon exactly how directions are assigned to the axes. Reference frames and their definitions are an important issue for research in spatial language, partly because prepositions are often ambiguous across reference frames. The prepositions *above, below, left, right*, *in front of* and *behind* in English are used when both an intrinsic and relative reference frame are employed and so the language is potentially ambiguous.

2.4. Reference frame structure

The reference frame typology that has most often been assumed in the literature is the tripartite one which divides reference frames into absolute, intrinsic and observer based. Therefore, I shall continue only to discuss the taxonomic systems which have such a tripartite nature. These are the reference frame system described by Levinson (1996, 2003) and the Traditional account, which has been used often in the literature (e.g. Miller & Johnson-Laird, 1976). Fortunately, it is not necessarily the case that these two are incompatible (although Levinson does present them as such) if we take the former as defining reference frames linguistically and the latter conceptually.

Reference frames are suggested as some form of mental representation that allows the description of location of objects. That is, they help to translate what is present in a perceptual representation into a linguistic representation. This also happens in reverse when a person hears a spatial description. Take, for example, Figure 2: Here a person has some perceptual representation of the scene. In order to describe the location of the dot they impose a reference frame onto the scene (this the conceptual reference frame) and then finally encode this linguistically (this is the linguistic reference frame).

2.4.1. Levinson's Primitives

Levinson (2003) provides a set of primitives all of which at least must be present to allow the description of object locations. There are four different types of primitives that he mentions. Firstly, there must be a set of linguistic labels that provide the names for the locations (e.g. *above* and *in front of*). These are language specific, and the regions which they describe may also vary by language. Second, he describes a system of co-ordinates that form the axes of the reference system. The co-ordinate systems are polar, there being one fixed axis which then defines the others by rotation from it. He also defines a series of processes, namely, reflection, translation, and rotation, whereby the axis of a primary co-ordinate system (C^1) can be altered to form a secondary co-ordinate system (C^2). The third set of primitives Levinson defines is a series of points that represent the scene. These are; the figure, the ground (reference object), viewpoint of an observer, the origin of the co-ordinate system and secondary co-ordinate systems, an anchor point which fixes labelled co-ordinates and a known landmark. Note that not all spatial relations require the representation of all points (e.g. a spatial description using an intrinsic reference frame does not require the representation of a viewpoint). Levinson also describes an anchoring system that is a description of how co-ordinates are labelled. The anchoring system maps the language specific labels onto the co-ordinate system used. In the case of the intrinsic reference frame the anchoring system is usually a part, or intrinsic side, of the reference object. As an example, this is how Levinson (2003) formally describes the intrinsic reference frame:

> An intrinsic spatial relator R is a binary spatial relation, with arguments F [figure] and G [ground], where R typically names a part of G. The origin X of the co-ordinate system C is always on the volumetric centre of G. An intrinsic relation R(F,G) asserts that F lies in a search domain extending from G on the basis of an angle or line projected from the centre of G, through an anchor point A (usually the named facet 'R'), outwards for a determined distance. F and G may be any objects whatsoever (including ego), and F may be a part of G. The relation R does not support transitive inferences, nor converse inferences. (Levinson, 2003, pp42-43).

Above I have described how the linguistic reference frames can be thought of as residing in a linguistic representation, but the primitives described by Levinson are clearly meant to be part of representations that are more than just linguistic in nature. The primitives that are described by Levinson are ones which would not be represented solely by the language faculty. In particular, the system of co-ordinates and the processes of reflection, translation and rotation, by which a secondary co-ordinate system is derived from a primary co-ordinate system, would seem to demand that there is some sort of spatial/conceptual representation. The rest of the primitives can be perfectly adequately accommodated as part of a linguistic representation that represents argument structure (Levin, 1993). In a similar fashion to verbs all the semantic arguments of a spatial expression are not necessarily explicitly realised.

2.4.2. Logan and Sadler's Reference Frame Structure

If Levinson describes the primitives that form the representation of linguistic reference frames, then there are a different set of primitives that form the primitives of the conceptual/spatial reference frames (the exact locus of this representation in the mental architecture is an issue of some debate; e.g. Carlson, 2000). Logan and Sadler (1996) provide the basis of one of the best computational models of just such a reference frame. They describe that:

> The reference frame is a three-dimensional coordinate system that defines an origin orientation, direction and a scale. It serves as a map between the conceptual representation and the perceptual representation, establishing correspondence between them. (Logan & Sadler, 1996 p499)

It seems from this that Logan and Sadler intend the reference frame to be a kind of process or representation located in a translation module (Jackendoff, 2002) that enables a correspondence between the conceptual representation and the perceptual representation. The conceptual representation is earlier described by Logan and Sadler (1996) in a way which sounds very similar to the linguistic reference frames described by Levinson (2003).

> The conceptual representation is a one-, two-, or three place predicate that expresses a spatial relation. The conceptual representation identifies the relation (e.g., it distinguishes above from below); it individuates the arguments of the relation, distinguishing between the reference object and the located [figure] object; it identifies the relevant reference frame (depending upon the nature of the reference object); and it identifies the relevant spatial template. The conceptual representation does not identify objects and relations directly in the perceptual representation; further processing and other representations are needed for that. (Logan & Sadler, 1996 p498).

This description is similar to that of the linguistic reference frames described by Levinson (2003) and therefore the two different theories of reference frames can be seen as compatible, with Levinson's linguistic reference frames taking the role of a conceptual/linguistic representation and Logan and Sadler's reference frames taking the role of a conceptual/perceptual representation.

The structure of the reference frames described by Logan and Sadler (1996) is based upon four parameters; the *origin*, *orientation*, *direction* and *scale*. Each of the three reference frames use the same representation, but differ in how the parameters are set. The parameters, Logan and Sadler (1996) argue, can be set consciously depending upon the goals of the speaker. The *origin* parameter refers to the object which the origin is situated upon. For example the deictic/viewer centred reference frame sets the origin upon the viewer. The intrinsic reference frame sets the origin upon the reference object. Logan and Sadler do not specify how the origin is set for the absolute reference frame. The *orientation* parameter determines which axes are the horizontal axes and which axis is the vertical axis. The orientation parameter can be set in line with an object or the viewer or an environmental conceptual plane (e.g. as in the cardinal directions). The *direction* parameter determines the end points of an axis (e.g. which is the *top* and *bottom* of the vertical axis). The orientation and direction parameters are represented hierarchically, with the orientation parameter being set prior to the direction parameter (Logan, 1995). Finally, the *scale* can either be set in terms of the viewer or an object; with the scale differing depending upon the size of the objects (e.g. *the ball is near the table* sets the scale parameter at a lesser value than the *planet is near the Earth*; Carlson & Covey, 2005). Carlson (2003) has argued that the term distance should be used as opposed to scale, because scale implies a series of discrete intervals, for which there is no evidence. In the remainder of the thesis I also shall use the term distance.

In addition to the parameters that are set for a reference frame, Logan and Sadler (1996) also describe four processes that are necessary for the apprehension of a spatial relationship. These are *spatial indexing*, *reference frame adjustment*, *spatial template alignment* and *computing goodness of fit*. *Spatial indexing* is the process which binds objects in a perceptual representation to the figure and reference arguments of the conceptual representation. For example, in the sentence *the ball is in front of the car*, the car is identified as the reference object and the ball as the figure object.

Reference frame adjustment is the process which sets the parameters of the reference frame. It is suggested that not all parameters of a reference frame are required to be set for all spatial terms. For example, *near* does not require the direction parameter to be set, whereas *above* does not require the distance parameter to be set. In the example *the ball is in front of the car*, if we assume this is according to an intrinsic reference frame, the origin is set on the reference object (the car). The orientation of the primary horizontal axis of the reference frame is aligned with the axis of elongation of the car (i.e. from the headlights to the taillights of the car). The direction of the primary horizontal axis is then set so the *front* end and *behind* end of the axis are the ends that correspond with the headlights and taillights of the car respectively. The *distance* parameter is not set because the spatial term *in front of* does not specify a distance. However, there is evidence to suggest that even though a parameter may not be necessary for the spatial term to be used it is still defined and represented (Carlson & Van Deman, 2004).

Spatial template alignment is the process where the spatial template (this is a representation which defines the regions of acceptability for a given spatial term around an object) is aligned with the reference frame. For example, the spatial template for the term *above* would be aligned with the top-bottom axis of the reference frame. So, continuing with the sentence, *the ball is in front of the car* the spatial template for the term *front* would be aligned with the front-back axis. The spatial template then parses space around the car into *good*, *acceptable* and *bad* regions of appropriateness for the term *front*.

Computing goodness of fit is the process which, once the spatial template is aligned, computes which spatial term is most applicable for the spatial relation. Logan and Sadler (1996) suggest this is done in parallel, where goodness of fit is computed for a whole range of candidate spatial templates, and weights are applied ranging from 1.0 - 0.0 for the regions around the reference object. However, there are other suggested models which compute goodness of fit based upon attention and the weights of vectors between the reference and figure object (Regier & Carlson, 2001).

There is good evidence that the different stages proposed by Logan and Sadler (1996) do exist. Logan and Compton (1996) showed that subjects were slower to verify whether one letter was above another letter when a distracter letter was between the letter that was the figure object and the letter that was the reference object. The time taken to verify the sentence increased as the number of distracter letters between the figure letter and the reference letter increased. The increase could have been due to the whole process of spatial apprehension being computed (i.e. the spatial relation between each of the distracter letters and the reference object being computed), or it could have been down to the process of spatial indexing alone and only when the correct letter was identified as the figure object was the spatial relation between the two computed. Carlson and Logan (2001) provided evidence that the latter was the case; that the effects of the distracter letters on verification time were due to difficulty in the stage of spatial indexing and that the spatial relation between the distracters and reference object was not computed. Carlson and Logan (2001) showed that the increase in time to verify the sentence was independent of the placement of the distracter. In this experiment, the distracters were placed not only in the same relation to the reference object as the figure object, but also in the opposite relation. Therefore, once the reference object was identified, attention could be moved in the appropriate direction and only encounter the figure object.

F	F
D	
R	R
	D

Figure 5: On the left a scene from Logan & Compton (1996), on the right a scene from Carlson & Logan (2001).

Figure 5 shows, on the left, a comparable scene from Logan and Compton (1996), where the figure object (F) is above the reference object (R) with the distracter object (D) between the two and therefore also above R. On the right of Figure 5, is a comparable scene from Carlson and Logan (2001), here the F and R are in the same position, but D is in the opposite spatial relation from F (i.e. below R). The effects on verification time were the same regardless of the distracter's placement (i.e. the same amount of time for the scene on the right of Figure 5 as the scene on the left). It took longer to verify the sentences as true or false when there was a distracter present compared to when there was no distracter present. This is strong evidence that the stage of spatial indexing is indeed a separate stage of computation and is independent of the other stages.

Further more direct support for the independence of the stages of spatial apprehension comes from Carlson, West, Taylor, and Herndon (2003). They used event related potentials (ERPs) to measure the electrical activity associated with verifying the spatial relation between a figure and a reference object, when the reference object was rotated 90° from its canonical position and so dissociating the intrinsic and absolute/relative reference frames (these two were coincident as the participant was upright). Participants were asked to make judgements about the spatial relation using one of the two alternative reference frames (intrinsic vs. absolute/relative) or both reference frames. The results distinguished between three distinct processes.

The effect of the rotation of the reference object was shown in a waveform located over the parietal region which differed when the reference object was in a non-canonical orientation. This reflects the increased difficulty with spatial indexing due to rotated objects being more difficult to identify than canonically oriented objects

(Jolicoeur, 1985). This was the first effect observed on the waveform which is consistent with the notion that spatial indexing occurs prior to the other processes of spatial apprehension (Carlson & Logan, 2001); although Logan and Sadler (1996) argue that the order of processes depends upon task demands.

When the reference object was in a non-canonical orientation the different reference frames were dissociated and so there was potential competition between reference frames for selection (Carlson-Radvansky & Logan, 1997). This was reflected in a frontal slow wave later in the time course than the parietal wave associated with spatial indexing. Frontal waves are associated with resolution of conflict such as is the case with Stroop like tasks (West & Alain, 1999). Therefore, Carlson et al (2003) argue that the frontal slow wave is a measure of difficulty in the process of setting the reference frame parameters.

Finally, there was a late positive waveform, occurring over the occipital-parietal region associated with the computation and comparison process. When the reference frames were aligned, there was a greater positivity to the waveform than when the reference object was in a non-canonical orientation and so the reference frames were dissociated. Carlson et al (2003) argue that this reflects the easier nature of computing the spatial relation when the reference frames are all aligned. Importantly, the late positive waveform was independent of whether or not the figure object was in the spatial relationship described by the sentence. This is because regardless of the veracity of the sentence the spatial relationship would still have to be computed in order to establish whether or not the sentence was true. Therefore, the benefit of having all reference frames aligned would still be evident. In addition to the other studies (Carlson & Logan, 2001, Logan & Compton, 1996) the results of Carlson et al (2003) are strong evidence that the process of spatial apprehension does indeed follow several separate distinct levels of processing.

According to Logan and Sadler (1996) a reference frame serves to enable a mapping from a perceptual representation to a conceptual representation. This is further explained by Logan (1995) who argues that a reference frame is a representation that allows attention to shift from one object to another. It is true that in order to apprehend a spatial relation attention is required (Logan, 1994). If a person is asked to indicate whether or not a plus above a dash (the target) is present amongst a group of distracters that are dashes above pluses then it takes longer the more distracters there are present. This is an indication that attention is required in order to 'see' the spatial relation; if attention were not required then the target would exhibit a 'pop out' effect and there would be no difference in the time to find the target regardless of the number of distracters (Treisman & Gelade, 1980, Treisman & Gormican, 1988).

2.5. Reference frame selection in production and comprehension

The ambiguity of spatial expressions across reference frames makes the choice of reference frame for a speaker describing a spatial relationship between two objects an important one. A speaker must choose a reference frame that the addressee will be able to understand. Similarly, in comprehending a spatial description the addressee must interpret the discourse using the correct reference frame in order to avoid misunderstandings. Often contextual factors make the decision for the speaker or addressee easy. For example, the reference object may have no intrinsic axes (e.g. *the chair is in front of the ball*) and so the description has to be interpreted using a relative reference frame. The syntax of the description may also play a role. For example, *the ball is on the chair's right*, implies an intrinsic reference frame. Intuitively it seems that addressees do not often misunderstand which reference frame a speaker is using (although this is an empirical question), but if this is the case then it is an important question as to why it is that people are very good at understanding which reference frame to use or is being used.

The issue of how people think about space (and to a large degree thinking about space is about which reference frame to use) has been around for a long time. Newton drew a distinction between absolute space, which was an infinite plane, and relative space, which located objects in terms of where other objects were. Leibniz took the idea of thinking about space *relatively* further and argued that when there are two objects, and one is moving, it is a matter of convenience as to which one we attribute motion and which we believe to be stationary. To a large degree the argument that humans think about space in a relative manner and in terms of their own bodily axes (deictically) has persisted through to recent history in cognitive psychology. Miller and Johnson-Laird (1976) argue that the default way a human speaks and thinks about space is deictically (i.e. in terms of their own bodily axes). Similarly, Levelt (1989) argues that the default reference frame in English is relative. These two claims are similar, but not quite the same. For Miller and Johnson-Laird (1976) the *relative* nature of human language is a universal common to all languages.

However, this appears not to be the case. Levinson (2000, 2003; Levinson, Kita, Haun & Rasch, 2002) has reported several languages (e.g. Tzeltal) that rely almost entirely on the absolute reference frame and use this for descriptions of even small scale scenes (e.g. the fork is north of the plate). Levelt (1989) on the other hand, claimed that the relative reference frame is the preferred reference frame in English. This is by no means a proven fact, and therefore it is necessary to assess the evidence regarding which perceptual layouts permit the use of which reference frames, and in which circumstances which reference frames have a preferred use.

2.5.1. The Principle of Canonical Orientation

Levelt (1984, 1996) presented the *principle of canonical orientation*, which argued when it is and when it is not appropriate for a person to use an intrinsic reference frame to describe the relationship between two objects. The principle of canonical orientation states that:

> "For the intrinsic system to refer to a relatum's intrinsic dimension, that dimension must be in a canonical respect to the perceptual frame of orientation." (Levelt, 1996 p92)

This essentially denies that a preposition can be used to describe an intrinsic spatial relationship, if the preposition is not being used to describe a spatial relationship in its canonical plane of orientation. For example, Levelt argues that intrinsic *above* cannot be used (Levelt argues there are very few instances where *above* and *below* can be used in an intrinsic sense) if doing so contradicts *above* in the absolute or deictic reference frame. Figure 6 shows a car tilted on its side so that its roof faces deictically *right*, the ball can be described as *right* of the car, but according to Levelt (1996) it cannot be described as *above* the car, despite being in a region that projects from the car's intrinsic top, because the scene violates the principle of canonical orientation. The violation comes from the fact that intrinsic *above* would be used to describe a spatial relationship in the horizontal plane and this is exactly what the principle of canonical orientation does not permit.

The principle of canonical orientation also operates on prepositions in the horizontal plane (e.g. *left, right, in front of, behind*). Figure 7 shows a cannon tilted so that it is facing upwards in the plane of gravity.

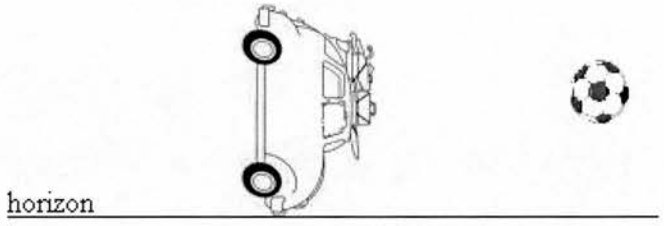


Figure 6: The ball cannot be described as *above* the car because it violates the principle of canonical orientation (Levelt 1984, 1996)

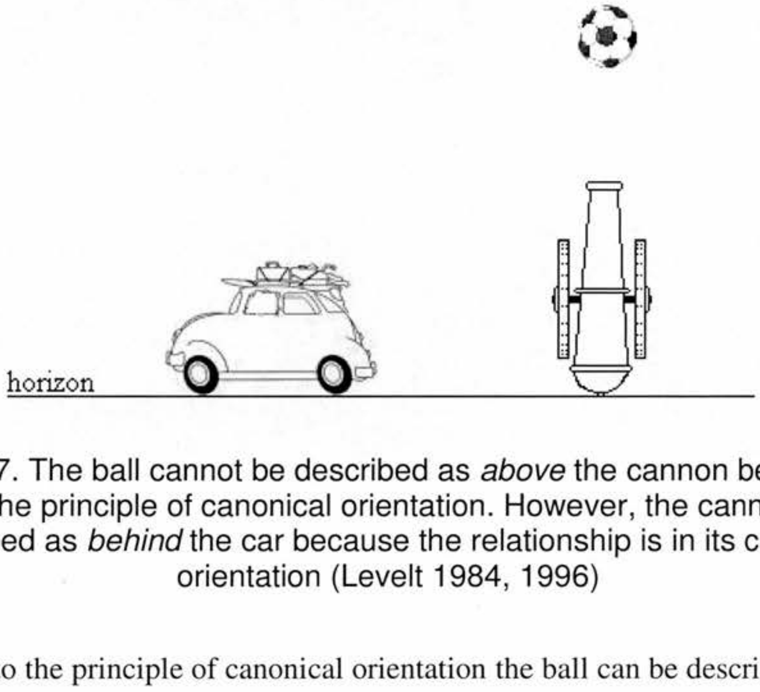


Figure 7. The ball cannot be described as *above* the cannon because it violates the principle of canonical orientation. However, the cannon can be described as *behind* the car because the relationship is in its canonical orientation (Levelt 1984, 1996)

According to the principle of canonical orientation the ball can be described as *above the cannon* (using an absolute and deictic reference frame), but it cannot be described as *in front of* the cannon using an intrinsic reference frame because *in front of* is canonically used for horizontal relationships and the relationship in Figure 7 is a vertical one. However, the cannon can be described as *behind the car* using an intrinsic reference frame because this relationship is in the horizontal plane.

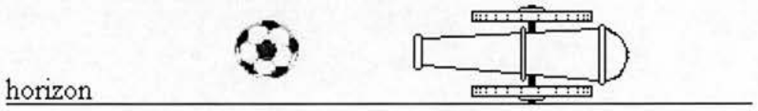


Figure 8. The ball can be described *as in front of* the cannon because it is in the horizontal dimension

It is important to note that the principle of canonical orientation does not mean that the reference object must be in its canonical orientation, just that the intrinsic axes of the reference object must be in canonical planes of orientation in order for the intrinsic reference frame to be used. Figure 8 shows the cannon again tilted on its side, but this time the ball can be described as *in front of the cannon* because it describes a horizontal relationship between the figure and reference object, and the front-back axis is canonically a horizontal axis.

These intuitions of Levelt (1984, 1996) have probably helped lead him to the conclusion that the deictic reference frame is the default in English because according to Levelt the intrinsic reference frame cannot be used to describe many situations. In addition the intrinsic reference frame, as already described, allows few inferences about a situation to be made and so may not be as useful as other reference frames. However, the principle of canonical orientation is largely based upon intuition and has been challenged both theoretically (Garnham, 1989) and empirically (Carlson-Radvansky & Irwin, 1993, Carlson-Radvansky & Logan, 1997).

2.5.2. The Framework Vertical Constraint

Garnham (1989) proposes an alternative to the principle of canonical orientation, *The framework vertical constraint*. The framework vertical constraint elevates the vertical axis to special status because of its universal solution to defining up and down (i.e. the direction of gravity). Garnham argues that the intrinsic reference frame cannot be used if the resultant relationship contradicts the relationship provided by the direction of gravity. Therefore, the ball in Figure 6 cannot be described as *above the car* because the intrinsic definition of *above* does not coincide with the absolute definition of *above*. In figure 7, the ball cannot be described as *in front of the cannon* because this also contradicts the relationship provided by the absolute reference frame. The framework vertical constraint again is largely based upon intuition, and is proposed as an alternative to the principle of canonical orientation, because Garnham (1989) argues that it can explain the use of the intrinsic reference frame with the postulation of a single rule (the framework vertical constraint), and is therefore more parsimonious than the principle of canonical orientation.

2.5.3. Empirical Evidence for Reference frame Selection

Empirically there is little support for either the principal of canonical orientation or the framework vertical constraint (Carlson-Radvansky & Irwin, 1993). Data from a variety of studies have shown that the intrinsic reference frame can be used in many more situations than either of the two theories allow. Carlson-Radvansky and Irwin (1993) presented subjects with a series of pictures where the reference object in a scene was either in its canonical orientation or in a non-canonical orientation, as a result of some plausible scenario (e.g. a car being picked up by a crane). Also included in the scene was a dot that represented a fly, which was the figure object. Subjects were shown a sentence describing the location of the fly in relation to the reference object and then rated how a good a description of the scene the sentence was. The results showed that although subjects did have a strong preference for using the deictic/absolute reference frames (the two were conflated because the subjects were upright), there were higher ratings of acceptability when the sentence described a relationship that was only acceptable using an intrinsic reference frame than when the sentence was incorrect using all reference frames. Further, the ratings for correct descriptions in the canonical orientation condition (when all three reference frames were aligned) were significantly higher than ratings for correct deictic/absolute descriptions in the non-canonical condition (when the intrinsic reference frame was dissociated from these two). What this indicates is that although there is a preference for using the deictic/absolute reference frame, the intrinsic reference frame is still an acceptable interpretation of a scene. A further experiment, in which subjects were asked to place the fly in accordance with the sentence's description, confirmed these findings. When the reference object was in a non-canonical position the fly was placed according to the deictic/absolute reference frame 83.8% of the time and according to the intrinsic reference frame 8.8% of the time. So although rare, people do violate the principle of canonical orientation and the framework vertical constraint. 7.5% of the dot placements were ambiguous between the intrinsic and the deictic/absolute reference frame which may reflect a combined usage of both reference frames (Carlson-Radvansky & Logan, 1997).

2.5.4 Functional Influences on Selection

These results are only partially contradictory of the Garnham (1989) and Levelt (1984) theories, because the still results show a high preference for the use of the deictic/environmental reference frame. However, this may be a context dependent factor. Carlson-Radvansky and Irwin (1993) increased the acceptability of the intrinsic reference frame by manipulating three factors. These were: Including an additional reference object, having only intrinsic descriptions, and decreasing the distance between the figure and reference object. The functional relationship between the figure and reference object also affect the reference frame selection process. Carlson-Radvansky and Tang (2000) conducted an experiment, using a similar method to that of Carlson-Radvansky and Irwin (1993), but in this experiment the figure and reference object could either be, functionally related and interacting (e.g. a sauce bottle over a hotdog where the bottle is upside down as if it was pouring sauce onto the hot dog), functionally related, but not interacting (e.g. the sauce bottle is in a position where it is not able to pour sauce on the hotdog), or unrelated (e.g. a bottle of ant killer and a hotdog). In the latter condition the reference object shares a similar geometry to the sauce bottle and so differences between the conditions cannot be attributed to geometric factors. However, ant killer and hot dogs are not typically functionally related. The results showed that subjects gave higher acceptability ratings for intrinsic descriptions when the figure and reference objects were related and interacting than when they were not interacting, or unrelated. Ratings for sentences describing a deictic/absolute relationship were not affected by this manipulation. Similarly, Carlson-Radvansky and Radvansky (1996) showed that subjects were more likely to describe the relationship between two objects using an intrinsic reference frame, when the two objects were related and interacting than when the two objects were related but not interacting, or not related.

The imposition of a reference frame is also affected by the functional relationship between the figure and reference object. Carlson-Radvansky, Covey and Lattanazi (1999) asked participants to place figure objects *above* reference objects that had functional parts offset from the centre of mass of the object (e.g. toothpaste above a toothbrush). They found that when the figure object was related functionally to the reference object, the placement of the figure object was displaced towards the functional part and away from the centre of mass. When the figure object was unrelated to the reference object, but of equivalent size and shape (e.g. a tube of oil paint), participants placed the figure object closer to the reference object's centre of mass. Interestingly, the displacement towards the functional part, for functionally related figure objects, was not complete. Instead the object was placed between the functional part of the reference object and its centre of mass, leading Carlson-Radvansky et al (1999) to suggest that there may be a functional reference frame in addition to a geometric reference frame and that these interact to produce the placement.

These results show that the relationship between the figure and reference object, and in particular the functional relationship, have a large effect upon reference frame selection (Carlson, 2000). The evidence suggests that there are a large variety of factors that influence whether or not an intrinsic reference frame is used to describe a spatial relationship or used to interpret a spatial description. These factors are contextual and pragmatic in nature for the main and show that broad rules, such as the principle of canonical orientation and framework vertical constraint, do not prevent the use of the intrinsic reference frame.

2.5.5. Influence of the Deictic Reference Frame

Evidence from Carlson-Radvansky and Irwin (1993) also questions the influence of the deictic reference frame on reference frame processing. Experiment 4 from this study dissociated all three reference frames by tilting subjects onto their sides, as is the case in Figure 2. The subjects were shown scenes similar to the production and ratings experiments and asked to produce descriptions that described the relationship of the figure object to the reference object. Less than 1% of the participants' descriptions used a deictic reference frame. In addition, participants used an absolute reference frame as often when it was dissociated from the deictic reference frame as when it aligned with the deictic reference frame. In the example of Figure 2 people would very rarely describe the dot as *left* of the chair (using a deictic reference frame), and that if the person in the figure were upright this would not increase the likelihood of them describing the dot as *above* the chair (using the now co-incident absolute and deictic reference frame). There was also the same number of intrinsic descriptions when the deictic and intrinsic reference frames were aligned as when they were dissociated, again highlighting the limited influence of the deictic reference frame. The lack of effect of the deictic reference frame may be because participants lying on their side mentally rotated themselves to an upright position. This would mean representationally the individual would be in line with the environment.

In fact, Friederici and Levelt (1990) have shown that astronauts in space are still able to describe the location of objects in relation to other objects, even though there is no gravitational frame to help them, by relying upon the deictic reference frame. Additionally, they showed that even when there were features of the environment with top-bottom axes (pictures of two trees) there was still a large preference to use the deictic reference frame. The evidence from the studies in space was also backed up by studies on Earth which required participants to describe the location of objects when the direction of gravity was not task relevant (by getting participants to lie on their backs). Again in this experiment participants preferred a deictic reference frame to a reference frame that was based upon features presented in the environment. So the deictic reference frame can be used to describe spatial locations in at least some situations when it is difficult to use the direction of gravity.

Carlson-Radvansky and Irwin (1993) attempted to show the default preferences people have for reference frame selection when describing the location of objects, but there are many other factors that affect reference frame selection. The functional relationship between two objects, mentioned previously, increases the likelihood that a person will use an intrinsic reference frame (Carlson, 2000). The functional relationship between two objects also has other effects on cognition. If two objects are functionally related then the spatial relationship between the two objects is more likely to be remembered and sentences describing the spatial relationship between the two objects are read faster (Radvansky & Copeland, 2000). Functional relations between figure and reference objects have an impact on many aspects of spatial relations (Carlson & van der Zee, 2004). In conversation the needs of an addressee, such as the difficulty in understanding the locative utterance, affects the reference frame used by a speaker (Schober, 1993). The nature of the environment can influence whether or not a scene is described as if walking through it or as if from above (Taylor & Tversky, 1992). There may also be a role to play for individual differences; different people may prefer to use one or the other reference frames (Carlson-Radvansky & Logan, 1997).

2.6. The Influence of Precedence on Reference Frame Selection

In addition to the factors mentioned in the previous section reference frame selection may also be biased by 'precedence'. That is, if one particular reference frame has been used before (either by oneself or ones partner in a conversation), then it is possible that a person is increasingly likely to continue using the same reference frame. This effect is demonstrated in other linguistic domains such as syntactic priming (Bock, 1986).

Prior to the work in this thesis there was no evidence that the use of a reference frame, on a previous occasion increased the likelihood of use on a subsequent occasion. However, Carlson-Radvansky and Jiang (1998) showed that comprehension of a scene using a particular reference frame, did affect the ease of comprehension of a subsequent scene, using an alternative reference frame. In this experiment, participants were asked to decide whether a dot was above a reference object or not. The reference object was rotated 90° in order to dissociate the relative and intrinsic reference frames. Participants were instructed to indicate that the dot was above the reference object if it was above according to either an intrinsic or a relative reference frame (the relative reference frame was actually a 'coincident' reference frame with the absolute reference frame). Unknown to the participants the trials of the experiment were paired into prime and probe trials. The conditions of the experiment are shown in Table 1. On the prime trial the dot was above the reference object according to a relative reference frame, as in the experimental prime trials shown in Table 1. Then on the probe trial the dot was either above the object according to the intrinsic reference frame, or below the object according to the intrinsic reference frame, shown in the matched and mismatched probe trials of Table

1. The control trials used an object without intrinsic sides on the prime trial (e.g. a ball) and the dot was placed above according to a relative reference frame. The results showed that participants were faster to verify that the dot was above the reference object on the probe trial for the control condition than the experimental condition. Furthermore, they showed that this effect occurred when the spatial terms matched from prime to probe and when they mismatched. In this experiment, interpreting the dot above the reference object on the prime trial using one reference frame made it more difficult to interpret the dot below or above the reference object using the alternative reference frame on the probe trial.

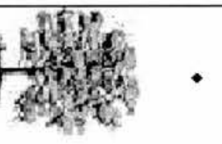
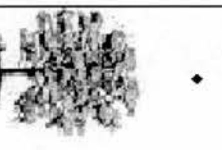
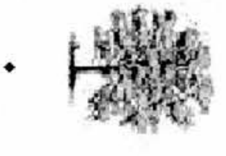
	PRIME TRIAL	PROBE TRIAL
Experimental Pair		
Control Pair		
Matched Relations	ABOVE	ABOVE
Experimental Pair		
Control Pair		
Mismatched Relations	ABOVE	BELOW

Table 1. The conditions used in Carlson-Radvansky and Jiang (1998). (note these are not the actual objects used in their experiment)

Carlson-Radvansky and Jiang (1998) argued that initially all reference frames were being imposed upon the scene and then the non-selected reference frames were inhibited, allowing the correct interpretation to be made. The inhibition of the non-selected reference frame on the prime trial creates difficulty via negative-priming on the probe trial. The control condition uses a ball as a reference object and so an intrinsic reference frame cannot be imposed and therefore inhibition is not required when using the relative reference frame.

In a second paper, Carlson and Van Deman (in press) extended these findings and showed that the negative-priming effect operates across multiple axes. Carlson-Radvansky and Jiang (1998) only investigated the effect of inhibition within a single axis (the top-bottom axis) and on the endpoint of an axis. Carlson and Van Deman (in press) used the same negative priming paradigm, but included the conditions shown in Table 2. (These are not the actual stimuli used in the experiment).

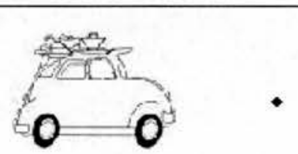
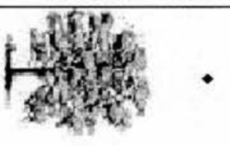
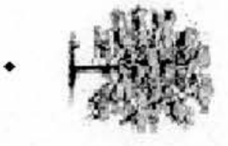
	PRIME TRIAL	PROBE TRIAL
Experimental Pair		
Mismatched Relations	FRONT	ABOVE
Experimental Pair		
Mismatched Relations	LEFT	BELOW

Table 2. The conditions from Carlson and Van Deman (in press) (not actual stimuli).

They showed that using front or back on the prime trial made it more difficult to use above and below in the alternative reference frame on the probe trial. The reverse was also true. This is evidence that there is inhibition between axes of a reference frame from the front-back axis to the top-bottom axis. The left-right axis exhibited a somewhat different pattern of results. In this experiment, participants found it very difficult to interpret the scene using the intrinsic left-right axis. As such, participants showed no inhibition of the left-right axis in this experiment. A significant boost in activation was shown for left and right when the same spatial term had been used on the prime trial. That is, when *left* was used on the prime trial there was an increase in speed using *left* in the same reference frame on the probe trial. This latter effect may, however, be a lexical effect.

2.6.1. Privileged Axes

Carlson and Van Deman (in press) argue that their pattern of results is explained by the privileged status of the front-back and top-bottom axes, compared to the left-right axis. The first two can be considered directed axes because they have a direction that is salient and discernable from the features of the reference object. The left-right axis is much more difficult to compute than the other axes in many tasks (Maki, Grandy & Hauge, 1979, Franklin, Tversky & Coon, 1992), because of the relative symmetry of this axis (Bryant & Wright, 1999). Because of the directional status of the front-back and top-bottom axes they are highly salient and so always active for a reference frame when computing the spatial relation between a figure and a reference object (obviously they are not active intrinsically if the characteristics of the reference object do not support an intrinsic reference frame e.g. a ball). Therefore, if a person wants to use the relative reference frame it requires inhibition of the privileged axes of the intrinsic reference frame and this makes these axes more difficult to use on an immediately subsequent occasion. The left-right axis does not need to be inhibited because it does not enjoy the same privileged status as the other two axes. The endpoints of the left-right axes do have an increase in activation if that endpoint was used on an immediately previous occasion. Carlson and Van Deman (in press) argue that this is evidence of a common representation that subserves both perception and language because the position in space and its relation to the sides of the reference object is represented along with the verbal label for the position.

The notion of privileged axes is supported by evidence from studies investigating access to reference frames from memory. It is known that people encode environments in memory from an egocentric perspective (Shelton & McNamara, 1997), and as such are faster and more accurate at indicating the spatial relationships between objects in an environment if they are imagining them from the perspective they originally viewed it. This has an effect on the way people are able to describe the location of objects in imagined environments, or environments which they are recalling from memory. Franklin and Tversky (1990) suggested the spatial framework analysis to account for how people access reference frames. This proposes that the different axes of the body are accessed with differing degrees of ease based upon various features of the axis. For an upright person, the head-feet axis is accessed fastest because this is an asymmetrical axis, both in terms of the physical features of the person and because in an upright person it is aligned with the direction of gravity. Following this, the front-back axis is accessed next fastest because of the asymmetry between the location of the individuals perceptual apparatus. The left right axis is accessed slowest because of the symmetry of this axis. There is much evidence to support this model of reference frame access. Franklin et al (1992) asked participants to read passages that described a scene and the location of objects in the scene. During the test phase, the participants imagined themselves moving in the described environment. Participants then saw a name of one of the objects in the scene and had to say the direction in which it lay as fast as possible. The results matched the predictions of the spatial framework analysis. Participants were fastest at giving directions of the objects that were above or below them, and they were slowest to name the direction of objects when the objects were to the left or right. Importantly, when the imagined change involved the participants imagining themselves lying down the results changed. Participants were still slowest to name the directions of objects that lay to their left and rights, but they were now fastest to name objects that were in front or behind them. This is congruent with the findings of Carlson and Van Deman (in press) arguing that certain axis are privileged based upon the directional salience of the axis. They argue that people do not represent the intrinsic left-right axis and therefore it does not need to be inhibited when the relative left-right axis is used and hence they found no negative priming of this axis.

2.7. Influence of Non-Selected Reference Frames

The work of Carlson-Radvansky and Jiang (1998) and Carlson and Van Deman (in press) showed that non-selected reference frames are inhibited. However, there is still evidence that non-selected reference frames have an influence on subsequent processing.

2.7.1 Spatial Templates

Carlson-Radvansky and Logan (1997) showed that before a reference frame is selected there are spatial templates aligned with multiple reference frames. A spatial template is a representation which parses space around a reference object into *good*, *acceptable* and *bad* regions of applicability for each spatial term. For polysemous spatial terms there are multiple spatial templates: Spatial templates do not have a one to one correspondence with language. Logan and Sadler (1996) investigated the spatial templates of several different spatial terms, using several different experimental methods. For example, for the spatial term *above* participants were asked to draw an X above a box. Logan and Sadler found that all participants placed the X in a position within the boundaries of the edges of the box and above it in the plane of gravity. In a second experiment, they had subjects rate the applicability of a sentence as a description of the spatial relation between the box and an X. In this experiment the X was positioned directly above the box, in various positions to the left and right of directly above the box, and in positions that were not above the box. Logan and Sadler (1996) found that the highest acceptability ratings for above were given when the X was positioned directly above the box. This was termed the good region. When the X was positioned in regions of space to the left and right of the *good* region the acceptability ratings of the sentence *the X is above the box* were lower. This was termed the *acceptable* region. The region termed *bad* had the lowest sentence acceptability ratings for the sentence the dot is above the box. The boundaries between the different regions of acceptability, Logan and Sadler (1996) argue, are graded rather than discrete. Figure 9 shows the spatial template for the spatial term *above*. A similar good, acceptable and bad structure of spatial templates was found for other spatial terms.

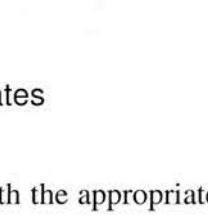


Figure 9. The spatial template for the spatial term *above*.

2.7.2. Multiple Spatial Templates

Spatial templates are aligned with the appropriate axis of the reference frame during the apprehension of a spatial relationship (Logan & Sadler 1996). For example, the spatial template for *above* is aligned so that the *good* region of acceptability coincides with the top half of the top-bottom axis. The top-bottom axis itself is aligned differently depending which reference frame is being used. Therefore Carlson-Radvansky and Logan (1997) asked two questions: The first was whether or not multiple spatial templates are active during a period in processing before a reference frame is selected, and the second was whether or not multiple spatial templates work to form composite spatial templates. They used a similar paradigm to that of Logan & Sadler (1996) (although focused solely on the spatial terms *above* and *below*), where participants had to rate the acceptability of a sentence describing a spatial relationship. Again the location of the figure object was varied around the reference object. The major difference in the Carlson-Radvansky and Logan (1997) study was that the reference object was rotated 90° so that the intrinsic and the coincidence deictic/environmental reference frames were dissociated. For the spatial term above, two potential spatial templates could be formed, one aligning with the intrinsic reference frame and one aligning with the deictic/environmental reference frame. In addition, the acceptable regions for both of these spatial templates would partially overlap. This is shown in Figure 10.

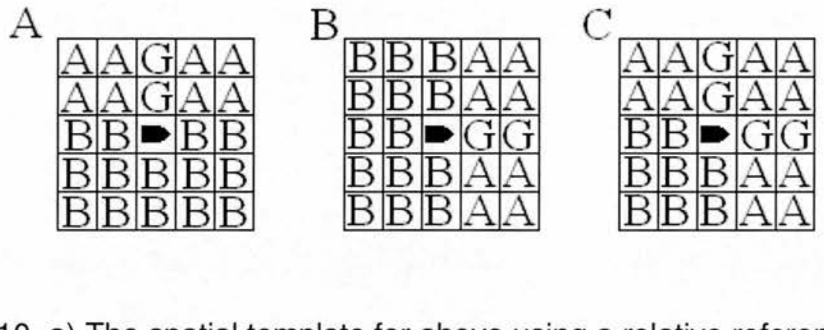


Figure 10. a) The spatial template for above using a relative reference frame when the reference object is non-canonically orientated. b) The spatial template for the above using an intrinsic reference frame when the reference object is non-canonically orientated. c) The combined intrinsic and relative spatial template for above.

The results showed that there were higher ratings of acceptability when the figure object was in the region deemed *good* according to either of the reference frames. This showed that both reference frames were active, at least on some of the trials for some of the participants. There were also higher ratings of acceptability when the figure object was in the region of acceptability where the two spatial templates overlapped (these are the four upper right boxes in Figure 10c), compared to the exclusive acceptable regions for each spatial template.

This indicates that not only does it seem possible that multiple reference frames and spatial templates are active during processing, but also that composite spatial templates can be formed from the spatial templates aligned with different reference frames. The experiment also noted that there was a degree of individual differences in the ratings. Some participants preferred one reference frame over the other, whilst some were happy to use both reference frames as a guide to their acceptability ratings. In Experiment two, Carlson-Radvansky and Logan (1997) used a speeded sentence picture verification task to measure the use of reference frames and spatial templates. The stimuli remained the same: a reference object rotated 90° to dissociate the reference frames and a figure object placed in different regions of acceptability (for different reference frames) around the object. There were three groups of participants: those who were instructed to indicate the figure object was above/below the reference object if it was so according to either reference frame, those who were to indicate the figure object was above/below the reference object if it was so according to a deictic/environment reference frame, and those who were instructed to indicate the figure object was above the reference object if it was so according to the intrinsic reference frame. The dependent variables in the experiment were the time taken and accuracy to indicate the acceptability of a sentence as a description of the spatial relationship between the figure and reference object when the figure object was placed in different regions of acceptability around the reference object.

The results showed, firstly that people were faster to respond to above spatial relations than below. Such a finding has been reported before (Clark, Carpenter & Just, 1973). The results for participants who had to respond using both reference frames to determine the spatial relationship gave faster and more accurate responses when the figure object was located in the region that was deemed acceptable by spatial templates aligned with both reference frames. When participants had to respond using just one reference frame, either the deictic/environmental or the intrinsic, there was an effect of multiple spatial templates also. Participants took longer and were less accurate to judge the sentence as not applicable to the spatial relation if the figure object lay in a position that was acceptable for the non-selected reference frame. This indicates that despite explicit instruction to only use one reference frame there was still activation to multiple reference frames and so an influence on processing. It seems that multiple reference frames are active despite a conscious knowledge that only one reference frame should be used.

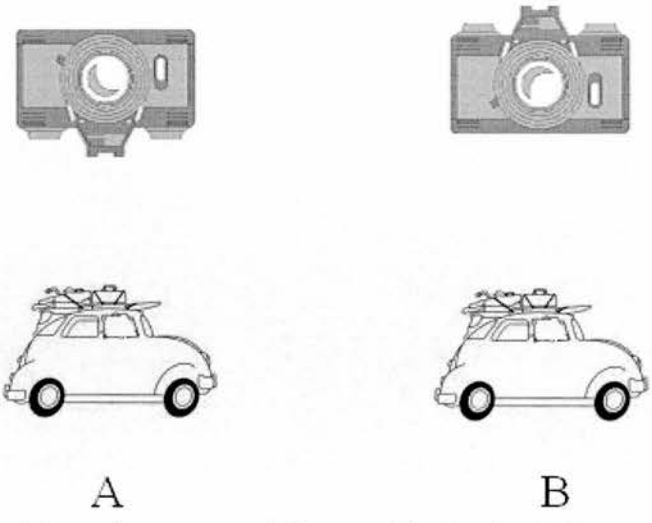


Figure 11. A). The reference and figure object orientation conflict. B). The reference and figure object orientation are aligned.

It is possible that reference frames are also applied to other objects in the scene that are not part of the spatial relation or even the figure object itself. Burigo & Coventry (2004) found that in scenes such as the one shown in Figure 11a where the figure object orientation conflicts with the reference frame orientation participants rated the acceptability of sentences such as *the camera is above the car* as less acceptable than when the figure object orientation is the same as the reference object orientation as in Figure 11b. This may be because participants are influenced by a reference frame imposed on the figure object.

However, Carlson et al (2003) obtained slightly different results when using an ERP methodology. Their behavioural data matched that of Carlson-Radvansky and Logan (1997), in that there was an influence of the reference frame that participants were not instructed to use on sentence verification time, indicating that the non-selected reference frame was still exerting some interference. The ERP data, on the other hand, showed that there was only evidence of competition when participants were instructed to use the intrinsic reference frame to make judgements and not the absolute/deictic reference frame (these two reference frames were coincident). When the participants were instructed to use the absolute/deictic reference frame, but not the intrinsic reference frame, there was no competition effect. The ERP data seems to suggest that there is activation of the absolute/deictic reference frame even when there is explicit instruction not to use this reference frame, but that it is possible to inhibit the intrinsic reference frame when instructed to do so. This is in contrast to the behavioural data reported both in Carlson et al (2003) and Carlson-Radvansky and Logan (2001). Carlson et al (2003) argue that the competition effect exhibited by the absolute/deictic reference frame when the instruction is to use the intrinsic reference frame represents non-conscious automatic activation of the reference frame because the conscious strategy ought to be to inhibit this reference frame.

In addition, there was also a condition in Carlson et al (2003) where participants were instructed to use *either* the absolute/deictic or the intrinsic reference frame in order to judge the veracity of the sentence. In this condition there was evidence of interference from the intrinsic reference frame shown in the ERP data. The crucial difference between this and the absolute/deictic only condition is that participants were instructed to use the intrinsic reference frame as well as the absolute/deictic reference frame in the former whereas in the latter participants were instructed only to use the absolute/deictic reference frame and therefore to suppress the intrinsic reference frame. Carlson et al (2003) argue that the interference exhibited by both the intrinsic and the absolute/deictic reference frames when participants were instructed to use either reference frame represents strategic conscious activation of reference frames. This is because there is no ERP evidence for interference by the intrinsic reference frame when participants were instructed to use the absolute/deictic reference frame only. Interestingly, the waveform, a slow frontal wave, was the same when the activation of the reference frame was automatic as when it was strategic, Carlson et al (2003) take this as evidence that the processes involved in activating a reference frame are the same whether it is strategic or automatic activation.

There are a great number of factors that influence reference frame selection. The evidence suggests that the environment provides the most influential reference system in the direction of gravity. Given its universal nature across all Earth environments and across time it makes most sense that humans have a possibly innate evolved grasp of the importance of gravity. Gravity, however, only operates in the vertical domain and in the horizontal plane the issue is much more complicated. There appear to be several factors that influence reference frame selection. Most of these are contextual effects such as the functional relatedness of the figure and reference objects (Carlson &Tang, 2000) and the layout of the environment (Franklin et al, 1992). There are also effects which would seem to occur earlier in processing such as previous use of a reference frame (or non-use) (Carlson-Radvansky & Jiang, 1998). It is also the case that multiple reference frames have an impact on processing even if they end up being non-selected (Carlson-Radvansky & Logan, 1997).

2.8. Spatial Language Summary

The issue for researchers in spatial language is to account for how a perceptual representation is mapped onto a linguistic representation so that we are able to talk about the location of objects in the world. The principal method proposed for how it is that humans achieve this is via the imposition of a reference frame which acts as a way of parsing space into regions around an object for production and as a method of directing attention from one object to another in terms of comprehension (Logan, 1994, 1995, Logan & Sadler, 1996). The process of spatial relation apprehension is argued to occur in several steps corresponding to processes that enable the translation of one representation into another (Jackendoff, 2002). A spatial or perceptual representation is translated into linguistic representation via processes that operate in a conceptual representation. The literature suggests that there is some form of conceptual reference frame which imposes a structure of space onto a scene (Logan and Sadler, 1996). This then is converted into a linguistic reference frame (Levinson, 1996, 2003) which contains information about the logical structure and various arguments in a propositional format and enables the translation of the information into syntactic structure and into a linguistic representation. The whole process is a translation of co-ordinate information about the spatial relations between objects, which is present in a perceptual representation into information that encodes the categorical relationships between objects which can be represented linguistically (Kosslyn et al, 1992).

Reference frames are a flexible representation and allow the perceptual representation to be encoded linguistically in numerous different ways by altering the settings of the parameters of the reference frame. The parameters can be set either automatically or strategically and there is competition between different settings of these parameters. The different settings of the parameters correspond to the three broad reference frames argued for in the literature, the absolute, intrinsic and relative/deictic. There are many known factors which influence the selection of a reference frame, function, environmental structure, goals of the speaker etc... However these have nearly all been discovered using a monologue situation and it remains important to investigate how reference frames are selected in more naturalistic uses of language.

2.9 Dialogue

The natural setting for language is dialogue (Clark, 1996). Everybody is able to participate in a conversation, but other types of language use such as giving lectures, writing novels and so on, require much practice and by many people are not mastered. Clark (1996) argues that dialogue is the primary use of language and that other forms of language are derivative from this. Therefore, a theory of language processing should explain how it is that language is processed in a dialogue context.

Many accounts of language have ignored dialogue in favour of investigations using monologue (e.g. Dell, 1986, Levelt, Roelofs & Meyer, 1999). From a practical aspect experiments which include just the language of one person are far easier to control. From a theoretical aspect dialogue has been viewed as the product of two monologues. The speaker produces language and the addressee comprehends what is spoken and then produces their own language. As a consequence of this view it has been implicitly assumed that theories of dialogue can be built from the combination of explanations of language production and language comprehension, which are largely built upon experiments involving a single person. However, there is much evidence to suggest that dialogue is an interactive joint process (Clark 1996) that is produced by all members of a conversation and that it is much more than two monologues put together. For example, Bavelas, Coates and Johnson (2000) have shown that the presence of an audience improves the ability of a speaker to tell a story, indicating that even in very monologic type situations listeners aid the speaker. It is also the case that addressees help complete the utterances of speakers when it appears the speaker is having difficulty, for example by offering possible words (Clark & Wilkes-Gibbs, 1986).

Clark (1996) has suggested that language is an interactive process which two interlocutors engage in. Over the course of the conversation interlocutors establish common ground and negotiate conceptual pacts. Clark and Wilkes-Gibbs (1986) showed that when interlocutors had to describe ambiguous tangram figures to each other the noun phrases used to describe the figures became less complex and shorter upon each subsequent exposure to the same tangram figure. For example, over the course of six trials one pair of participants went from describing one tangram as "All right, the next one looks like a person who's ice skating, except they're sticking two arms out in front" to "Um, the next one's the person ice skating that has two arms?" on the second trial, by the sixth trial they referred to the tangram as "The ice skater".

(Clark & Wilkes-Gibbs, 1986 p12). The interlocutors in this experiment were implicitly agreeing upon conceptualisations for each of the figures, but they were doing so interactively. A process such as this demonstrates how language is much more than two monologues put together.

2.10. Gricean Maxims

It is clear that in conversation sentences are interpreted within the context of the conversation rather than as they would be if they were uttered without a context. Addressees make assumptions about the utterances of a speaker and try their best to interpret the utterance as if it added to the conversation. Grice (1975) listed four maxims which addressees and speakers use to interpret and produce utterances. These maxims express the fact that in general people who are in conversation are being co-operative. The four maxims are: *The co-operative principle*: This is a general principle which says that people should provide contributions to the conversation which are required for the purpose of the conversation. This principle subsumes the four maxims below. *The maxim of quality*: This is the principle that people should say only that which they believe to be true. *The maxim of quantity*: This is the principle that people should provide as much information as required and no more. *The maxim of relevance*: This is the principle that people should provide contributions which are relevant to the situation. *The maxim of manner*: This principle states that people should be clear and easy to understand by being unambiguous, brief and orderly.

There is some evidence that people do not adhere to these four maxims all of the time (e.g. Brennan & Clark 1996), but Levinson (1983) argued that at a deeper level the maxims are adhered to even when they appear not to be. People strive to interpret sentences in terms of these maxims and assume that there is some degree of implicature when the maxims are flouted. For example, consider the following exchange

A: *Where's Bill?*
B: *There's a yellow VW outside Sue's house* (Levinson 1983 p102).

The answer of B seems to violate the maxim of relevance as the question had nothing to do with the location of a yellow VW. However, if A assumes that B is adhering to the maxim of relevance then A can interpret it as a response indicating that Bill owns a yellow VW and therefore may be at Sue's. Levinson (1983) also argues that the maxims can be violated in order to make points or to add linguistic emphasis. When a speaker does violate one of the maxims the addressee still assumes that the speaker is adhering to the co-operative principle and so assumes that the speaker is implicating a different meaning from what the speaker's utterance would normally imply; one that does convey some appropriate information. In addition, speakers must assume that addressees share enough *mutual knowledge* with them to enable the addressees to work out the implied meaning of the utterance.

2.11. Common Ground

Common ground is essential for two people to successfully take part in any joint activity. It refers to the representation of mutual knowledge; that is knowledge which all participants in a joint activity are aware of. The role of mutual knowledge is of the highest importance to explanations of dialogue. In order for an utterance to be understood the addressee must have knowledge about what the words of the speaker's utterance refer to. The speaker must also have knowledge that the addressee understands what the different words refer to, and the addressee must understand that the speaker understands that the addressee understands and so on ad infinitum. Clearly people do not represent such a list of infinite recursions in order to represent common ground. Clark (1996) has suggested that common ground is represented in a different manner. He calls it *shared common ground.* Something is considered to be part of common ground if every member of a community (this can be anything from two to an entire population) has a basis (details of what constitutes evidence is discussed in section 2.11.1) which indicates that every member of the community also has the basis and also gives the information that is part of common ground. In simpler terms there must be a basis which suggests to all members of the community that all members of the community are aware of the basis, and the basis must also give the additional information that is part of common ground. Clark (1996) likens common ground to self awareness, because members of a community must have common knowledge that the basis is common and that the information is common. Underlying the representation, however, there still lies an assumption by each individual that a basis exists for each piece of common ground.

> *The principle of justification.* In practice, people take a proposition to be common ground in a community only when they believe they have a proper shared basis for the proposition in that community. (Clark, 1996 p96)

It is the fact that the common ground rests upon the belief that something is common ground which leads to confusions occasionally, when something is believed to be common ground, but is not. When people lie they are taking advantage of this belief by deliberately manipulating the basis which is available to community members.

2.11.1. Evidence that something is part of Common Ground

There are a large number of different types of evidence which count towards being a basis for shared common ground. These can be social, as in the knowledge that someone belongs to a particular group, or perceptual. People make the assumption that others have similar perceptual apparatus and so things which are perceptually salient will be part of common ground. Clark (1996) therefore identifies three different types of common ground: 1) Initial common ground. 2) Current state of the joint activity. 3) Public events so far.

Initial common ground is the knowledge which is assumed or inferred to be possessed by a current partner in a dialogue. This knowledge can be inferred from our background knowledge about people who are members of a given community. At a basic level, a person can suppose that a native English speaker will have a larger English vocabulary than many non-native English speakers and so will understand the meaning of more words and idioms which are in the language. If a person is known to come from a particular town then it can be assumed they have some knowledge of the geography of that town (Isaacs & Clark, 1987). It can be expected that people who are of the same nationality possess similar knowledge about how to behave in given public situations. For example, that in Britain people drive on the left hand side of the road. It also means people who are not members of communities can infer what types of things other people will know. If a person is a surgeon, a non-surgeon may not have the same medical common ground, but will know the types of information that the surgeon will know. This is communal common ground, but we also possess personal common ground which is the history of interactions and knowledge about different people we know. Clark (1996) argues that it is this which separates friends from acquaintances. We have a great deal of personal common ground with friends, more with close family members, but probably less with the local shop owner.

The current state of the activity is a representation of the state the activity is at. At the beginning of a conversation the state is the initial state, but as each turn of the interaction takes place the state advances. For Clark (1996) a dialogue is a joint activity and is more than just the language of the participants. He describes it as having things in common with dancing with a partner or playing chess, both of which are also joint activities which contain different amounts of linguistic interaction. In order to fully understand the meaning of a dialogue it needs to be interpreted in terms of the context it occurs in and this includes many non-linguistic parts of the joint activity, for example the position of participants relative to each other, the goals of participants, the gesticulations of participants. The part of common ground which Clark (1996) describes as the current state of the activity are external representations which mark the state of affairs the joint activity is at. In a football match this includes things such as the scoreboard which represents the score of the match; the ball's location is also an external representation accessible to players on the pitch (and spectators). A dialogue often has some sort of goal which the participants wish to achieve, although sometimes the goal may be vague (e.g. killing time or catching up with a friend). The current state of events may be represented externally and may mark progress towards the achievement of the goal. For example, two people may be having dinner and the things on the table will represent the current state of the activity if the dinner is in a restaurant a menu on the table indicates that the meal has not yet started. External representations provide information which are accessible to participants and therefore can be assumed to be in participants' common ground.

The type of common ground Clark (1996) describes as public events so far are a representation of the events which have publicly occurred so far during the joint action. This is probably most often recorded as a mnemonic representation of the events of the joint action, but could also be external as in the case of minutes at a meeting, or court records. In the case of a dinner between two people each participant will probably keep a record of what has been spoken about, how they arrived at the restaurant, where the exit is and so on.

2.12. Use of Common Ground

Common ground is important for a conversation to be successful. The internal representations of common ground which two people share must be similar enough, though they need not be identical. There is, however, some degree of egocentricity displayed in dialogue. People are more likely to assume a member of a community knows a piece of information if they are also a member of that community and know the information (Ross, Greene & House, 1977). Similarly, in goal driven experiments people are likely to assume that their goal is shared with a partner even when there is evidence that this is not the case (Russell & Schober, 1999).

2.12.1. Common Ground in Production

It remains a question as to whether interlocutors do engage in audience design: Do speakers use common ground when formulating utterances, and if they do, whether they do so initially or after monitoring. It remains possible that speakers formulate their utterances based upon egocentric knowledge exclusive to them, but for the majority of time this happens to also coincide adequately enough with an addressee's knowledge (Horton & Keysar, 1996). It is only when there is a clear violation of what is known by an addressee that the speaker invokes common ground. According to Horton and Keysar (1996) the initial utterance is based solely upon knowledge which is known to the speaker. This knowledge can either be known also by the addressee or be exclusive to the speaker. The initial utterance is monitored and alterations are made if the utterance violates common ground. An internal monitor is thought to monitor speech for the existence of errors prior to production (Levelt, 1989), and Horton and Keysar (1996) propose that this is also extended to monitor for audience design.

Horton and Keysar (1996) investigated the different predictions made by the *initial design* model, which proposes that common ground is taken into account when the message is planned, and the *monitoring and adjustment model,* which proposes that initial utterance planning is based on egocentric knowledge and monitored for contradictions with common ground. They did this by manipulating the contrast set of a target object which was to be described by a participant. The speaker had to describe a target object (e.g. a circle) which was going to move from their side of the computer screen to the listener's (played by a confederate) side of the computer screen (the screen was separated by a barrier). At the bottom of each side of the screen (on both the speaker and listener's side) there was a contrastive object which was the same as the target object only larger or smaller (e.g. a larger circle). In this case the speaker could describe the target object as a small circle. Importantly, for some of the trials there was no contrastive object on the listener's side of the screen. In this case, the contrastive description of small circle is not appropriate for the listener and the use of it violates common ground and the Gricean maxim of quantity. Horton and Keysar (1996) found that speakers used the contrastive description to describe the moving object even when the contrast set was not part of common ground, but only when there was time pressure on producing the utterance. They present this as support for the monitoring with adjustment model suggesting that the initial utterance is formed using privileged knowledge only, and when time pressure is applied there is no time for monitoring and repair to occur and so the utterance is egocentric.

2.12.2. Common Ground in Comprehension

The results of Horton and Keysar (1996) support the claim that initially utterances are produced on the basis of egocentric knowledge and only later repaired on the basis of common knowledge. Keysar et al (2000, Keysar et al, 1998) used an eye-tracking paradigm and showed that addressees interpret descriptions based upon egocentric knowledge and not common knowledge. This study again used contrast sets to dissociate common and privileged knowledge. The speaker (again a confederate) could see two objects, for example two candles. One of the candles was bigger than the other and so the smaller of the two could be described as the small candle. The addressee, however, could see three candles, two which the speaker could see and one which was smaller still. Therefore, the small candle described by the speaker was actually the middle sized candle for the addressee. When the speaker asked the addressee to move the small candle, the addressee still initially considered the smallest candle as a candidate referent even though it was only part of the addressee's knowledge. The results of Keysar et al (2000) suggest that addressees also initially use an egocentric strategy to interpret a speaker's utterances and only later in processing consider a strategy based upon common ground. Hanna, Tanenhaus and Trueswell (2003) did, however, show some evidence of early use of common ground to understand a speaker's utterance, but also showed that information which was exclusive to the addressee affected processing at an early stage. They argue for a weaker version of the monitoring with adjustment model that is probabilistic in nature, where the salience of the exclusive information affects the weighting and likelihood of it affecting understanding early in processing.

2.13. The Interactive Alignment Model of Dialogue

Regardless of the way in which common ground is incorporated into processing the above studies all rely, at least implicitly, on a representation of common ground that is something like that proposed by Clark (1996). That is, a representation of common ground that is explicit and updated during the course of conversation. According to Clark (1996) each turn and additional phase of a conversation adds to the common ground that is shared by the participants. Therefore, updating common ground is a continuous process, and an individual must keep track, constantly monitoring information which has since become common ground. The explicit modelling of common ground is a cognitively expensive process; it requires holding a model of that which is known exclusively and also that which is known mutually. It is partially the cognitive expense that is associated with representing common ground which led to Horton and Keysar (1996) proposing the monitoring with adjustment model.

However, there are theories which can negate the need to explicitly represent common ground. To an extent this is what the strong version of monitoring with adjustment is attempting to do. Rather than use common ground participants use the heuristic that their own knowledge of a situation is likely to be very similar to that which is shared by other people in the same situation and so rely solely on that in dialogue. However, if this initial stage is monitored for whether the utterance violates common ground, then it is just as cognitively expensive, because a model of common ground is required to monitor the utterance. Pickering and Garrod (2004) have proposed the *interactive alignment model of dialogue*.

The interactive alignment model (Garrod & Pickering, 2004, Pickering & Garrod, 2004) is based on the multitude of findings which report that interlocutors over the course of a conversation come to talk about things in a similar fashion (e.g. Branigan et al, 2000, Brennan & Clark, 1996, Cleland & Pickering, 2003, Levelt & Kelter, 1982). The argument is that over the course of a conversation interlocutors align the different representations associated with speech production and comprehension, so that essentially each interlocutor is using the same representations as each other to produce and comprehend utterances. Speech production is known to take several different stages from intention to articulation, to echo the title phrase of Levelt's (1989) seminal work (Dell 1986, Levelt 1989, Levelt et al, 1999). These different stages are different representations each of which performs a different function. The exact processing details vary from theory to theory, for example whether or not the process is heavily modular or whether it is feed forward only (Dell 1986, Levelt et al 1999), but all agree that there is some form of conceptual level where the message is represented, followed by a grammatical level which encodes the sentence structure and then a phonological level where the sounds are implemented, and finally the articulatory plan which encodes how the muscles will move to form the speech. It is these different levels which are argued to be aligned between interlocutors partaking in conversation together.

2.13.1. Lexical Alignment

It is a well known finding in dialogue research that interlocutors come to talk in the same way over time. Brennan and Clark (1996) showed that participants aligned lexically during a conversation. In their experiments participants were put into pairs, one of which was the director and the other the matcher. The director had a set of picture cards laid out in front of them in a grid. Their goal was to describe the order of their cards so that the matcher could put their identical set of picture cards into the order on the director's grid. Over the course of an experiment, pairs repeated the task with the same and different sets of cards, so that each pair sorted the same deck up to four different times. The results showed that directors tended to persist in their use of referential terms for items. There was a large proportion of consistency between the terms used to refer to an item the first time and on each subsequent occasion. In addition, there was greater within pair consistency for referential term use than between pair consistency. Brennan and Clark (1996) argue that the reason for this consistency is that interlocutors form temporary partner specific conceptual pacts. These are tacit agreements, contributed to by both interlocutors, on lexical terms used to refer to items. To support this, they showed that interlocutors would violate the Gricean maxim of quantity in order to maintain a conceptual pact. For example, in the presence of two potential referential candidates for the basic term *shoe*, directors may choose to use the subordinate term *pennyloafer* to distinguish between the two candidates. However, on a subsequent occasion, when only the pennyloafer is present, the director would still persist in the use of the subordinate term, even though the term *shoe* would now have been sufficient. The use of the subordinate term was, however, partner specific. When the director had a new matcher, the director was more likely to revert back to using the basic level term. This is evidence that interlocutors tend to use the same lexical terms as each other and persist in the use of lexical terms over the course of a conversation.

2.13.2 Syntactic Alignment

There is also strong evidence showing that interlocutors align syntactically. Levelt and Kelter (1982) phoned shops and asked them what time they closed. They worded the enquiry either "What time do you close?" or "At what time do you close?" (The actual investigation was carried out in Dutch). The results showed that people answered in a form that mirrored the question, either "Five o'clock" or "At five o'clock" respectively. The alignment in Levelt and Kelter's (1982) study may have been due to lexical repetition of the word *at* and so may not necessarily have been syntactic alignment. However, other studies have shown syntactic alignment independent of lexical items. Branigan et al (2000) used a confederate priming paradigm where a confederate described a picture on a card (*prime*) and the participant had to choose which card matched the description. The participant then described the picture on their card (*target*) under the impression that the confederate was choosing which card matched the description. The experimental descriptions given by the confederate were all descriptions using a di-transitive verb (e.g. *give*). These are verbs with three syntactic arguments (often indicating the transfer of an object from one person to another) and can typically be used in two different syntactic constructions; a double object construction (e.g. *the chef giving the waiter the plate*) or a prepositional object construction (e.g. *the chef giving the plate to the waiter*). These constructions differ very little difference in their meaning. Branigan et al (2000) found that participants were more likely to use a construction after they had heard that construction being used by the confederate, than after they heard the confederate use the alternative construction. Furthermore, Branigan et al (2000) also reported that the participants were more likely to use the same syntactic construction as the confederate when there were overlapping lexical items between the prime and the target than when there were not. In this experiment, the overlapping lexical item was the verb describing the transfer. There was a greater amount of alignment when the verb was the same between prime and target compared to when it differed.

Pickering and Garrod (2004) argue that this is because alignment at one level of representation increases the amount of alignment at other levels of representation. In the case of Branigan et al (2000) lexical repetition is leading to an increased amount of syntactic alignment. The inference extended from this is that the more things in common between the prime and the target the greater the amount of alignment will be exhibited.

Cleland and Pickering (2003) have shown increased syntactic alignment in a confederate priming experiment when nouns in the prime and target were semantically related. There experiments investigated priming of pre-nominal descriptions (e.g. *the red sheep*) and relative clause descriptions (e.g. *the sheep that's red*). As in Branigan et al (2000) participants were more likely to use the same syntactic structure as the confederate, but this tendency increased when the nouns were semantically related. For example when the confederate described *the sheep that's red* the participant was more likely to use the relative clause when describing the goat that was red as opposed to a knife that was red.

Syntactic priming has also been shown between languages with bilingual interlocutors. Hartsuiker, Pickering and Veltkamp (2004) used a confederate priming paradigm and showed that participants were more likely to use an English passive sentence (e.g. *the truck is being chased by the taxi*) after hearing the confederate use a Spanish passive than after hearing the confederate use an active construction.

2.13.3. Conceptual Alignment

There is also evidence that interlocutors align the way they conceptualise the world. Garrod and Anderson (1987) had participants take part in a game where they each had to traverse a maze, but doing so required the co-operation of both interlocutors. Each participant saw the same maze on a computer screen (they could only see their own screen), but each maze differed slightly in the presence or absence of gates and switches. The gates on each screen could only be opened or closed by switches on the other person's screen and so opening a gate on one screen required a person's partner moving to a switch on their screen. Navigating through the maze therefore required the co-operation of each person. The mazes were grid like, formed from a number of boxes with connecting routes between them. There are many different ways that participants could talk about the maze, but the descriptions generally fell into four different categories, Path, Line, Figural and Co-ordinate. Path descriptions described the route to be taken from present location to a goal (e.g. *go along two then up one*). The line descriptions described the line and column people were on (e.g. *I'm on the top row second column in*). The figural descriptions divided the maze up into areas (e.g. *at the top of the left indicator*). Finally, co-ordinate descriptions labelled the maze using a grid like system (e.g. *I'm at C2*).

The results showed that once a system of describing the maze was agreed upon pairs tended to stick to that system. The agreement process was not overtly negotiated. Individuals in the pairs adopted different roles where one was a leader who initiated corrections to the other person, but the agreement on which descriptive system to use was implicit. There was also much greater between-pair than within-pair variability in the descriptive systems used: Members of pairs tended to describe the maze in the same way. The description of the maze, Garrod and Anderson (1987) argue, is a reflection of the way the individual has built a mental model (i.e. a conceptual representation of maze) of the maze. Therefore, pairs converging upon their descriptions of the maze is taken as evidence that during the course of the conversation they are aligning mental models. Garrod and Anderson (1987) also show alignment at levels of representation other than conceptual. There are several ways of describing the various parts of the maze, for example, the boxes could be described as boxes or nodes, and the use of the terms columns and rows differed from pair to pair. However, partners also lexically aligned, in the same fashion as Brennan and Clark (1996) showed, so that both used the same term to describe the same thing.

Garrod and Clark (1993) extended this work to show that conceptual alignment occurs at all ages. They had school children perform the maze task and found that they too showed alignment of description types and lexical items. However, there were some differences from the adults; younger children appeared to show evidence of aligning only superficial features of the lexicon and description scheme. It was only the older children that had deep levels of alignment which was important for achieving mutual understanding of the dialogue. This was evident because the success rate of the exchanges in conveying the intended meaning was much lower for younger children than older children. Younger children would co-ordinate, but were unable to change to other forms of description when it was clear that the selected method of description was failing.

2.13.4. Alignment at other levels of representation

The evidence is strong that interlocutors align representations at several levels. In addition to the alignment of lexical, syntactic and conceptual representations, interlocutors also align speech rate (Bard & Aylett, 2004), accent (Giles & Powesland, 1975) and other non-linguistic behaviours such as, foot rubbing or touching one's face (Chartrand & Bargh 1999), respiration (McFarland, 2001) and posture (Shockley, Santana, & Fowler, 2003). Pickering and Garrod (2004) take this evidence as support for the interactive alignment model of dialogue. According to this theory successful dialogue is achieved when the representations of each interlocutor are in alignment. Alignment is achieved without resort to overt explicit negotiations (Garrod & Anderson, 1987), except in rare circumstances when the normal alignment mechanisms fail.

2.13.5. Implicit Common Ground

Alignment of representations leads Pickering and Garrod (2004) to argue for *implicit common ground*. This is in direct contrast to Clark (1996) who argues that the representation of common ground is explicit. Implicit common ground does not involve mutual knowledge, rather, it is incidental and arises from automatic alignment of representations. As alignment occurs between interlocutors the representations they use to produce and comprehend utterances become very similar. Essentially both are using egocentric perspectives, but alignment leads to both people's egocentric representations being sufficiently the same that it is as if they were using common ground. The advantage of implicit common ground is that it is cognitively inexpensive, alignment leads to interlocutors' utterances being produced using the same representations and therefore it operates in the same manner as if each were using common ground to form utterances.

2.13.6. Mechanisms of Alignment

Alignment is argued to occur via a resource free priming mechanism that operates through direct channels between the different levels of representation. This occurs via input-output co-ordination. The same processes comprehend speech as produce speech and therefore there is co-ordination between the two systems. The resource free nature of alignment allows for a cognitively inexpensive mechanism to account for common ground. The channels of alignment are argued to be direct because the speech production and comprehension system is modular; with each level interpreting and producing specific kinds of representations. Therefore, although all information is present only in the speech stream each level of representation interprets only the representation which it can produce.

The interactive alignment model explains how common ground can be established implicitly between two people at a very local level, but can it establish co-ordination at a more widespread scale in the sense of Clark's (1996) communal common ground? The evidence seems to suggest it is at least possible. Garrod and Doherty (1994) conducted a version of the maze task where 15 dyads were formed to play the game at the same time. After each game was completed partners swapped so that each person was now playing with someone new from the community of dyads and the conventions established with their old partner would have to be re-established with their new one. Garrod and Doherty (1994) found that very few partner changes were required before the entire community had converged upon single way of representing the maze. Similarly Barr (2004) has modelled the process of alignment across a population of artificial agents and found that very few interactions between partners are required before global convergence is achieved. Barr (2004) also found evidence of local 'dialects' occurring in groups of the artificial agents when interaction was more likely with recent partners.

Taken together this is evidence that a large amount of consensus and convergence can be gained from cognitively inexpensive alignment. The notion of implicit common ground also bypasses any problems with respect to iterative representations or reflexive representations (Lee, 1999) and is therefore a strong alternative to explicitly represented common ground (Clark, 1996).

2.14. Partner Specific Effects

One of the predictions of low level priming as a mechanism of alignment is that there is little conscious control over the process. As a consequence, there should be little effect of partner specificity (but see Horton & Gerrig, 2005b, who suggest a mechanism that can account for automatic effects of partner specificity). For example, if two interlocutors lexically entrain upon a referential term then they should be more likely to use that term when they switch to speak with new partners.

There is evidence to suggest that there are no partner specific effects which determine the intelligibility of expressions in dialogue. It is known that the first time a word is produced it tends to be clearly pronounced, but on the second production of the word it is less intelligible because it is *given* information (Fowler & Housum, 1987). Bard, Anderson, Sotillo, Aylett, Doherty-Sneddon, and Newlands (2000) have shown that what matters is that the information is *given* for the speaker not the addressee. The intelligibility of the expression is reduced regardless of whether or not the addressee has heard the expression before.

Similarly, Barr and Keysar (2002) have shown no effect of partner specificity in comprehension for lexical alignment. When addressees hear a speaker describe a picture of a flower as a *carnation* and a picture of a car as a *sports car* they become entrained and expect a speaker to continue using those lexical terms. If the addressee then hears the speaker use the term *car* to refer to the sports car the addressee experiences a moment of confusion as they initially expect the word *car* to be the onset of the word *carnation*. However, if this is the first time the addressee has heard the speaker use the word *car* then there is no period of confusion. The addressee expects the carnation to be referred to as a *flower* because there has been no lexical entrainment, therefore the addressee expects the speaker to use basic level terms to refer to the objects (e.g. *car* and *flower*) as opposed to the sub-ordinate terms (e.g. *sports car* and *carnation*; Rosch, Mervis, Gray, Johnson, & Bayes-Braem, 1976).

Barr and Keysar (2002), however, had one speaker use the initial subordinate terms (*carnation* and *sports car*) and then a second speaker use the term *car*. They found that addressees still responded initially with looks towards the carnation despite the fact the speaker was different to the one the addressee had entrained on the terms *carnation* and *sports car* with. Barr and Keysar (2002) argue that this is evidence that interlocutors do not show effects of partner specificity and instead use an egocentric perspective to comprehend dialogue.

Metzing and Brennan (2003), however, have shown effects of partner specificity in comprehension. In this experiment, a director (a scripted confederate) gave instructions to the matcher (participant) to move objects around a grid. The matcher therefore entrained on the names the director was giving to objects (e.g. *shiny cylinder*). The director was then either changed or remained as the original director, and the director either used the original term to refer to the object or a new term. When the director used the original term there was no difference in the time to look at the target object whether the director was the original one or new one. This is analogous to the effect obtained by Barr and Keysar (2002). However, when the term used to refer to the object was new (e.g. *silver pipe*) addressees were slower to look at the target object when the original director used the term compared to when the new director used the term. Participants in this experiment demonstrated partner specific effects and so were slower, both to look at an object and to move an object, when the original director used a *new* term compared to when a new director used a *new* term.

Horton and Gerrig (2005a, 2005b) have argued that interlocutors can show partner specific effects without having to explicitly represent common ground. This would be compatible with the notion of implicit common ground obtained through alignment of representations (Pickering & Garrod, 2004). Horton and Gerrig (2005a) argued that implicit associations between information and different partners can lead to partner specific effects; information can become associated with a specific person so that when in a conversation with that person the associated piece of information is more likely to be active in memory.

Despite the debate over when partner specific information is considered in the formation of utterances it is the case that for most situations an egocentric perspective can be adopted and used to form utterances which are understandable. This is especially the case if the egocentric perspectives of the two interlocutors are over the course of the conversation aligned in the manner described by Pickering and Garrod (2004).

2.15. Spatial language and Dialogue

In many respects the same issues surrounding language in a dialogue context also are present for spatial language in dialogue. Of much importance is the issue concerning which perspective to take upon a scene. With regards to non-spatial language, perspective is used in a very broad sense, referring to the other person's knowledge base as a whole. When interlocutors take their partner's perspective they are interpreting or producing an utterance with their partner's knowledge in mind (although as mentioned previously there is an issue as to whether interlocutors do take each other's perspective at all). When it comes to spatial language the issue of which perspective is adopted is literally spatial. This is because interlocutors can physically have different perspectives on an environment, and they can choose which perspective to describe the location of an object from (e.g. does a speaker describe an object as on their left or on their partner's right). There has been relatively little work investigating spatial language in dialogue, at least in comparison to the amount of work on spatial language and on dialogue as individual areas. Most of the work that there has been has investigated whether or not speakers use an addressee's (allocentric) perspective or a speaker (egocentric) centred perspective when describing the location of objects.

2.15.1. Perspective Taking

Imagine a scene as in Figure 12: A and B are standing facing the side of a car. B is looking for his ball which he kicked in the direction of the car. A who was standing at the car probably knows where the ball is and so B asks, "Did you see my ball?" A (understanding the implication of the question) replies "Yes. It's behind the car". From the position where A and B are standing the ball can either be located at point 1 if the utterance was using a relative reference frame or at point 2 if the utterance was using an intrinsic reference frame.

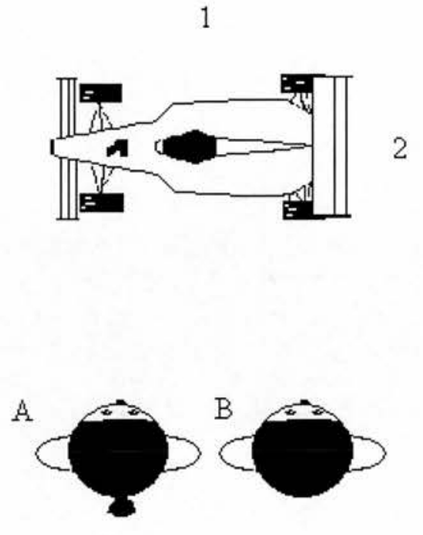


Figure 12: The ball can either be in location 1 or 2 if the description is "The ball is behind the car" depending upon which reference frame is used.

In order for B to understand where the ball is he must use the correct reference frame to interpret A's utterance. In this situation an incorrect interpretation will have little cost to B, but in others, for example when asking for directions, the cost could be greater.

However, the situation is more complex than this because interlocutors can be viewing a scene from different perspectives. As an example of the problem let us take the situation described in Figure 13: Here we see that B has lost his ball again. This time, however, he is viewing the scene where he has lost his ball from a different perspective to A. Once again, A is aware of the ball's location and B asks her where his ball is. The conversation could go like this:

B: *Did you see my ball?*

A: *Yes. It is behind the car and to the right*

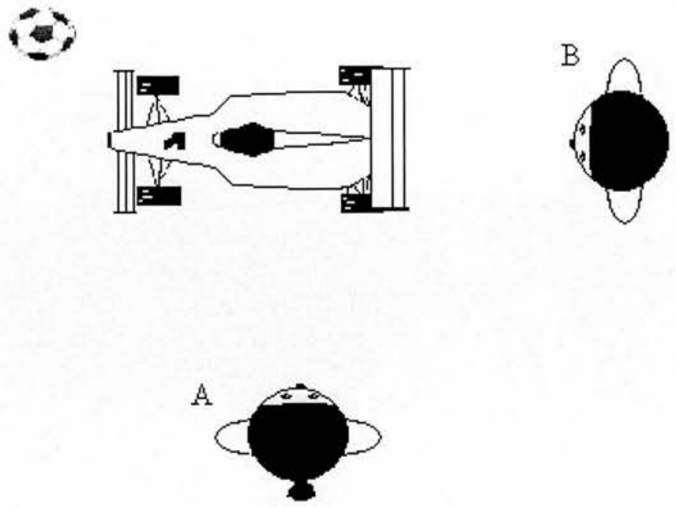


Figure 13: The ball is in front of the car according to an intrinsic reference frame. The ball is behind and to the right based upon a B's relative reference frame and behind and to the left based upon A's relative reference frame.

In this utterance A is using B's perspective. However, A could have said *it is behind and to the left*; in which case she would be using her own perspective. Alternatively A could have said *It is in front of the car*; in which case she would be using an intrinsic reference frame and using neither her nor B's perspective. As in the situation described for Figure 12, B must choose the correct perspective and reference frame in order to know where his ball is.

As a result of the potentially different viewpoints interlocutors can have on a scene, it is a question for spatial language research as to which reference frame a speaker is likely to use to produce an utterance and that an addressee is likely to use to comprehend an utterance. On the one hand, it is possible that a speaker could adopt an egocentric perspective, which should be cognitively easier as this is the viewpoint from which he experiences the scene. On the other hand, the speaker could adopt an addressee's perspective which would lessen the cognitive effort required by the addressee to interpret the scene. Of course, this is assuming that they are both viewing the scene from a different perspective. If they are viewing it from the same perspective then the description will most likely be from both interlocutors' perspectives (i.e. the difference between a speaker and an addressee centred perspective is only important if the two perspectives differ).

Buhl (2001) asked participants to watch a video which described the layout of an environment containing buildings, trees, phone boxes etc... The participants then had to imagine standing in a certain position in the environment and describe a route to one of the objects in the environment to an imaginary second person. The results showed that participants were more likely to describe the route from the perspective of their addressee. However, in this experiment the addressee was only imaginary and therefore could not give any feedback, and as a consequence it is not clear how generalisable this is to real dialogue. The importance of the addressee being able to give feedback on understanding has been shown to be important for many non-spatial dialogue situations (e.g. Bavelas et al, 2000, Clark & Krych, 2004). Investigations on spatial language have also shown that speakers are more likely to use an addressee centred perspective when talking to a tape recorder compared to when talking to a real partner able to give feedback (Schober, 1993).

2.15.2. Partner Effects on Perspective taking

It is known that using a reference frame on one occasion increases the difficulty of using an alternative reference frame on a subsequent occasion (Carlson-Radvansky & Jiang, 1998). This, however, was shown in a monologue situation and it is a question how one interlocutor's use of a reference frame will influence another interlocutor's use of reference frames. Interlocutors work together to agree on how it is that other forms of language are used to refer to objects (Clark & Wilkes-Gibbs 1986), and interlocutors align other linguistic representations to reach tacit agreements about how to conceptualise the world (Garrod & Anderson, 1987, Pickering & Garrod, 2004). It is therefore a strong possibility that interlocutors will have an effect on how each other describes the locations of the objects in a scene.

Schober (1993) has shown that interlocutors are more likely to use their addressee's perspective when describing the location of objects. In this experiment, one participant was assigned the role of director, whose job was to describe which of two circles was marked to the other participant, who was assigned the role of matcher. The director saw a large circle which contained two smaller circles, one of which was indicated with an arrow. The perspective which the director and matcher were viewing the scene from was also indicated (in this experiment the director and matcher could either view the scene from the same position, offset by 90° or offset by 180°). The director had to describe which of the two circles was indicated by the arrow to the matcher, so that the matcher could mark the correct circle on their sheet. Schober found that directors were more likely to use the matcher's perspective (as the director is using the matcher's perspective this is also an addressee's perspective) when describing the circle. However, directors did not make exclusive use of the matcher's perspective. Instead they switched between their own perspective and the matcher's perspective. In one condition, there was no matcher present and instead the director was told to describe which circle was marked to a tape recorder as if it were a matcher. This resulted in the directors producing more matcher centred perspectives probably as a result of there being no feedback to indicate understanding. Therefore, the director is under pressure to make the utterances as simple and easy to understand as possible.

The evidence is suggestive that speakers are more likely to use an addressee centred reference frame. How does this influence which reference frame the addressee uses if they wish to make a reply? Interlocutors align many levels of representation but do they align the representations which underlie spatial descriptions? The question is not so simple because there are different levels which underlie a spatial description. Consider again the situation in Figure 13:

B: *Did you see my ball*?

A: *Yes. It is behind the car and to the right*

Here A is using the matcher's perspective and also the addressee's perspective and is using a relative reference frame because the description is dependent upon B's point of view. B, however, is still not sure of the ball's location and so queries the description. Now B has several options as to how to clarify the question. B could say:

B: *Behind and to the right?*

Here B is aligning with A in the sense that both were using the matcher's perspective. However, he is using an egocentric perspective, whereas A used an allocentric, addressee centred perspective and therefore in this sense B is not aligning with A.

B: *To the left of the car?*

Here B is using A's perspective and therefore an addressee's perspective and so he is aligning at this level because A in her first description also used an addressee centred perspective (although who the addressee is has switched). However, B is using a director centred perspective whereas A used a matcher centred perspective. Like the above option this is a description using a relative reference frame and so is aligning at this level, although this description is relying upon A's not B's viewpoint. B could also even say:

B: *You mean in front of the car?*

Here B is actually changing to an intrinsic reference frame and so is not aligning with A at any level at all (Although he is still using the same reference object and it is a possibility that interlocutors also align which reference objects they use).

In the above exchange, A and B can align and both use the matcher's or the director's perspective when describing the location of the ball (this is equivalent to both using A's perspective or both using B's perspective). This is *aligning at the task level* (Schober 1995). When individuals do this they are making things easier for one of the individuals with a specific task. A and B can also align by both choosing to use either a speaker or an addressee perspective. This is *aligning the choice of perspective*. It is important to remember that the role of director and matcher remains constant throughout the exchange, whereas the role of speaker and addressee changes with who is speaking. Klabunde and Porzel (1998) suggest that the relative social status of the interlocutors and perceived linguistic ability has an effect on whose perspective is used. A and B could also align reference frames and choose to always use an intrinsic or always use a relative reference frame. This is *aligning reference frames*. The latter is closely related to aligning choice of perspective, but can be different depending upon how reference frames are defined. If Levinson's taxonomy of reference frames is used, then aligning reference frames is independent of aligning perspective because a relative reference frame can be based upon any person's perspective and an intrinsic reference frame can still be based upon the internal axes of a speaker or an addressee. If the Traditional taxonomy is used, then perspective alignment and reference frame alignment are closely related because both are determined by the directional axes of one person or another (although a deictic reference frame is sometimes argued to encompass both speaker and addressee centred descriptions Levelt, 1989).

The task in Schober (1993) did not involve the participants switching roles between the director and matcher. However, matchers did ask for clarifications of the descriptions which directors gave. As reported above directors tended to use their addressee's perspective in descriptions; this is also using a matcher centred perspective. When the matchers replied, they also used an addressee's perspective; in this case this is a director's perspective. So it appears from Schober (1993), that in this task, participants align levels of perspective and not at the level of task.

Schober (1995), however, found that participants aligned at the task level, not the choice of perspective. In this experiment participants were once again assigned the role of director and matcher. The director was presented with a large circular display which contained a number of objects of three different kinds. Fifteen of these objects were numbered 1 -15, and the director had to indicate which ones had which numbers, so the matcher could mark them on their identical circular display (minus the numbers). Again directors used the addressee centred perspective, and matcher centred perspective. However, this time when the matchers asked for clarification of the location of an object the matchers used their own perspective. This is a speaker centred perspective, but still a matcher centred perspective. In this experiment interlocutors were aligning at the task level, not at the level of perspective.

The reason for these opposite results between Schober (1993) and Schober (1995) may lie in the complexity of the tasks involved. The displays used in Schober (1995) were much more complicated than those used in Schober (1993). As a consequence it is possible that the matcher's task was perceived to be more difficult in the later study. Both interlocutors might have realised this and so sought to minimize difficulty for the matcher by always using their perspective.

Schober (1993, 1995, 1998) interprets both sets of results as being indicative of interlocutors co-ordinating the use of perspective in conversation. Unlike Pickering and Garrod (2004), who interpret alignment during dialogue to be due to low level processes, Schober (1995) argues that the alignment is due to conscious strategies on behalf of the participants. Either way, the results are suggestive that interlocutors do align representations that underlie the production of spatial descriptions, but there are still many questions regarding the issue. For example, what contextual factors may contribute to interlocutors aligning different kinds or aspects of representations that contribute to spatial language? Do interlocutors align reference frames in the sense defined in the spatial language literature? (e.g. Levinson, 2003, Miller & Johnson-

Laird, 1976). Does alignment occur in a transient trial by trial fashion as with alignment of other representations in dialogue? (e.g. Branigan et al, 2000) and can the alignment be attributed to low level factors as suggested by Pickering and Garrod (2004)

Chapter 3: Alignment Experiments

3.0. Chapter Overview

This chapter describes three experiments that investigate whether interlocutors align reference frames during dialogue. There is a large amount of evidence showing that interlocutors come to talk in the same way during a conversation. They use the same syntactic structures (Brangian et al, 2000); the same lexical items (Brennan & Clark, 1996); and build the same mental models as each other (Garrod & Anderson, 1987). Pickering and Garrod (2004) have argued that this reflects alignment of representations, which leads to implicit common ground being built up between interlocutors.

The three experiments described in this chapter used a confederate priming paradigm, where a scripted confederate described the location of a dot in relation to a reference object to a naïve participant, who then also described the location of a dot in relation to a reference object. Experiment 1 investigated whether or not interlocutors align the entire reference frame during dialogue. The results were suggestive that interlocutors do indeed align reference frames, but that this may be due to lexical alignment of the prepositions. Experiment 2 investigated whether or not interlocutors align just the axes of reference frames during dialogue. The results showed that interlocutors did align the axes of reference frames, and that this was independent of lexical alignment. However, it was noted that in Experiments 1 and 2 the participants were describing the left-right intrinsic axis in a different manner to the confederate. This may have been the cause of the lexical alignment effect in Experiment 1. As a result, in Experiment 3 the confederate's use of the left-right intrinsic axis was changed in line with participants' interpretations from Experiments 1 and 2. The results showed a significant effect of alignment of reference frames that was independent of any lexical alignment. This is argued to be strong evidence that interlocutors do align reference frames and is interpreted as support for the Pickering and Garrod (2004) interactive alignment model of dialogue.

3.1. Introduction

Research on dialogue has shown that interlocutors tend to align their utterances. That is, over the course of a conversation interlocutors will tend to communicate in a similar fashion to each other. Alignment occurs at several levels of communication, including: conceptual (Garrod & Anderson, 1987), lexical (Brennan & Clark, 1996) and, syntactic (Branigan et al, 2000).

Pickering and Garrod (2004) put forward the interactive alignment model of dialogue as an explanation for why interlocutors align. According to this theory, alignment is the basis for successful dialogue. Alignment occurs when interlocutors' representations become similar at different levels of language production and comprehension. Alignment is achieved in most circumstances without the interlocutors resorting to overt negotiation (Garrod & Anderson, 1987), the reason for this being that the mechanism of alignment is low-level and cognitively inexpensive. This allows alignment to be achieved quickly and efficiently, without reliance upon time-consuming strategies of open negotiation. Indeed, such strategies are only employed when the primitive mechanisms fail.

Dialogue research has shown alignment of linguistic representations, but alignment is hypothesized also to occur for conceptual representations, such as those associated with object location. A speaker's conceptual representation of where objects are located is reliant upon an overall spatial representation, which underpins the use of spatial language. In order to describe object locations effectively, it is important that both interlocutors take the same perspective (Levelt, 1989) concerning the objects they are locating. For example, an addressee must understand whose left a speaker is talking about. In the same way that interlocutors align on which lexical terms should be used to describe a scene, it would be advantageous for interlocutors to align on which perspective a scene should be described from. When describing the location of an object a speaker has a choice of reference frame to use to produce the utterance, and the addressee has a choice of reference frame to use to comprehend the speaker's description. As mentioned in Chapter 2, there are several different ways of defining reference frames, but for the sake of this chapter I shall adopt Levinson's taxonomy (Levinson, 1996, 2003) (the distinction between taxonomies is irrelevant for the experiments in this chapter). Levinson's Taxonomy divides reference frames into three types: *absolute*, *intrinsic* and *relative*. The experiments in this chapter investigate interlocutors' alignment of the intrinsic and relative reference frames.

The intrinsic reference frame locates an object based upon the directional axes of the reference object. The dot in Figure 14 can be described as *above the chair* because it is in a region of space that extends from the top of the chair.

The relative reference frame locates an object in relation to the viewpoint of an observer. The axes of the reference frame are labelled based upon the features of the person upon whose viewpoint the location is based. In Figure 14 the dot would be described as *left of the chair* using a relative reference frame. (In many cases the relative reference frame is used from the viewpoint of a speaker or an addressee, but it can also be from a third person perspective.)

When describing an object's location, an individual can select one of these reference frames to use in preference to either of the other two. Carlson-Radvansky and Jiang (1998) showed that reference-frame selection is achieved via inhibition of non-selected reference frames. In their experiment, when participants used a relative reference frame to identify an object's location, they were slower to describe an object's location using an intrinsic reference frame immediately afterwards. Reference frame inhibition operates not only on the endpoint of an axis, but on at least the entire axis. For example if *left* (intrinsic) is inhibited, then using *right* or *left* (intrinsic) in the subsequent description will take longer than using a relative reference frame.

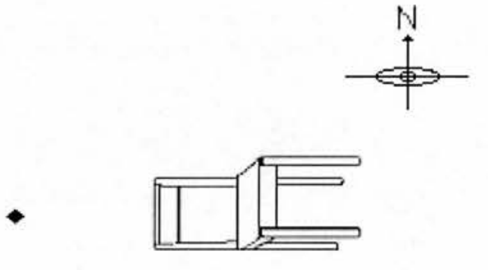


Figure 14. The dot can be described as *west* using an absolute reference frame, *above* using an intrinsic reference frame or *left* using a relative reference frame.

The findings of Carlson-Radvansky and Jiang (1998) suggest that reference frames are influenced by low-level priming. This is the type of low-level mechanism that is a candidate for alignment of reference frames in dialogue. Carlson-Radvansky and Jiang's (1998) experiment only investigated inhibition of the endpoint of an axis and the inhibition of the axis itself. Carlson and van Deman (in press) extended the investigation to the entire reference frame and showed that negative inhibition does operate between axes of the relative reference frame. That is, using intrinsic *above* on trial n led to inhibition of the relative reference frame, and therefore the participants were slower to use relative *left* or *right* on trial n+1. However, the results did not show inhibition from using the intrinsic left and right axis. Using intrinsic *left* on trial n did not lead to an increase in reaction time to use the relative reference frame on trial n+1. These findings suggest that reference frames are amenable to low-level processes which could lead to alignment during dialogue. However, the results do not establish whether this occurs. Carlson-Radvansky and Jiang (1998) (also Carlson & van Deman, in press) used reaction time as a measure of cognitive effort, whereas in dialogue any effect of priming must manifest itself by a change in the person's linguistic behaviour. Therefore it is unclear whether this kind of priming is enough to cause the alignment of reference frames between interlocutors in the manner described by Pickering and Garrod (2004).

With regards to the use of reference frames in dialogue, Schober (1993, 1995) has shown that the reference frame that a speaker selects is affected by their partner in a conversation. In Schober (1993), directors described the location of an object to an addressee (matcher) who viewed the scene from a different perspective and directors were more likely to use the addressee's (matcher's) perspective when describing the location. When the addressee queried the director's descriptions, they used their own perspective to describe the object's location. In this case interlocutors were both using the matcher's perspective to describe the location of the object. From this Schober concluded that interlocutors use conscious strategies to collaborate in ways of describing object location. However, in Schober (1995) the results were slightly different. In Schober (1995) once again there were directors who described the location of objects to matchers, but this time both the describer and the matcher used a matcher-centred reference frame. This time the director was using an addressee-centred perspective, but the matcher was using a speaker-centred perspective. This is different to Schober (1993) where both the director and the matcher used an addressee-centred perspective, making the task of comprehension easier for the person who at the time was the addressee, regardless of their role as matcher or director in the task. In contrast, in Schober (1995) both the director and the matcher used a matcher-centred perspective, making the task of comprehension easier for the person whose role was the matcher, regardless of whether at a given moment they were the speaker or the addressee.

Schober's (1993, 1995) results suggest that interlocutors affect each other's choices of ways to describe objects' locations. However, it is not clear that this is necessarily alignment of reference frames. In his experiments, two participants interacted freely, allowing little control over what was said by each pair. This means that pairs of participants may be reverting to a default reference frame (e.g. there may be a preference for interlocutors to use the perspective of the person who is receiving the information because the information receiver may be perceived to have the more difficult task). If this is the case the results do not tell us what happens when a speaker uses a dis-preferred reference frame.

This chapter presents three experiments that use a confederate priming paradigm to investigate whether or not interlocutors align reference frames during dialogue. Experiment 1 investigated *between* axes alignment of reference frames, using the intrinsic reference frame as defined by theory. Experiment 2 investigated *within* axis alignment of reference frames, using the intrinsic reference frame as defined by theory. However, it was noticed that participants used a different interpretation of the intrinsic reference frame to that used by the confederate in Experiments 1 and 2. Therefore Experiment 3 investigated *between* axes alignment of reference frames, using the intrinsic reference frame in the same way as participants had in Experiments 1 and 2.

3.2. The Confederate Priming Paradigm

The Experiments in this chapter all use the confederate priming method, set up in the following manner (e.g. Branigan et al, 2000). In this paradigm naïve participants were introduced to the confederate as if she was a second participant. The confederate and participant sat facing computers that were positioned back to back as in Figure 15. They were unable to see each other or each other's screen.

The confederate had a script, which appeared as a sentence on their computer screen. The confederate read this sentence, which took the form of a description of the location of a dot in relation to a reference object (*Prime Sentence*). At this point the participant had to decide which of two scenes matched the description given by the confederate (*Match task*). The two scenes in the match task differed only in the location of the dot in relation to the reference object. The participant was then instructed to describe the location of a dot in relation to a reference object on their screen (*Target*). The participants believed that on the confederate's screen there was now a match scene, containing two reference objects with dots, and that the confederate had to choose which of these matched the description. For experiments 1-3 the confederate's description used either an intrinsic reference frame or a relative reference frame. The confederate was unable to tell which condition each trial was in as all she saw was a sentence on a screen. The dependent variable was the reference frame used by the participant to describe the location of the dot on the Target scene.

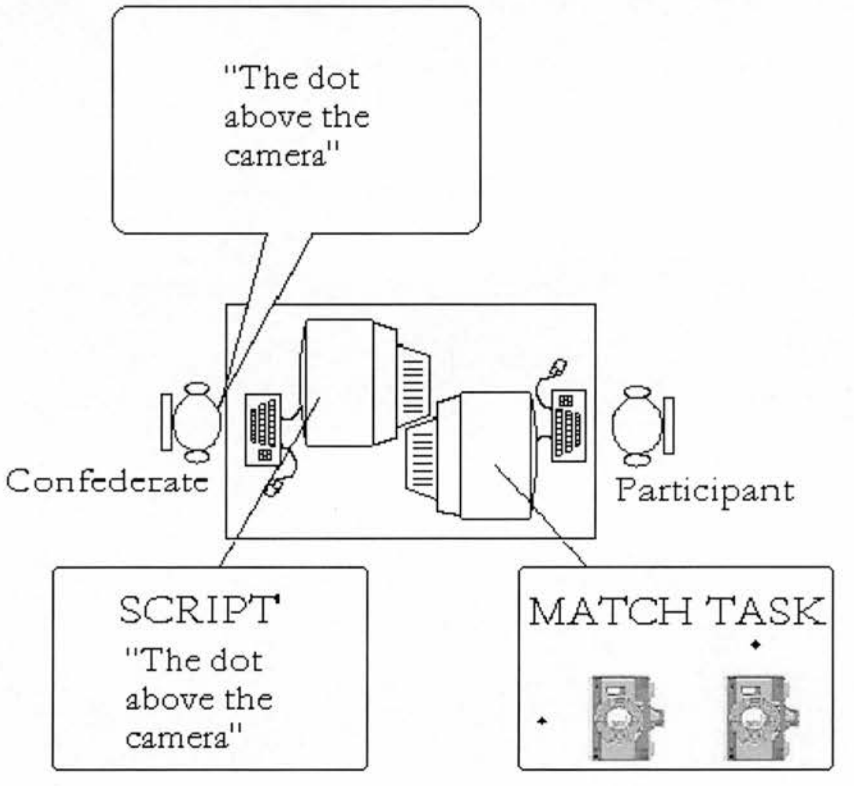


Figure 15: The experimental set up for Experiments 1-3.

3.3. Experiment 1

Experiment 1 was designed to investigate whether interlocutors aligned *between* axes of a reference frame. Carlson-Radvansky and Jiang (1998) have shown that negative priming operates along the axis of a reference frame. It has also been shown in monologue that negative priming operates at the level of the entire reference frame for the relative reference frame, but not for the intrinsic reference frame (Carlson & Van Deman, in press). If interlocutors do align reference frames then they will use a reference frame significantly more when they have just heard an utterance using that reference frame than when they have just heard an utterance using an alternative reference frame. Alternatively interlocutors may select a reference frame based solely upon the perceptual properties of the spatial array, in which case they should be unaffected by the reference frame just used by their partner.

The experiments also set out to separate reference frame alignment from lexical alignment. If alignment of reference frames occurs independently of lexical alignment, we can expect participants to use a reference frame significantly more if they have just heard an utterance using that reference frame, regardless of whether or not using the same reference frame requires using the same spatial term. If reference frame alignment does not occur independently of lexical alignment, then participants will only use the same reference frame as the confederate, when doing so also uses the same spatial term.

Finally, it is possible that there will be a greater amount of reference frame alignment when using the same reference frame as the confederate requires using the same lexical term. This would reflect an increase in alignment due to a 'lexical boost'. Previous work has shown that the more factors in common between the prime and target the greater the amount of alignment (Branigan et al, 2000, Cleland & Pickering, 2003). However a lexical boost has been shown not to occur when the lexical items that are repeated are closed class as is the case with prepositions (Bock, 1989). If lexical repetition increases the amount of reference frame alignment then participants will be significantly more likely to use the same reference frame as the confederate, when doing so requires using the same spatial term than when doing so requires using a different spatial term. Participants will still be more likely to use a reference frame if they have just heard the confederate use that reference frame compared to an alternative reference frame, independent of the spatial term required to use the same reference frame.

3.3.1. Participants

12 students of the University of Edinburgh were paid volunteers in the experiment, which lasted 20 minutes. All were native English speakers. The confederate was a female postgraduate at the University of Edinburgh.

3.3.2 Materials and Design

The experiment was run on two computers positioned back to back, using E-prime software. One program was created for the confederate and consisted of prime sentences positioned in the centre of the screen of the form *The dot above the chair.* This formed the script for the experiment. The second program was for the participant and displayed pictures for the match and target phases of the experiment.

12 monochrome objects, positioned in the centre of the screen, were used as reference objects. Two versions of each object were used, one rotated 90° clockwise and one rotated 90° anti-clockwise from a canonical orientation. The figure object was a square rotated so that its vertices were the top, bottom, left, and right most points. The figure object was located either above, below, left, or right (in a relative reference frame) of the reference object.

There were three within-participants and within-items factors: Prime reference frame (Relative vs. Intrinsic); Preposition (Same vs. Different); and Target orientation (Vertical vs. Horizontal), yielding a total of 8 conditions. The Prime reference frame factor refers to which reference frame the confederate used on the prime phase. Note that the confederate did not actually see the scene, but read a sentence on her screen, which accurately described one of the two scenes displayed on the participant's screen (Match phase).

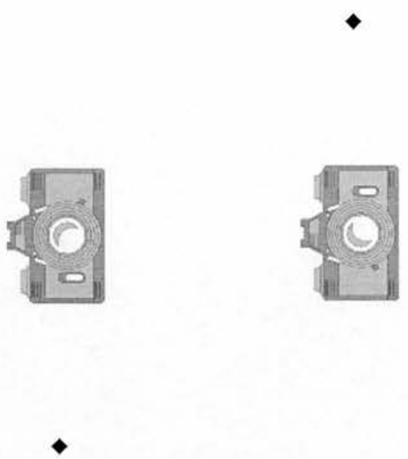


Figure 16: A match scene from Experiment 1-3.

Figure 16 shows a match scene from the experiments. The confederate's description was accurate for only one of these scenes on a given trial. In the following example the description is of the left hand scene. The confederate used either an intrinsic description, in which case the description was *The dot right of the camera,* or a relative description, in which case the description was *The dot below the camera.* The scene described by the confederate is the *Prime scene* (in Figure 16 the left hand scene); the scene that is not described by the confederate is the *Negative scene* (In Figure 16 the right hand scene). Together the prime scene and the negative scene form the *Match scene* (Figure 16).

The Preposition condition refers to whether or not the target scene which the participant describes is the same scene as the prime scene for a given trial. In Figure 16 the left hand picture is the prime scene, therefore in the *Same* condition, Figure 17 would be the target scene the participant had to describe. In the *Different* condition the participant would have to describe the target scene shown in Figure 18

Note that for the *Same* condition if the participant uses the same reference frame as the confederate then they will also have to use the same preposition as the confederate, in this example either *below* (relative reference frame) or *right* (intrinsic reference frame. In the *Different* condition if the participant uses the same reference frame as the confederate then they will be using a different preposition to describe the location of the dot. In this example if the confederate uses an intrinsic reference frame to describe the prime scene in Figure 16 they will describe the dot as *right* of the camera. If the participant then uses an intrinsic reference frame to describe the target scene in Figure 18 they would describe the dot as *above* the camera.

Figure 17: The target scene for the *Same* condition following the match scene in Figure 16.

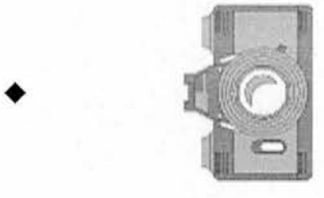


Figure 18: The target scene for the *Different* condition following the match scene in Figure 16.

For each prime scene in the Different condition there was only one possible location for the figure object on the target scene, which was not the same relation according to the alternative reference frame, and where the figure object was not in the opposite position to the prime scene. The dot on the target scene was positioned so that it was not possible for the participant to describe the location of the figure object on the target scene by using the same preposition as the confederate (in the same reference frame), by using the antonym preposition to that used by the confederate, or by using the same preposition as the confederate (in the alternative reference frame).

The Target orientation reflects whether or not the figure and reference object are aligned vertically or horizontally. In Figure 17 the target scene is aligned vertically, in Figure 18 the target scene is aligned horizontally.

The match scenes contained two scenes side by side, the prime scene and the negative scene. The prime and negative scenes differed only in the spatial relation between the figure and reference object; the reference object was the same for both scenes. The description of the confederate was an accurate description of the prime scene only. The negative scenes did not match the description of the confederate according to *any* reference frame. In Figure 16 the prime scene on the left can be described as either *the dot below the camera* (relative) or *the dot right of the camera* (intrinsic). Neither of these descriptions can describe the negative scene shown on the right hand side. The negative scene can only be described as *the dot above the camera* (relative) or the *dot left of the camera* (intrinsic). There were two possible negative scenes for each prime scene. That is, there are two scenes that were not described by the description of the prime scene regardless of which reference frame was used to interpret the scene.

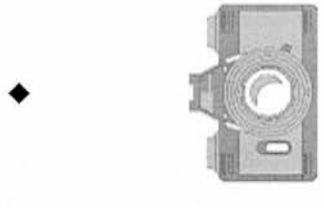


Figure 19: The alternative negative scene for Figure 16 (this would replace the right hand picture).

In Figure 16 the left hand scene shows one appropriate negative scene. Figure 19 shows the alternative negative scene which could be used in Figure 16. Figure 19 can either be described as *the dot left of the camera* or *the dot above the camera*. Once again neither of these is compatible with the description used for the prime scene. As a consequence of there being two different versions of the match scene for each trial two lists were created; one version was used on each of the lists. In addition half of the negative scenes on the match scenes in each list had the dot in the opposite position to the dot on the prime scene; this is the case for the negative scene in Figure 16. The other half had the figure object in the alternative position, this would be the negative scene shown in Figure 19. Finally half the trials had the prime scene on the left and half had the prime scene on the right.

In the experiment all 96 trials, consisting of a prime sentence, match scene and target scene, were experimental trials; there were no fillers. The trials were completely randomised, with a different randomisation for each of the 12 participants. Each of the twelve reference objects (see appendix A) appeared in each of the 8 conditions. The reference object was the same for both the prime scene and the target scene.

3.3.3 Procedure

The two participants were introduced to each other (throughout the experiment, the experimenter treated the confederate as if she was a naïve participant). The participant and confederate were then seated at a computer each. Viewing distance was not controlled for; participants viewed the screen from a distance that was comfortable for them. The computers were situated back to back so that neither could see each other, or the other's screen.

After hearing instructions, participants pressed the space bar to begin a practice session of 8 trials, one trial corresponding to each of the 8 conditions. Instructions then appeared on the screen signalling the end of the practice session and the start of the experiment. Each trial proceeded as follows: After participants pressed <space> to begin, the match screen appeared. The match screen contained two examples of a reference object (both the same, one on the left, one on the right) and a dot located above, below, left or right of each one. The confederate gave a description of the location of the dot in relation to the object. The participant then chose which of the two examples on the screen matched the confederate's description of the dot location accurately, pressing the M key if it was the right-hand example and the Z key if it was the left-hand example. Participants were told that if they were not sure which picture matched their partner's description to pick the one they thought matched most closely.

After selection the match scenes disappeared (no feedback was given) and a fixation cross appeared in the centre of the screen. This remained on screen for 1000ms. The fixation cross was then replaced by a reference object in the centre of the screen with a dot above, below, left or right of it. Participants then described the location of the dot in relation to the object. After describing this they pressed space and the scene disappeared. It was replaced by a fixation cross in the centre of the screen for 500ms. This then disappeared and the next trial began with a match task. This was the procedure for all practice and experimental trials.

3.3.4 Results

In this experiment all participants' responses were coded as using an intrinsic reference frame or a relative reference frame. Descriptions were coded as using an intrinsic reference frame if the prepositions shown in Figure 20 were used by the participant to describe the dot in those locations. Descriptions were coded as using a relative reference frame if the prepositions used in Figure 21 were used by the participants to describe the dot in those locations.

Figure 20: The prepositions used to describe the regions of space surrounding an object when an intrinsic reference frame is used.

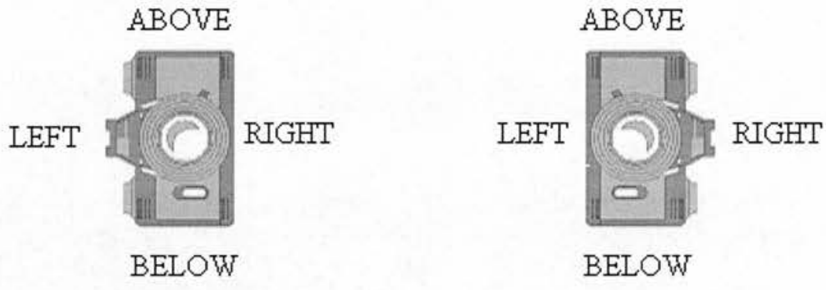


Figure 21: The prepositions used to describe the regions of space surrounding an object when a relative reference frame is used.

When participants made left/right errors these were coded as the appropriate reference frame. This was possible because it was clear that the participant was not intending to use the alternative reference frame as the left-right axis of one reference frame was the top-bottom axis of the non-selected reference frame. Errors of this nature occurred primarily on the left-right axis, but did occur once on the top-bottom axis when the reference object was a medal. This may have been because the medal had a more ambiguous top-bottom axis than the other reference objects. In the analysis participants' first responses were used to code the reference frame they used. If they used one reference frame, but then switched to the alternate reference frame the response was coded as using the first reference frame. This only happened on two trials throughout the twelve participants. Figure 19 can be described as the *dot to the left of the camera* using a relative reference frame or it can be described as the *dot above the camera* using an intrinsic reference frame. If the participant described this scene as *the dot above... sorry I mean to the left of the camera* then this response was coded as intrinsic because this reference frame was used prior to the participant using an intrinsic reference frame. First responses were only coded if the preposition was produced and the reference frame could be determined. If the description of the participant was *the dot a...to the left of the camera* then this was coded as a relative reference frame. Similarly, if the description was *the dot to the...above the camera* then this was coded as intrinsic.

Results are reported in terms of the percentage of intrinsic responses. This is arbitrary as all responses were either intrinsic or relative. The results showed that when the confederate used an intrinsic reference frame to describe the prime scene the participants used an intrinsic reference frame 62.9% of the time to describe the target scene. When the confederate used a relative reference frame to describe the prime scene the participants used an intrinsic reference frame 53.3% of the time. Therefore, participants used an intrinsic reference frame 9.6% more following comprehending an intrinsic reference frame, than following comprehending a relative reference frame.

The percentage intrinsic responses were analysed using two 2x2x2 repeated measures ANOVAs (by participants (F1) and by items (F2)), with Prime reference frame (intrinsic vs. relative), Preposition (Same vs. Different), and Target orientation (Vertical vs. Horizontal). Table 3 shows the mean percentage of intrinsic responses given by participants in each of the 8 conditions. There was a significant main effect of Prime reference frame (62.9% vs. 53.3%; $F1(1,11) = 26.86$, $p < .01$; $F2(1,11) = 9.35$, $p < .05$). Participants were significantly more likely to use an intrinsic reference frame after the confederate had used an intrinsic reference frame, compared to when the confederate had used a relative reference frame.

When the target scene was oriented vertically, participants used an intrinsic reference frame 48% of the time compared to 68% when the orientation was horizontal. This difference was significant ($F1(1,11) = 8.07$; $p < .05$; $F2(1,11) = 101.17$; $p < .01$), showing that participants were significantly more likely to use an intrinsic reference frame when the objects were oriented horizontally than when they were oriented vertically.

As predicted there was no effect of Preposition ($F1(1,11) = 4.34$; $p > .05$; $F2(1,11) = 1.39$; $p > .05$). Participants used an intrinsic reference frame as much when the prime and target scenes were the same as when they were different. This is regardless of which reference frame the confederate used.

	Relative Prime		Intrinsic Prime	
	Same	Diff	Same	Diff
Vertical	32.3	52.8	54.9	52.3
Horizontal	63.6	64.4	73.3	70.9

Table 3: Mean percentage of intrinsic responses in Experiment 1.

There was a significant two-way interaction between Prime Reference Frame and Preposition ($F1(1,11) = 13.07$; $p < .01$; $F2(1,11) = 6.19$; $p < .05$). All other two-way interactions were non-significant (all Fs < 3.1; $p > .05$). There was a significant three-way interaction between reference frame prime, Preposition and Target orientation ($F1(1,11) = 5.6$; $p < .05$; $F2(1,11) = 12.5$; $p < .05$).

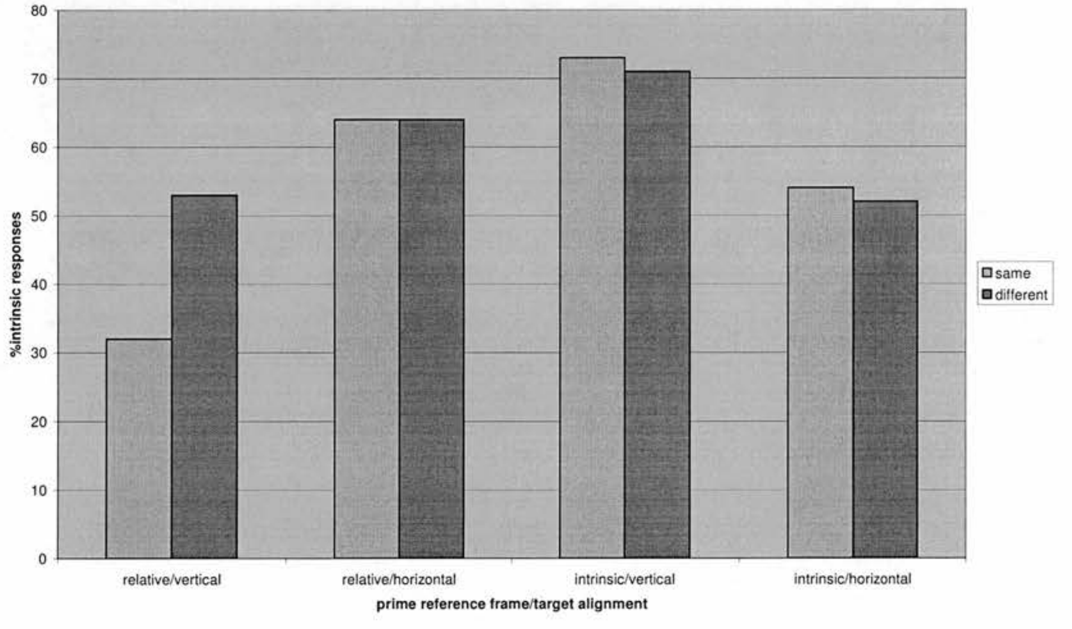


Figure 22: Mean % intrinsic responses for the eight conditions in Experiment 1.

Figure 22 shows that the interaction between Prime Reference Frame and Preposition appears to arise because of a difference between two of the eight conditions; *relative*, *same*, *vertical* and *relative*, *different*, *vertical*. The former yielded 32.3% intrinsic responses, whereas the latter yielded 52.8% intrinsic responses.

Post-hoc analyses showed that this was the only significant difference between same different pairs in the prime overlap condition ($t(23) = -2.91$; $p = .01$).This means that participants were more likely to use a relative reference frame when the reference and figure object were aligned vertically (i.e. they would use *above* and *below* to describe the dot's location) following the confederate using a relative reference frame when there was preposition overlap (i.e. the confederate used *above* or *below*) than when there was no preposition overlap (i.e. the confederate used *left* or *right*).

3.3.5 Discussion

The results of Experiment 1 show an effect of alignment of reference frames. Participants were more likely to use an intrinsic reference frame after the confederate had used an intrinsic reference frame than after the confederate had used a relative reference frame.

There was also a significant main effect of target orientation. Participants were more likely to use an intrinsic reference frame when the figure and reference object were horizontally aligned, compared to when the figure and reference object were vertically aligned. This is because participants preferred to use the lexical terms *above* and *below* as opposed to *left* and *right*. This accords with the well established finding that the top-bottom axis is easier to use than the left-right axis. This has been argued to be as a result of the asymmetry of the top-bottom axis relative to the symmetry of the left-right axis (e.g. Bryant & Wright, 1999; Franklin et al., 1992; Taylor et al., 1990).

It was important to demonstrate that any alignment effect was due to the participants aligning reference frames with the confederate rather than aligning lexically. That is, participants were not just using the same reference frame as the confederate, as a result of using the same spatial term as the confederate. If participants were aligning reference frames with the confederate then the participants should use the same reference frame as the confederate as often when the target scene was different to the prime scene as when the target scene was the same as the prime scene. When the target scene was the same as the prime scene the participant could use the same reference frame as the confederate, just by repeating the preposition used by the confederate to describe the spatial relation. If the participant was more likely to use the same reference frame as the confederate when the target scene was the same as the prime scene, compared to when the two were different, the alignment effect could be attributed to lexical alignment. The significant interaction between prime reference frame and Preposition suggests that the alignment effect is due to lexical alignment. This interaction reflects the difference between the *Same* and *Different* cells for each prime reference frame shown in Table 3. This result indicates that participants are using the same reference frame as the confederate more when doing so involves using the same preposition. However, there is a caveat to the above explanation: The post-hoc analyses indicate that the interaction is due to a difference between one of the four Preposition pairs. For each prime reference frame condition, there are two target orientation conditions (*horizontal* and *vertical*), and then two Prime target overlap pairs (*same* and *different*) for each of these, yielding four pairs. The percentage intrinsic responses for these four pairs are shown in Figure 22. If participants align reference frames the percentage intrinsic responses should be the same within each pair (although different between pairs). This is the case for three of the four pairs (the three rightmost pairs). The percentage intrinsic responses only differs from a *Same* to *Different* condition when the prime reference frame is relative, and the figure and reference object of the target scene are vertically oriented. The percentage intrinsic responses for this pair are represented in the two bars on the leftmost of Figure 22 and the two cells in the upper left of Table 3. This suggests that it is only when this is the case that there is no reference frame alignment and only lexical alignment. It would be expected that the alignment effect would be lexical for all of the conditions, however this is clearly not the case as for the other *Same/Different* pairs the percentage intrinsic responses are statistically the same.

The reason for apparent lexical alignment of only one Same/Different condition pair is not totally clear, but may stem from the observation that participants used intrinsic left-right in the opposite manner to that used by the confederate. Figure 20 shows the areas of space described as intrinsic left and right by the confederate, however the participants consistently described the areas described by the confederate as intrinsic left as right and vice-versa (in the analysis these were coded as intrinsic as they appear to be left-right errors, albeit very consistent ones as they occurred for all trials). This may have caused the participants to interpret some relative prime sentences as intrinsic prime sentences. The reason for this is that when the confederate used relative left and right the description would be accurate for both scenes of the match phase, according to the participants' interpretations of intrinsic left and right (for one scene it would be an accurate relative description, as intended, for the other it would be an accurate intrinsic description, as interpreted by the participant). If the participant chose the intrinsic interpretation then they would be more likely to use an intrinsic reference frame to describe the target scene because of reference frame alignment.

The second caveat for not adopting a lexical alignment explanation is the selection of items used in the experiment. This is related to the first caveat and was possibly a reason why participants were unable to use the intrinsic left/right correctly. Some of the items used as reference objects in Experiment 1 did not have intrinsic fronts and backs (e.g. a bottle). As the intrinsic left and right are assigned by anchoring the direction to the front of the reference object the participants would not have been able to do use the intrinsic reference frame for left and right with some of the reference objects. This may have led to some confusion and could have been the reason they chose to use the intrinsic reference frame in the way they did and consequently why they may have aligned with the unintended reference frame in some of the conditions.

The results of Experiment 1 showed that participants were more likely to use a reference frame if they had just heard an utterance using that reference frame compared to when they have just heard an utterance using an alternative reference frame. The results also suggest that this effect may have been due to lexical alignment rather than reference frame alignment. However, because the lexical effect only occurred when the prime reference frame was relative and the figure and reference object of the target were vertically aligned the lexical explanation cannot be wholly accepted. It is possible that the lexical effect may have arisen because of participants falsely interpreting left and right intrinsic and that this may have been because of an artefact of the materials used as reference objects; specifically that some of the reference objects had no intrinsic fronts. Experiment 2 was carried out to address these issues.

3.4. Experiment 2

In Experiment 2 the objects used as reference objects were changed from those used in Experiment 1, so that all now had intrinsic fronts. This was so participants were able to assign the intrinsic left-right axis in the correct manner. In addition, Experiment 2 was changed so it investigated alignment of reference frames *within* an axis, as opposed to the *between* axis investigation of Experiment 1. This meant that the Preposition Different conditions were replaced with *Antonym* conditions, where the figure object was located in the opposite position on the target scene to that on the prime scene (Experiment 3 returns to a between axes investigation). If interlocutors do align reference frames, participants will be significantly more likely to use a reference frame if they have just heard the confederate use that reference frame, than if they have just heard the confederate use an alternative reference frame. If the alignment is due to reference frame alignment and not lexical alignment, participants will use the same reference frame as the confederate, when doing so requires using a different spatial term, as much as when using the same reference frame requires using the same spatial term.

3.4.1 Participants

12 students of the University of Edinburgh, none of whom had taken part in Experiment 1, were paid volunteers in the experiment, which lasted 20 minutes. All were native English speakers. The confederate was a female postgraduate at the University of Edinburgh.

3.4.2. Materials and Design

In total 12 reference objects were used with the same dimensions as those used in Experiment 1. The reference objects from Experiment 1 that had intrinsic fronts were used. Those that did not have intrinsic fronts were replaced with reference objects that did (see appendix B).

The design was the same as Experiment 1, except that the Preposition *Different* condition was replaced with an *Antonym* condition. Figure 23a-d shows the conditions, used in Experiment 2: Figures 23a and 23b both represent the Preposition *Same* condition. The prime and target scenes are the same, so using the same reference frame as the prime to describe the target will require using the same preposition as on the previous trial. Figure 23a shows the target scene vertically oriented; Figure 23b shows the target scene horizontally oriented. Figures 23c and 23d both represent the Preposition *Antonym* condition. The prime and target scenes are different, so using the same reference frame to describe the target scene as the prime scene will require using the antonymous preposition as on the prime. Figure 23c shows the target scene vertically oriented; Figure 23d shows the target scene horizontally oriented. Apart from replacing the *Different* condition all other aspects were the same as Experiment 1.

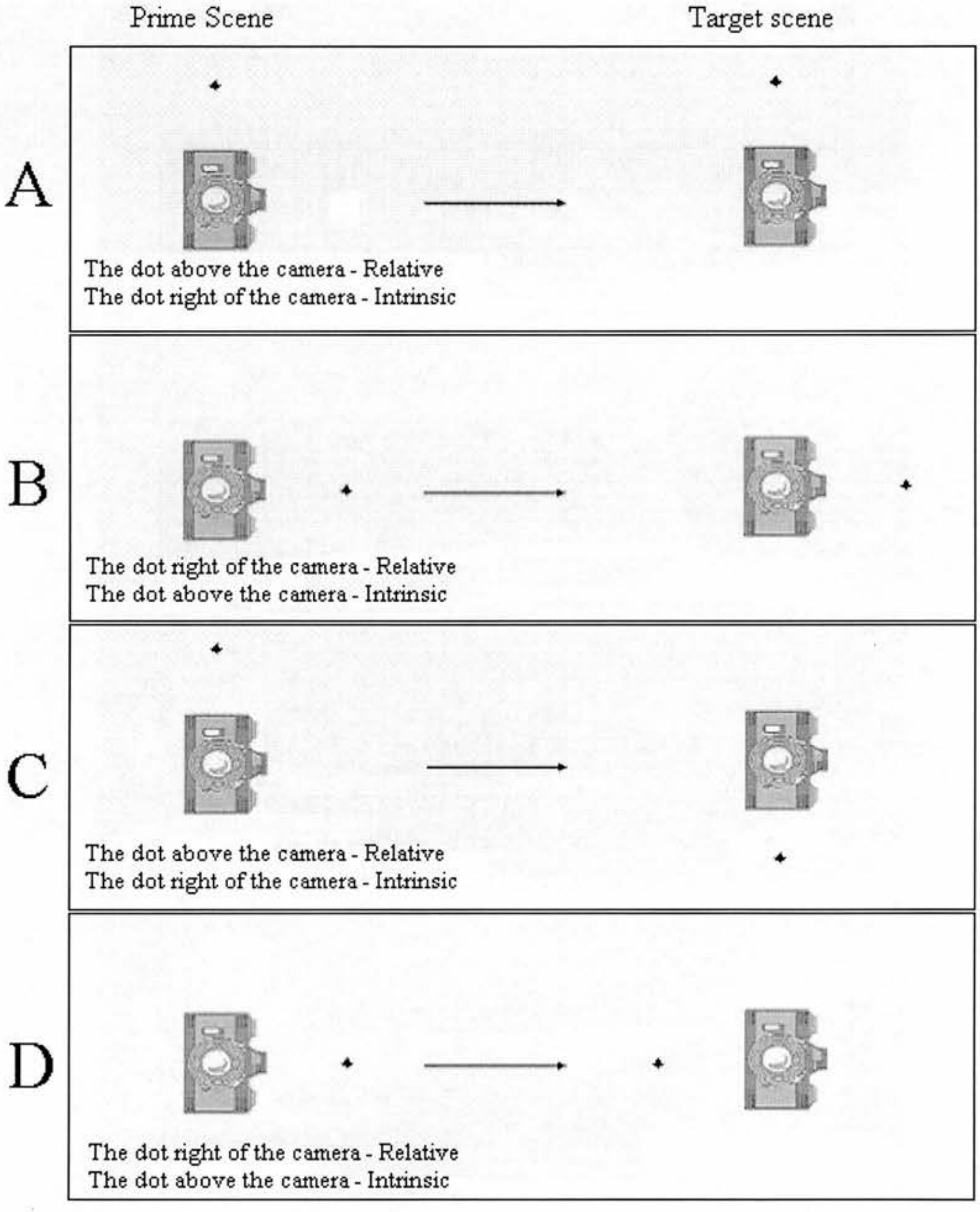


Figure 23a-d: The conditions in Experiment 2.

3.4.3. Procedure

The procedure was the same as Experiment 1.

3.4.4. Results

Analyses were based on participants' initial responses; if they used one reference frame but then changed to the alternative reference frame, their first response was taken. This occurred on 42 (3.6%) trials in the experiment. Utterances were coded in the same way as Experiment 1. The results showed that, as in Experiment 1 participants consistently used intrinsic *left* and *right* in the opposite manner to that used by the confederate. That is, the position which the participants consistently described as *left* using an intrinsic reference frame was described by the confederate as *right* using an intrinsic reference frame. Trials that could be interpreted as *left* and *right* intrinsically (and therefore *above* and *below* relatively) were excluded from the initial analyses. This comprised all the trials in which the figure and reference objects were aligned vertically. Analysis 1 is performed on responses from trials where the target scene was aligned horizontally only. Analysis 2 included all orientations of the target scene.

3.4.5. Experiment 2: Analysis 1

The percentages of intrinsic responses were analyzed using two 2x2 repeated measures ANOVAs, one treating participants (F1) and the other treating items (F2) as the random effect, with Prime reference frame (Intrinsic vs. Relative) and Preposition (Same vs. Antonym), as within-participants and items factors (as all responses were intrinsic or relative, choice of intrinsic responses as the dependent measure is arbitrary).

Table 4 shows the mean percentage of intrinsic responses in each condition (including the vertically aligned conditions excluded from initial analyses). Most important, there was a significant main effect of Prime Reference Frame (35.4% vs. 46.8% ; $F1(1,11) = 8.68$, $p < .05$; $F2(1,11) = 517.6$, $p < .01$), showing that naïve participants were more likely to use an intrinsic reference frame after the confederate had used an intrinsic reference frame than after the confederate had used a relative reference frame.

Additionally, there was no main effect of Preposition ($F1(1,11) = 1.61$, $p > .05$; $F2(1,11) = 1.36$, $p > .05$). Participants were no more likely to use an intrinsic reference frame when the prepositions were the same in the prime and target as when they were antonyms. The interaction between Prime reference frame and Preposition overlap was not significant ($Fs < 1$). Hence participants were just as likely to use the same reference frame as the confederate had just used, when this involved using the antonymous preposition as when it involved using the same preposition.

	Relative Prime		Intrinsic Prime	
	Same	Diff	Same	Diff
Vertical	18.7	17.9	26.3	24.8
Horizontal	34.8	36.1	45.8	48.6

Table 4: Mean percentage of intrinsic responses in Experiment 2.

3.4.6. Experiment 2: Analysis 2

In a second analysis the data from the vertically aligned figure and reference object conditions were included, treating participants' intrinsic left-right orientation as a use of the intrinsic reference frame, despite the fact it was opposite to the way the confederate was using intrinsic left and right. For this analysis the descriptions were coded as intrinsic, if when the dot was in positions 1 and 2 as shown in Figure 24 they were described as right and left of the reference object respectively. Whether or not this is truly an example of the use of an intrinsic reference frame will be discussed further in the General Discussion of this chapter.

The percentages of intrinsic responses were analyzed using two 2x2x2 repeated measures ANOVAs, one treating participants (F1) and the other treating items (F2) as the random effect, with Prime reference frame (Intrinsic vs. Relative), Preposition (Same vs. Antonym), and Target orientation (Vertical vs. Horizontal) as within-participants and items factors.

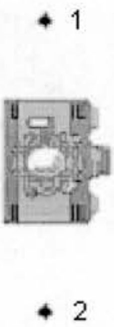


Figure 24: Participants consistently described dot 1 as *left* and dot 2 as *right*, despite this being opposite to the description used by the confederate.

There was a significant main effect of Prime Reference Frame (26.9% vs. 36.2% ; $F1(1,11) = 7.93$, $p < .05$; $F2(1,11) = 167.17$, $p < .01$), showing that naïve participants were more likely to use an intrinsic reference frame after the confederate had used an intrinsic reference frame than after the confederate had used a relative reference frame.

Additionally, there was no main effect of Preposition ($Fs < 1.8$).). Participants were no more likely to use an intrinsic reference frame when the prime and target scene were the same than when they were different. There was a main effect of Target orientation (significant by items and approaching significance by participants) (21.9% vertically aligned vs. 41.3% horizontally aligned; ($F1(1,11) = 4.11$, $p = .067$. $F2(1,11) = .904$, $p<.05$). This suggests that naïve participants were more likely to use the intrinsic reference frame when the target scene was vertically aligned, than when it was horizontally aligned. This effect may reflect the preference of participants to use *above* and *below* rather than left or right.

There was a two-way interaction of Prime Reference Frame and Target orientation by participants alone ($F1\ (1,11) = 6.21$, $p < .05$; $F2\ (1,11) < 1$). If real, this suggests that naïve participants align more with the confederate when the target scene is horizontally aligned (24%) than when the target scene is vertically aligned (7%). All other interactions were non-significant (all Fs < 2).

3.4.7. Discussion

The results of Experiment 2 show that again participants were more likely to use a reference frame after just hearing the confederate use that reference frame than after hearing the confederate use an alternative reference frame. Importantly, in Experiment 2 there was no interaction between Prime Reference Frame and Preposition. This was the case both for the initial analysis, which excluded the vertically aligned target scenes, and for the second analysis, which included them. This indicates that participants used the same reference frame as the confederate regardless of whether the target scene was the same as the prime scene or not. Therefore, participants used the same reference frame as the confederate as much when doing so required using a different preposition to the confederate, as when using the same reference frame required using the same preposition as the confederate. This means the alignment effect cannot be attributed to lexical alignment and is indeed the participant aligning reference frames with the confederate.

As in Experiment 1 there was a significant effect of Figure/reference object target alignment. This indicates that participants were more likely to use the relative reference frame when the target scene had the figure and reference object vertically aligned, and that participants were more likely to use the intrinsic reference frame when the target scene had the figure and reference object horizontally aligned. This reflects the preference for participants using the spatial terms *above* and *below* rather than *left* and *right* (Franklin, Taylor & Tversky, 1992).

Once again, as in Experiment 1, participants consistently used intrinsic left and right in the opposite manner to that used by the confederate. In Experiment 1 this interpretation may have been a consequence of some of the reference objects not having intrinsic fronts, however in Experiment 2 all of the reference objects did have intrinsic fronts. As a consequence, trials in which the target scene could be described using intrinsic left and right were omitted from the initial analysis. In a second analysis these trials were included and coded as intrinsic responses. Both analyses revealed the same pattern of results and both showed a reference frame alignment effect independent of lexical alignment.

Experiment 2 showed that participants did align reference frames with interlocutors and that this was independent of lexical alignment. The reference frame alignment effect was shown to operate within an axis of the reference frame. That is, because the figure object was always located on the same axis of the reference frame for the target scene as on the prime scene. This is consistent with the results of Carlson-Radvansky and Jiang (1998), who showed negative priming within the axis of a reference frame. The lexical alignment result of Experiment 1 is therefore possibly due to an artefact of the confederate's descriptions. As a consequence, Experiment 3 was designed to replicate Experiment 1 in its investigation of between axis reference frame alignment, but the confederate's use of intrinsic left and right was reversed in line with participants' descriptions. If it is the case that the lexical alignment effect evident in Experiment 1 was due to the confederate's use of intrinsic left-right, it should not be present in Experiment 3.

3.5. Experiment 3

Experiment 3 was a replication of Experiment 1 with the exception that the confederate described intrinsic left and right in the same manner as participants had in Experiments 1 and 2. Experiment 3 investigated *between* axis alignment of a reference frame, as the figure object in the target scene was either in the same position as on the prime scene, or was located upon a different axis. It was predicted that reversing the left-right intrinsic axis in line with participants' interpretations would result in participants using a reference frame more following hearing the confederate use that reference frame than following hearing the confederate use the alternative reference frame. If there is reference frame alignment then participants should be just as likely to use the same reference frame as the confederate when doing so requires using a different spatial term as when doing so requires using the same spatial term (i.e., there should be no lexical only alignment effect suggested in the results of Experiment 1).

3.5.1. Participants

16 further students from the University of Edinburgh were paid to take part. All were native English speakers and none had participated in Experiment 1. The confederate was a female postgraduate student.

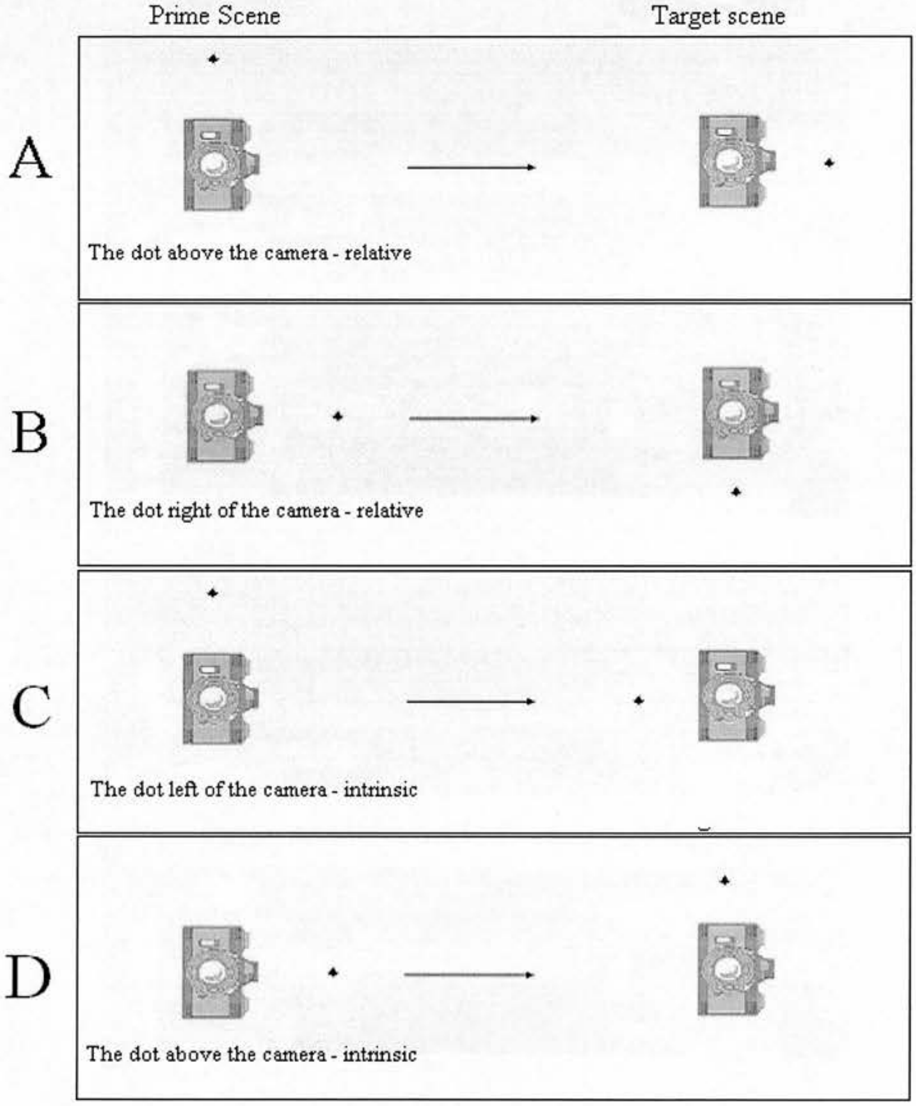


Figure 25a-d: The Relative and Intrinsic 'Different' conditions in Experiment 3.

3.5.2. Materials and Design

This was the same as Experiment 1 except that the confederate used *left* and *right* intrinsic in the opposite manner in line with participants' interpretations from Experiments 1 and 2. Figure 25a-d shows the Preposition *Different* conditions. Figures 25a and b show the relative *Different* conditions. Figures 25c and d show the intrinsic *Different* conditions. These replace the Preposition *Antonym* conditions from Experiment 2 shown in Figure 23c and d.

3.5.3. Procedure

This was the same as Experiments 1 and 2.

3.5.4. Results

The participants' responses were coded in the same manner as that used for coding the results of Experiment 1, with the exception that intrinsic left and right was coded in the opposite manner to that used in Experiment 1. The mean percentages of intrinsic responses for each of the 8 conditions are shown in Table 5.

	Relative Prime		Intrinsic Prime	
	Same	Diff	Same	Diff
Vertical	32.1	26.6	41.2	39.1
Horizontal	39.6	38.1	52.1	48.9

Table 5: Mean percentage of intrinsic responses in Experiment 3.

The percentage of intrinsic responses were analyzed using two 2x2x2 repeated measures ANOVAs, one treating participants (F1) and the other treating items (F2) as the random effect, with Prime reference frame (Intrinsic vs. Relative), Preposition (Same vs. Different), and Target orientation (Vertical vs. Horizontal) as within-participants and -items factors. As in Experiments 1 and 2, there was a significant main effect of Reference Frame ($F1(1,15) = 6.79$; $p < .05$; $F2(1,11) = 24.36$; $p < .01$): Participants used an intrinsic reference frame more often following an intrinsic description (45.3%) by the confederate than following a relative description (34.2%) by the confederate. This is an alignment effect of 11.1%. The main effect of Target orientation was significant by participants only ($F1(1,15) = 6.1$; $p < .05$; $F2 < 1$). If this effect is real, then it shows participants preferred to use a relative reference frame when the target figure and reference objects were vertically aligned, and an intrinsic reference frame when the target figure and reference objects were horizontally aligned. Importantly, the interaction between Reference Frame and Preposition did not reach significance. Indicating that there was no effect of using the same lexical item for the prime and target ($F1(1,15) = 0.18$; $p > .05$; $F2(1,11) = 2.43$; $p > .05$). No other effects approached significance (all $p > .05$).

3.5.6. Discussion

Experiment 3 showed that participants were more likely to use a reference frame if it had just been used by the confederate than if the confederate had just used the alternative reference frame. Unlike the results of Experiment 1 the results of Experiment 3 cannot be attributed to lexical alignment. Participants were just as likely to use the same reference frame as the confederate when doing so required using a different lexical term, as when using the same reference frame required using the same lexical term. The results also extend those of Experiment 2 to show that participants align between axes of a reference frame.

3.6. General Discussion

The results of this chapter show that interlocutors align reference frames when describing objects' locations. Importantly, the results indicate that alignment is not due to lexical priming caused by the experimental participant repeating the preposition just used by the confederate. Repetition of a lexical item by the participant did not increase the amount of alignment as has been shown in previous confederate priming experiments (e.g. Branigan et al, 2000).

The apparent lexical priming effect shown in Experiment 1 was due to the participants interpreting left and right intrinsic differently to what was intended by the confederate. In Experiment 2 nine of the reference objects used were changed from those used in Experiment 1, so that all had an intrinsic front as this was hypothesised to be the cause of the misinterpretation of intrinsic left-right in Experiment 1. The misinterpretation, however, persisted in Experiment 2. Despite this in initial analyses there was no evidence of lexical alignment. In Experiment 3, the confederate's use of intrinsic left and right was reversed in line with participants' interpretations from Experiments 1 and 2. The confederate and participants were now using the intrinsic reference frame in the same way, and the lexical alignment effect from Experiment 1 was no longer evident.

The results of the three experiments support the interactive alignment model of dialogue (Pickering & Garrod 2004). Interlocutors align at multiple levels of linguistic representation: These results extend this to show that interlocutors also align aspects of a conceptual representation (i.e. reference frames). This helps an addressee choose the correct reference frame for interpreting a partner's spatial description and helps a speaker select a reference frame which their addressee is most likely to comprehend. In short, the alignment of reference frames allows interlocutors to adopt the same perspective on a scene (conceptually) and therefore ease dialogue. The act of comprehending a partner's locative description involves the interpretation of their utterance using a reference frame. In using a reference frame to comprehend an utterance it increases the likelihood that that reference frame will be used on a subsequent occasion. The results of this experiment show that this happens fast and in a transient manner. Schober (1993, 1995) argued that interlocutors converge on the perspective which they choose to describe a scene from. However, in those experiments interlocutors did not take turns as the information giver and receiver. The present experiments allow an examination of how understanding a locative utterance affects subsequent reference frame selection on a trial by trial basis. The results suggest that interlocutors are very flexible in their choice of reference frames and are able to switch between one reference frame and another if that is what their partner used immediately previously.

What is notable about these results is that there was no cumulative effect of lexical alignment and reference frame alignment: There was no increase in the amount of reference frame alignment when using the same reference frame as the confederate required the participant to use the same spatial term as the confederate, compared to when aligning reference frames required the participant to use a different spatial term to the confederate. Other studies have shown larger alignment effects when more factors are common between the prime and the target (e.g. Branigan et al., 2000;

Cleland & Pickering, 2003), and Pickering and Garrod (2004) have argued that alignment at one level of representation percolates through the language system and increases the likelihood of alignment at other levels of representation. However, in these cases the repetition was of open class lexical items (nouns, verbs etc.), and other studies (e.g. Bock, 1989) have found no increase in the amount of syntactic priming when there was lexical repetition of closed class items. As prepositions form a closed linguistic class this may be the reason why there was no evidence of lexical alignment. Alternatively, it may be because the prepositions were used to refer to both their intrinsic relation and relative relation, and so held little meaning independent of the reference frame selected regarding the location of the figure object. As such there would be little gained by aligning lexically. A prediction of this latter hypothesis is that if different lexical terms were used by the confederate for each reference frame (e.g. *over* for relative and *above* for intrinsic) then there should be an effect of lexical alignment, as in this scenario the preposition used would provide information, independent of the reference frame, regarding the location of the figure object.

The results support the work showing negative priming of reference frames (Carlson-Radvansky & Jiang, 1998, Carlson & Van Deman, in press). Carlson-Radvansky and Jiang (1998) showed negative priming of a reference frame along a single axis of a reference frame, but they did not investigate negative priming between axes of a reference frame. The results presented here show that alignment occurs between axes of a reference frame. In line with Pickering and Garrod (2004) the mechanism of alignment is argued to be, like negative priming, low-level (negative priming may be a mechanism for alignment of reference frames, but this cannot be established from the present data). Carlson and Van Deman (in press) showed that negative priming also operates between axes of a reference frame. However, the results presented here were not entirely the same as those of Carlson and Van Deman (in press). In their study they showed that there was no between axes effect for intrinsic left and right. What this means is that when participants had to use relative left or right they did not have to inhibit intrinsic left and right and so the intrinsic reference frame as a whole was not inhibited. They interpret this as showing that the intrinsic left-right axis cannot be used easily and so is not represented on the axes of the reference frame. This is not the case in the experiments presented here: Participants were able to use intrinsic left and right, and the use of intrinsic left and right by the confederate led to the use of intrinsic above and below by the participant and vice versa.

There are two possible reasons for the discrepancy between the present results and the results of Carlson and Van Deman (in press): The first is that in their study they explicitly instructed participants to use the intrinsic reference frame in the theoretically correct way. Participants in Carlson and Van Deman (in press) had to understand that this was the correct way to use the intrinsic reference frame before the experiment could begin. It is evident from the present studies that participants preferred a different orientation of the intrinsic left-right axis (i.e. the one in Experiment 3) and as such the way it was implemented in Carlson and Van Deman (in press) made it difficult for it to be inhibited. The second possibility rests on the issue of automaticity and conscious strategies. Carlson and Van Deman (in press) used reaction time to measures participants' apprehension of a spatial relation with a given reference frame. This is something which is less likely to be affected by a conscious strategy. Participants are presumably trying to be as fast and as accurate as possible and therefore any reaction time differences are caused by unconscious mental processes. It is possible that alignment of reference frames is due to a conscious strategy (cf. Schober, 1995), as opposed to a pre-attentive low-level mechanism (cf. Pickering & Garrod, 2004). It is also possible that there is a division between these two mechanisms whereby the intrinsic left and right is aligned using conscious strategies, whereas the other axes of the intrinsic reference frame and the relative reference frame are aligned using low level mechanisms.

3.6.1. Participants use of the Intrinsic left-right axis

One of the issues that arose from these experiments is the orientation of the intrinsic left-right axis that has been used by both participants and the confederate. In Experiments 1 and 2 the confederate used intrinsic *left* and *right* as indicated in Figure 24. In Experiment 3 the orientation of the intrinsic left-right axis was reversed in line with the interpretations of the participants. Up until now I have referred to these both as examples of the intrinsic reference frame. An important question is, which reference frame are these two uses actually employing?

It seems clear that the confederate in Experiment 1 and 2 is using *left* and *right* intrinsic in an unambiguous fashion. If a person were to put themselves in the place of the reference object then their left and right sides would be in the same position as left and right of the reference object as described by the confederate. However, participants interpreted *left* and *right* intrinsic in the opposite manner.

There are two possible explanations here: one is that it participants are employing a secondary relative reference frame, the other, that this is an utterance using an intrinsic reference frame.

According to Levinson's (1996, 2003) theory of reference frames it is possible to create a secondary co-ordinate system (C^2) by performing different processes on a primary co-ordinate system (C^1). C^2 can be derived from C^1 by either a process of *rotation, reflection*, or *translation*. For example, in Figure 26 the ball is between the observer and the tree. In English the observer would describe the ball as *in front of the tree.* In this situation Levinson (2003) argues that the observer's relative C^1 is being *reflected* onto the tree (which has no intrinsic axes in the horizontal plane) to form C^2. The tree then inherits the directional axes of C^2.

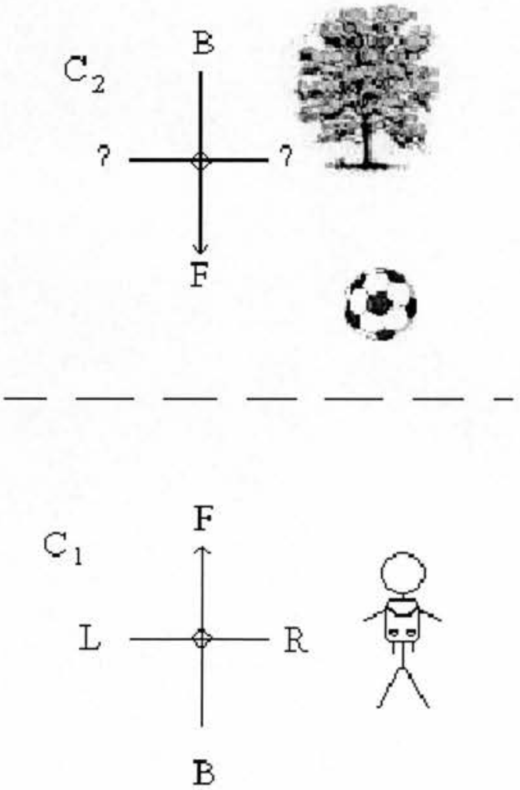


Figure 26: In English the C^1 is reflected in the dashed line to form C^2 and the tree inherits these directional axes.

In other languages (e.g. Hausa), the ball is described as in front of the tree, when the situation in Figure 27 is the case (i.e. the ball is on the far side of the tree from the observer). In this case, C^1 is being *translated* to form C^2. The tree still inherits the axes, but the different process leads to a different orientation of C^2.

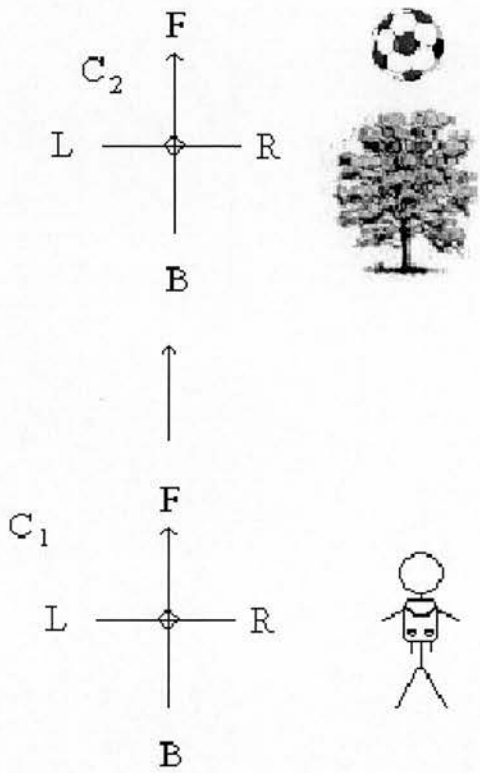


Figure 27: In Hausa C^1 is translated to form C^2 and the tree inherits different directional axes to those in English

The participants' use of left-right intrinsic in these experiments may have actually been the use of a secondary relative reference frame. What participants may have been doing was rotating their relative reference frame, this being the primary co-ordinate system, so that their top-bottom axis was aligned with the reference object's top-bottom axis (see Figure 28). If the left-right axis followed suit then it would result in the pattern of utterances displayed in the experiments. According to Levinson (2003) the participants are in fact using a secondary relative reference frame, which can be termed the *rotated relative*.

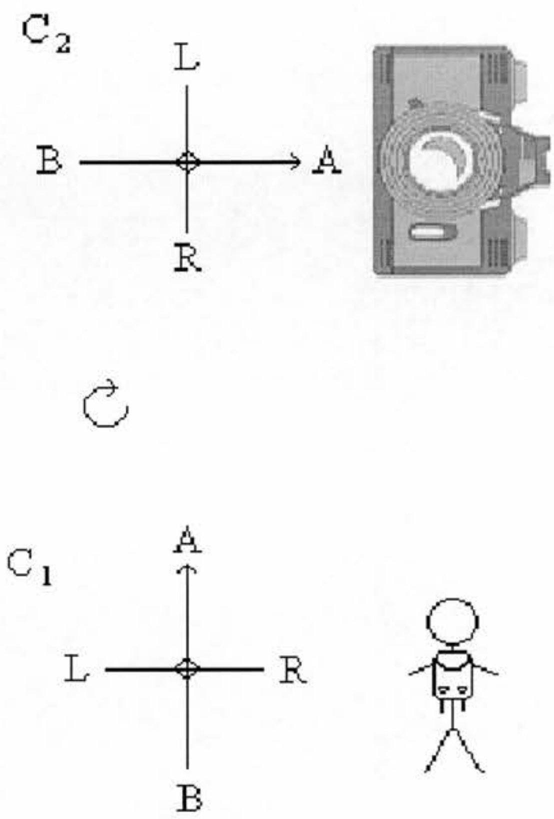


Figure 28: The 'intrinsic' reference frame in Experiments 1-3 may have been a rotated relative.

Despite the argument that participants are using a rotated relative reference frame there is an equally strong argument that they are in fact using an intrinsic reference frame, although not an intrinsic reference frame that is defined in the literature. The argument that participants are using is an intrinsic reference frame rests upon the patterns of alignment shown in the experiments. The conclusion of these experiments is that participants are more likely to use an utterance that is based upon the same underlying reference frame as they have just heard used by the confederate. Following from this, if an utterance is more likely to be used by the participants following a given utterance spoken by the confederate then the assumption is that both utterances share a common underlying representation.

In the present experiments when the confederate described the scene in Figure 29a as *the dot above the camera* the participants were more likely to describe the scene in Figure 29b as *the dot left of the camera*, compared to when the confederate had described Figure 29a as *the dot right of the camera* (the opposite was also the case when Figure 29b was the prime and 29a the target).

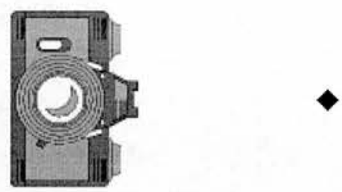


Figure 29a: The dot is above the camera (intrinsic) and right of the camera (relative).

Figure 29b: The dot is left of the camera (intrinsic) and below the camera (relative).

According to the argument laid out above this indicates that both of these utterances (*the dot above the camera* and *the dot left of the camera*) share a common underlying representation (reference frame). Previous literature has almost always argued that the description of Figure 29a as *the dot above the camera* is a use of the intrinsic reference frame (e.g. Carlson-Radvansky & Irwin, 1993, Carlson-Radvansky & Logan, 1997, Levelt 1989), therefore as participants are more likely to describe Figure 29b as *the dot left of the camera*, it follows that this utterance too employs the intrinsic reference frame. Alternatively, both descriptions could be an example of the rotated relative reference frame. If, however, the latter scenario was the case then it would not be clear that the intrinsic reference frame was anything other than a transformed relative reference frame.

It is not possible to tell from these data alone which of these accounts is correct and therefore which reference frame it is that participants are using when they are describing intrinsic (rotated relative) *left* and *right*. However, it is clear that participants are aligning at some level: If we accept that the reference frame used is the intrinsic reference frame then participants are aligning reference frames with the confederate. If on the other hand we accept that the participants are using the rotated relative reference frame, then participants are aligning the process of rotating C^1 relative to form C^2 relative.

What is important is that it is clear that interlocutors are aligning reference frames, or some aspects of a reference frame during conversation. The experiments presented here also show that spatial language is amenable to investigation using dialogue paradigms. This is important as most research on spatial language has used monologue contexts, whereas investigating reference frames using dialogue paradigms provides a more naturalistic context. This is of particular interest as in other areas of psycholinguistics the use of language in dialogue has been shown to involve more than is evident from investigations in monologue (e.g. Brennan & Clark, 1996, Schober, 1993) and it is likely that this is also the case for spatial language. Further, it is possible that spatial language is used in different ways during conversation than in monologue laboratory experiments.

Because these experiments show that spatial language can be investigated using dialogue paradigms it allows the investigation of reference frame structure (as well as other aspects of spatial language) in this manner. Chapter 4 presents a series of experiments investigating the typology of reference frames using the confederate priming paradigm, based upon the premise that interlocutors align reference frames in conversation.

3.7. Conclusion

This chapter raises many issues about the representation of reference frames and the processes that operate on them. Whilst these are undoubtedly interesting in their own right they are not the main focus of the chapter, nor was the initial goal of the experiments presented here to investigate them. The finding of primary importance in this chapter is that interlocutors align reference frames during dialogue and that this is a transient phenomenon allowing speakers to switch from one reference frame to another on an utterance by utterance basis. This is in line with the interactive alignment model of Pickering and Garrod (2004) and suggests that in natural language use one of the factors that play a role in a speaker's reference frame selection is the reference frame that their addressee has used on previous occasions.

Chapter 4: Reference frame taxonomy

4.0. Chapter overview

This chapter builds upon the findings presented in Chapter 3, that interlocutors align reference frames during dialogue. Four experiments are presented, which investigated the representation of different reference frames with the goal of determining the appropriate way to categorise reference frames. The alignment literature (e.g. Branigan et al, 2000; Haywood, Pickering & Branigan, 2005; Pickering & Garrod, 2004; Watson, Pickering & Branigan, 2004) suggests that an utterance is more likely to be produced by an interlocutor, if it is based upon the same underlying representation as was just used by their partner to produce an utterance. In terms of spatial language, a speaker is more likely to use a locative utterance that uses the same reference frame as was just used by their partner.

The confederate priming paradigm was used in this series of experiments to investigate which locative utterances use which reference frames. This is an important question to investigate as the spatial language literature has yet to establish an accepted taxonomy of reference frames (cf. Levinson, 1996, 2003; Jackendoff, 1996, Miller & Johnson-Laird, 1976; Herskovitz, 1986). This Chapter seeks to establish empirical evidence to differentiate between two different reference frame taxonomies, that of Levinson's (1996, 2003) linguistic reference frames, which divides reference frames into *absolute, intrinsic* and *relative*, and the Traditional taxonomy, which divides reference frames into *absolute, intrinsic* and *deictic* (e.g. Carlson-Radvansky & Logan, 1994; Miller & Johnson-Laird, 1976).

Experiments 4-7 investigated the different predictions made by Levinson's taxonomy and by the Traditional taxonomy. The results lend support to the Traditional taxonomy over Levinson's taxonomy, but this interpretation is made cautiously because the results also show that reference frame alignment is influenced by the position of the reference object and the results do not rule out alignment of the origin of the reference frame independently of the reference frame itself.

4.1. Introduction

There has been little conclusive agreement regarding how reference frames should be divided into distinct types. Chapter 1 described several different taxonomies of reference frames, (e.g. Levinson, 1996, 2003; Jackendoff, 1996, Miller & Johnson-Laird, 1976) each differing in their characteristics. This Chapter investigates the differences between two of these taxonomies; the Traditional taxonomy and Levinson's taxonomy. In Chapter 3 the differences between the reference frame taxonomies had no bearing upon the design or results, but for the sake of consistency I adopted Levinson's taxonomy of *absolute*, *intrinsic* and *relative*. This reflects a recent move in the literature (e.g. Carlson, 2002, Logan & Sadler, 1996) to a preference for this taxonomy rather than the Traditional taxonomy of *absolute*, *intrinsic* and *deictic*. The shift towards Levinson's taxonomy has largely been for theoretical reasons; however there is little empirical evidence to support either taxonomy over the other. The focus of the experiments in this chapter was to test the different predictions of Levinson's taxonomy and the Traditional taxonomy.

4.1.1. Traditional Taxonomy (Miller & Johnson-Laird, 1976)

The Traditional taxonomy divides reference frames into three distinct categories: *absolute*, *intrinsic*, and *deictic*. The absolute reference frame is in all respects the same as the absolute reference frame defined by Levinson (1996) and as it is not investigated in this chapter it shall not be discussed further.

The *intrinsic* reference frame has its directional axes set in line with the intrinsic axes of the reference object. For example, the front of the object can be defined in terms of the direction of motion. The assignment of the directional axes of an object can be done in several different ways (described in the introduction, but see Jackendoff, 1996), what is important is that once the directional axes are defined these are then used to label the regions of space around the reference object. For example, a ball is in front of a car, using the intrinsic reference frame, if it is in the region closest to the car's headlights.

The *deictic* reference frame has its directional axes set egocentrically (it is an issue as to whether or not a reference frame which has its directional axes set in accordance with an addressee is also a deictic reference frame. This issue is dealt with in

Experiment 6). A figure object is to the left of a reference object if it is to the ego's left of the object. The defining characteristic of the deictic reference frame is that it is a *speaker centred*, *egocentric*, reference frame.

According to the Traditional taxonomy a reference frame type is defined by the object that defines the *orientation* parameter of the reference frame and the *origin* parameter. If this is the speaker then it is a deictic reference frame. If it is another object then it is the intrinsic reference frame.

4.1.2 Levinson's taxonomy (Levinson, 1996, 2003)

Levinson's taxonomy also divides reference frames into three distinct types: *absolute*, *intrinsic* and *relative*. As mentioned above, the absolute is the same for both the Traditional and Levinson's accounts.

Levinson's intrinsic reference frame has a *binary* argument structure, where the two arguments of the relation are the figure object and the reference object. The reference object is the object that determines the directional axes of the reference frame. This means that if a spatial description uses the speaker as the reference object the utterance uses an intrinsic reference frame. According to Levinson's taxonomy the following sentences all employ an intrinsic reference frame:

1). *The ball is in front of the car* (where the ball is in the region projected from the car's headlights)

2). *The ball in front of you* (where the ball is in the region projected from your front side)

3). *The ball in front of me* (where the ball is in the region projected from my front side).

The feature that these three sentences have in common, and the feature which, according to Levinson, defines them as using an intrinsic reference frame, is they all have a binary argument structure. That is they have a reference object (car, you, and me respectively) as one argument and a figure object (a ball) as a second argument. This means that if the reference object in any of these scenes is rotated then the description is no longer accurate, but the direction the scene is viewed from is irrelevant when using the intrinsic reference frame.

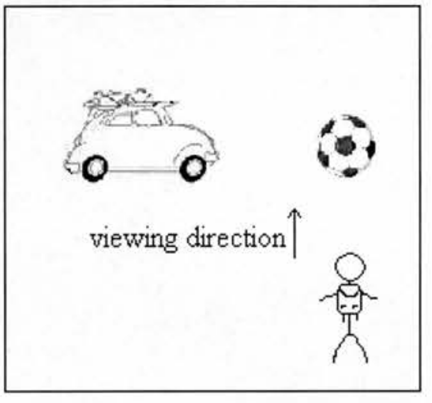
Levinson's relative reference frame is based upon the viewpoint of an observer, who can be the speaker, addressee or a third person. The observer defines the directional axes of the reference frame. This then becomes a *ternary* relationship (as opposed to the intrinsic binary one), where there is a figure object, reference object, and an observer that is *not* the reference object. This reference frame is reliant upon the viewpoint of the observer being maintained; if the viewpoint changes, then the spatial description is no longer accurate.

Levinson's Taxonomy classifies reference frames independently of the location of the origin. In sentence 3, above, the origin is egocentric, but the reference frame used is intrinsic. In contrast, a sentence such as *the ball is to the left of the car* (from my perspective) the origin is also egocentric, but the reference frame is relative. It is possible that interlocutors, rather than aligning the reference frame, align origin of the reference frame. If this is the case then interlocutors would be aligning representations *across* reference frames.

4.1.3. Traditional and Levinson's taxonomic differences

In essence the difference between the two taxonomies is based upon the method of reference frame categorisation. The Traditional taxonomy uses the object that sets the orientation parameter of the reference frame as the defining characteristic. If it is an object other than the ego then it is *intrinsic*, if it is egocentric (or addressee centred) it is *deictic*. The Traditional taxonomy therefore, affords a degree of special status to a reference frame that is egocentric. Levinson's taxonomy uses the argument structure of the reference frame as the defining characteristic: If it is *binary* and dependent upon the orientation of the reference object the reference frame is *intrinsic*. If it is *ternary* and based upon the viewpoint of the observer then it is a *relative* reference frame. Note that according to Levinson's taxonomy, no special status is afforded to the egocentric reference frame and so sentences like *The ball is in front of me* use an *intrinsic* reference frame, whereas according to the Traditional taxonomy it uses a *deictic* reference frame.

In this chapter sentences such as *the ball is in front of me* will be described as *ego-referent.* This is because the sentence has the speaker as the reference object and the different taxonomies argue that this sentence uses different reference frames, *deictic* or *intrinsic* for Traditional and Levinson's respectively. Sentences such as *The ball is in front of the car*, where the ball is at the car's headlights, will be described as using an *intrinsic* reference frame because both taxonomies agree that this sentence uses an intrinsic reference frame. Finally, for sentences such as *the ball is to the left of the car*, where the ball is left of the car from an observer's perspective, will be described as *relative/deictic*_reference frame because both taxonomies can treat this sentence as equivalent, although they label the reference frame differently. Table 6 shows a scene containing a ball, a car, and a person (whose position the reader should assume).



	Traditional Taxonomy	Levinson's Taxonomy	These experiments
The ball is in front of me	Deictic	Intrinsic	Ego-referent
The ball is in front of the car (at the car's headlights)	Intrinsic	Intrinsic	Intrinsic
The ball is to the right of the car (to my right of the car)	Deictic	Relative	Relative/Deictic

Table 6: The reference frame that the Traditional and Levinson's taxonomies attribute to different sentence forms describing a scene.

Table 6 also contains descriptions of the location of the ball and which reference frames the two taxonomies attribute to which descriptions; in addition it also includes how the reference frames will be referred to when discussing them in the context of these experiments. The four experiments presented here use a confederate priming paradigm to determine which of the two taxonomies describes the correct way to categorise reference frames. This operates in much the same way as the confederate priming paradigm used for the experiments in Chapter 3: The confederate was introduced to the participant as if she was another participant and both were led to the experiment room. The only real difference between the confederate priming paradigm used in this chapter and in Chapter 3 is that the participant and confederate were seated side by side in these experiments (see Figure 30), with a divide between them to prevent them seeing each other's screen, as opposed to opposite each other.

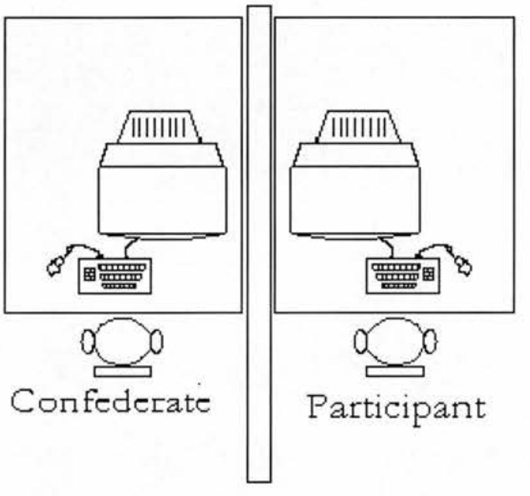


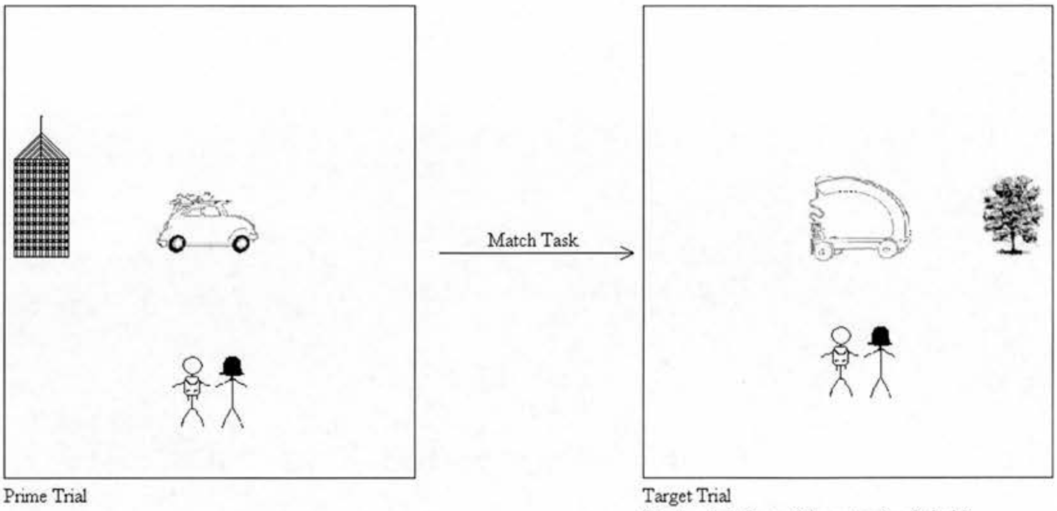
Figure 30: The participant and confederate configuration for experiments 4-7

4.2. Experiment 4

Experiment 4 investigated whether or not the sentence *The ball is in front of me* uses the same reference frame as the sentence *The ball is in front of the car* (at the car's headlights), or uses the same reference frame as the sentence *The ball is to the left of the car* (to the left side from my perspective). According to Levinson's taxonomy, *The ball is in front of me* uses an intrinsic reference frame, and therefore the same reference frame as *The ball is in front of the car*, which uses an intrinsic reference frame according to both Levinson's and the Traditional taxonomies. According to the

Traditional taxonomy, *The ball is in front of me* uses a deictic reference frame because it is egocentric, and therefore uses the same reference frame as the sentence *The ball is to the left of the car* which uses a relative/deictic reference frame (depending upon the taxonomy).

This experiment investigated which reference frame the sentence *The ball is in front of me* uses. As the two different taxonomies (Levinson's and the Traditional) make different predictions about which reference frame is being used, the results may provide empirical support for one taxonomy over the other. In the experiment the confederate described the location of a figure object to the participant using an *ego-referent* reference frame; a sentence which in Levinson's taxonomy is *intrinsic*, but in the Traditional taxonomy is *deictic*, (e.g. *The ball is in front of us* - because there are two people, *us* is used rather than *me*). This is the *Prime sentence*. The participant then decided whether or not the prime sentence was an accurate description of the location of the object displayed on their screen. This is the *match phase* containing the *match scene*. The participant then described the location of a figure object, but was presented with a situation where they could either use an *intrinsic* reference frame (one that is *intrinsic* according to both taxonomies) or a *relative/deictic* reference frame (one that it is *relative* according to Levinson' Taxonomy and *deictic* according to the Traditional taxonomy). This is the *Target trial* containing the *target scene*. Figure 31 shows an example of a trial from Experiment 4. The match scene was described by the confederate as either: *The skyscraper is left of the car* (relative/deictic condition); *The skyscraper is behind the car* (intrinsic condition); or *The car is in front of us* (ego-referent condition) (the two people represented the confederate and the participant).



Prime Trial
The skyscraper is to the left of the car - relative/deictic
The skyscraper is behind the car - intrinsic
The car is in front of us - ego-referent

Target Trial
The tree is in front of the catapult - (intinsic)
The tree is right of the catapult - (relative/deictic)

Figure 31: An example of a trial from Experiment 4.

The participant then had to describe the location of the tree on the Target trial. This could be done using either an intrinsic reference frame (*The tree is in front of the catapult*) or a relative/deictic reference frame (*The tree is right of the catapult*).

If Levinson's taxonomy is correct and *The ball is in front of us* uses an intrinsic reference frame, then participants should be more likely to use an intrinsic reference frame after hearing the confederate use an ego-referent reference frame than after hearing the confederate use a relative/deictic reference frame. There should be no difference in the likelihood of the participant using an intrinsic reference frame after hearing the confederate use an ego-referent reference frame than after hearing the confederate use an intrinsic reference frame.

If the Traditional taxonomy is correct and *The ball is in front of us* uses a deictic reference frame then participants will be more likely to use a relative/deictic reference frame after hearing the confederate use an ego-referent reference frame than after hearing the confederate use an intrinsic reference frame. There should be no difference in the likelihood of the participant using a relative/deictic reference frame after hearing the confederate use an ego-referent reference frame than after hearing the confederate use a relative/deictic reference frame.

4.2.1. Participants

12 students of the University of Edinburgh were paid volunteers in the experiment, which lasted 20 minutes. All were native English speakers who had not participated in Experiments 1-3. The confederate was a female undergraduate at the University of Edinburgh. Six further participants were excluded from analyses because they interpreted the confederate's intrinsic reference frame descriptions using a relative reference frame (i.e. they consistently responded that the match scene did not match the confederate's description for the experimental intrinsic trials).

4.2.2. Materials and design

The Experiment was run using two E-prime computer programs. One computer program presented locative sentences upon a computer screen and was used by the confederate. A second program was created for the participant which presented pictures for the match phase and the describe phase of each trial.

Each of the scenes used in the experiment were created from a picture of two people, one object viewed from the side which had intrinsic sides, termed a *tri-axial* object (e.g. a car) and one object which had no horizontal intrinsic axes, termed a *bi-axial* object (e.g. a tree) (Carlson-Radvansky & Jiang, 1998). Together the two people fitted into a 69x69 pixel square. The tri-axial objects were fitted into a rectangle approximately 100x90 pixels in dimension. The bi-axial objects fitted into a rectangle approximately 90x60 pixels in dimension (see appendix C).

All the *match* and *target* scenes contained the two people at the centre and bottom of the screen, as well as one tri-axial object and one biaxial object. The tri-axial and bi-axial objects were selected from three *land* bi-axial items (mast, tree, and skyscraper) and three *land* tri-axial objects (car, catapult, and train) or three *sea* bi-axial objects (buoy, lighthouse and island) and three *sea* tri-axial objects (ship, windsurfer, and hovercraft) (see appendix C). The sea and land items were never paired together in a picture. One of the bi-axial or tri-axial objects was presented in the centre of the screen and one to the left or the right of the central object. All of the match and target scenes in the experiment therefore appeared in an arrangement as in Figure 32.

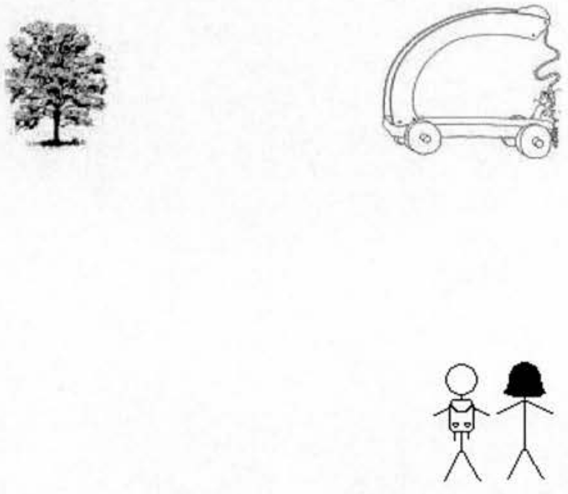


Figure 32: The arrangement of objects in the prime, match and target scenes of Experiment 4.

There was one within-participant, but between-items factor of Prime reference frame (ego-referent, intrinsic, and relative/deictic). There were 108 trials each consisting of a *prime sentence* where the confederate read a sentence which described the location of one of the objects on the match scene, a *match phase* where the participant saw the match scene on their screen and had to decide whether or not the prime sentence was an accurate description of the match scene, and finally a *target phase* where the participant was told, via a word on the screen, to describe the location of one of the objects on the target scene (see figure 31).

54 of the trials were experimental trials, 18 in each of the three conditions. For all experimental trials the prime sentence was an accurate description of the match scene. The match scenes were created by pairing each of the tri-axial objects with each of the bi-axial objects, within the land and sea types (e.g. the car was paired with the mast, the tree, and the skyscraper to form three different scenes and the hovercraft was paired with the buoy, the island, and the lighthouse). This was the case for each of the other tri-axial objects. The spatial relationship between the bi-axial and tri-axial objects was counterbalanced across the three different conditions

so that for each of the match scenes the bi-axial object was to the left of the tri-axial object an equal number of times as to the right of the tri-axial object and also so that the bi-axial object was intrinsically in front of the tri-axial object and intrinsically behind the tri-axial object an equal number of times. Further, the positions of the objects' were counterbalanced so that each tri-axial object appeared as often as the central object as each bi-axial object. However, on experimental target scenes the tri-axial object was always the central object.

The target scenes were created from the same pool of objects as the objects used in the match scenes. Again the 18 target scenes in each condition were formed from combinations of the tri-axial and bi-axial objects. The target scene always contained different objects (except for the two people who were present on all scenes) from the match scenes, with the stipulation that the objects were of the same type as in the match scene (e.g. if the match scene used the sea objects then so did the target scene). On experimental target scenes the tri-axial object was always in the centre of the screen and the bi-axial object was to the left or right. For the experimental trials the participant always had to describe the location of the bi-axial object. This meant that the participants could either describe the location of the bi-axial object as in front of or behind the tri-axial object (if they used an intrinsic reference frame) or to the right of or to the left of the bi-axial object (if they used a relative/deictic reference frame). The spatial relationship between the tri-axial and bi-axial objects was always different on the target scene from the match scene. The arrangements of the objects on the target scenes was counterbalanced across the three conditions so that the bi-axial object was to the left an equal number of times as to the right of the tri-axial object and so that the bi-axial object was intrinsically in front of the tri-axial object and intrinsically behind the bi-axial object an equal number of times. This resulted in a between-items design because the spatial relations between the objects were different for different conditions (all objects appeared an equal number of times in the target scenes in each condition, however).

The match scenes for the intrinsic prime reference frame condition all had the tri-axial object as the reference object. This is because the bi-axial object was unable to support the intrinsic reference frame. The match scenes for the relative/deictic prime reference frame condition all had the bi-axial object as the reference object this was so bi-axial and tri-axial objects were in each role(figure and reference object) an equal number of times the throughout experiment.

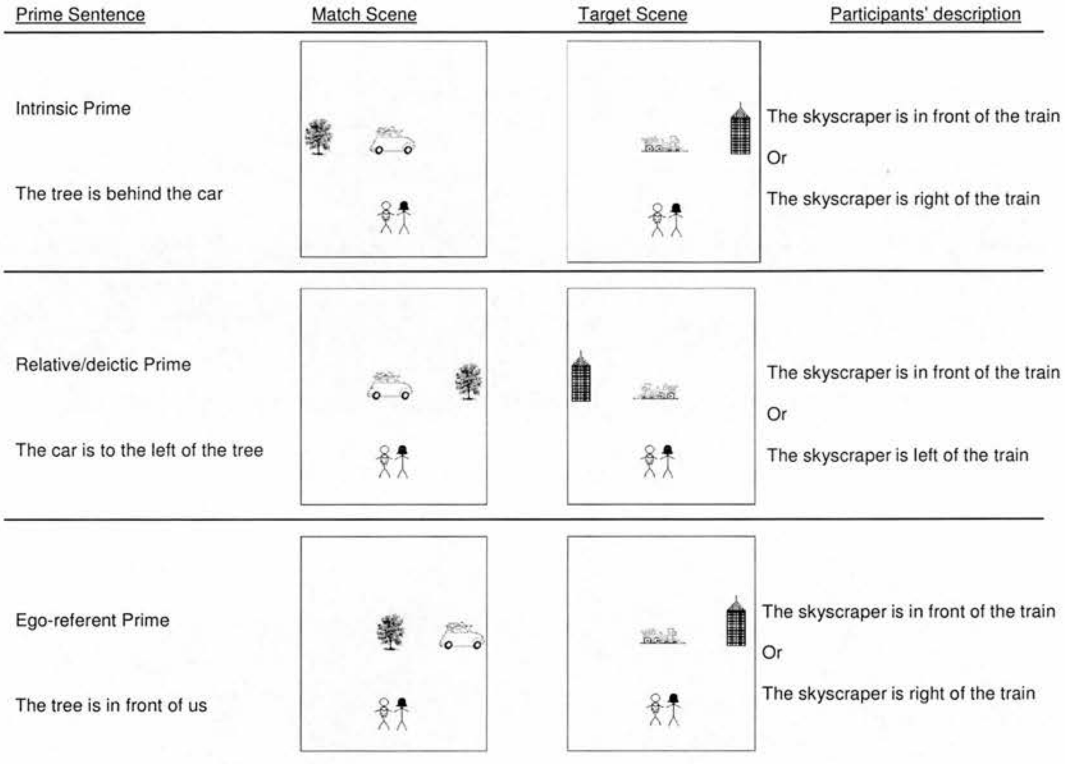


Figure 33: The Experimental conditions from Experiment 4.

The ego-referent reference frame condition match scenes all had the two people as the reference objects. Participants were told that these two people represented the confederate and themselves and so the confederate described the figure object as being *in front of us*. Half of these match scenes had the bi-axial object as the figure object and half had the tri-axial object as figure object, again this was so that each type of object was the figure and reference object an equal number of times. The three conditions with each of their scenes are exemplified in Figure 33. Notice that the prime sentence is always an accurate description of the match scene and that the spatial relationships between the bi-axial and tri-axial objects is different for the target and match scenes.

The 54 filler trials all had match scenes that were an incorrect match for the prime sentence. The objects for the match scenes were the same as described by the prime sentence, but the spatial relationship was different. The participant therefore had to compute the spatial relationship between the objects in order to complete the match task. 18 of the filler trials had a relative reference frame prime sentence, 18 had an intrinsic reference frame prime sentence and 18 had the ego-referent prime sentence. On the filler trials the object on the target scene, which the participant had to describe the location of was always the tri-axial object. This meant that on the filler trials there was no possibility of using the intrinsic reference frame because the bi-axial object was the reference object and these objects have no horizontal intrinsic axes. Figure 34 exemplifies the filler trials, using an intrinsic prime.

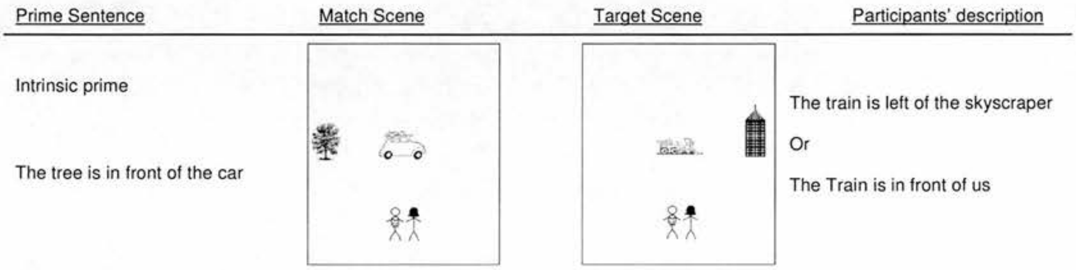


Figure 34. An example of the filler trials from Experiment 4, using an intrinsic prime, where the prime sentence does not describe the match scene.

4.2.3. Procedure

The participant was introduced to the confederate as if she was another participant. The confederate and the participant were then seated at computers on two desks side by side, with a divide between them to prevent them seeing each others' screens. Participants positioned themselves at a comfortable viewing distance from the computer screen. The participant and the confederate were told that they would be describing the location of objects to each other and then deciding whether a scene on their screen matched the description of their partner's scene. They were told that they would take it in turns to be the person describing and the person matching. In addition they were told that all the pictures contained two people at the bottom of the screen which represented the two participants, and that they should treat themselves as part of the scene. They were also told that all the pictures would contain two other objects, one in the centre of the screen and one that would be one side or the other of the central object. Participants were told not to use the screen to describe the location of the objects; for example they should not say *the mast is to the left of the screen.*

After hearing instructions the participants pressed the space bar to begin the practice session. The practice session lasted six trials, one for each of the three experimental conditions, and three fillers. Instructions then appeared on the screen signalling the end of the practice session and the start of the experiment.

Each trial proceeded as follows: After participants pressed the space bar to begin they were presented with the match scene. The confederate then described the location of an object which was present on the participant's screen. The participant then decided whether the match scene matched the confederate's description, pressing the M key if they thought it matched and the Z key if they thought it did not match. Participants were told that if they were unsure whether the scene matched or not then they should make their best guess, rather than ask questions. After selection the match scene disappeared (no feedback was given) and a fixation cross appeared in the centre of the screen. This remained on screen for 1000ms.

The fixation cross was then replaced by a word naming the object that was to be the figure object on the target scene. This word stayed on the screen for 1500ms. When the word disappeared it was replaced by the target scene. Participants then had to describe the location of the object which was indicated by the word. After describing the location of the figure object on the target scene participants pressed the space bar and the target scene was replaced with a fixation cross. The fixation cross remained for 500ms, before being replaced with the next match scene. The confederate pressed the space bar after hearing the participant's description; this caused the next prime sentence to appear on their screen. This was the procedure for all practice, experimental and filler trials.

4.2.4 Results

In the analysis participants' first responses were used to code the reference frame they used. If they used one reference frame, but then switched to the alternate reference frame the response was coded as using the first reference frame (for more details see Chapter 3).

In this experiment responses were coded as either using an *intrinsic* reference frame, *relative/deictic* reference frame or *other*. The *other* category included responses where participants described the location of the wrong object or gave a description that did not describe the location of the figure object using a reference frame (e.g. *the windsurfer is travelling towards the buoy*). There were very few responses which did not use an intrinsic or a relative/deictic reference frame: For each of the three prime reference frame conditions intrinsic, relative/deictic and ego-referent the percentage of *other* responses was 5.1%, 8.8%, and 6.9% respectively. Descriptions were coded as relative/deictic if the figure object was described as *left of* or *right of* the reference object (e.g. *the buoy is to the left of the windsurfer*). Descriptions were coded as intrinsic if the figure object was described as *in front of* or *behind* the reference object (e.g. *the buoy is behind the windsurfer*). In order to just analyse responses that used an intrinsic or a relative/deictic reference frame for each participant in each condition, the percentage of intrinsic reference frame responses was divided by the percentage of intrinsic reference frame responses plus the percentage of relative/deictic reference frame responses. This gave a score (reference frame index, RI) between 0 and 1 for each participant in each condition. A score of 1 indicated that all the responses which were valid used an intrinsic reference frame. A score of 0 indicated that all the responses which were valid used a relative reference frame.

Table 7 shows the mean reference frame index for each of the three conditions. There were approximately 0.08RI more responses using the intrinsic reference frame following an intrinsic prime than following a prime using the other two reference frames. The results were analysed using two 1-way ANOVAs, with Prime Reference Frame as the condition (relative/deictic, intrinsic, and ego-referent), one within-participants (F1) and the other between-items (F2). There was a significant effect of Prime Reference Frame by participants and approaching significance by items ($F1(1,22) = 3.40$; $p = .05$; $F2\ (2,51) = 2.90$; $p = .066$). Planned comparisons showed that the difference between the *intrinsic* Prime Reference Frame and the *relative/deictic* Prime Reference Frame conditions approached significance by participants and items ($t1(11) = 2.03$; $p = .069$, $t2(17) = 1.84$; $p = .072$), and that the difference between the *intrinsic* Prime Reference Frame and the *ego-referent* Prime Reference Frame condition also approached significance ($t1(11) = 2.12$; $p = .058$, $t2(17) = 2.25$; $p < .05$). There was no difference between the *relative/deictic* prime condition and the *ego-referent* prime condition ($t1(11) = 0.27$; $p > .05$, $t2(17) = 0.42$; $p > .05$). Participants were slightly more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame than after hearing the confederate use a relative/deictic reference frame or an ego-referent reference frame. Further, participants were just as likely to use a relative reference frame after hearing the confederate use an ego-referent reference frame as after hearing the confederate use a relative/deictic reference frame.

Condition	Relative/deictic prime	Intrinsic prime	Ego-referent prime
Mean Reference frame index	0.32	0.40	0.32
Mean percentage intrinsic responses	28.7%	38.4%	29.2%
Mean percentage relative responses	62.5%	56.5%	63.9%

Table 7: The mean reference frame index, the mean percentage of intrinsic responses and the mean percentage of relative responses for each of the three conditions.

4.2.5. Discussion

The results show that participants are marginally more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame, than after hearing the confederate use a relative/deictic reference frame. This is a result in line with the findings of the experiments in Chapter 3. It confirms that interlocutors do align reference frames during dialogue and contributes to the evidence for the interactive alignment model of dialogue (Pickering & Garrod 2004). Importantly it also shows that the task in this experiment elicited the alignment effect that is central to distinguishing between Levinson's and the Traditional taxonomy of reference frames.

The critical condition in Experiment 4 was the ego-referent prime condition. The results showed that there was no significant difference in the likelihood of the participants using a relative reference frame following the ego-referent prime reference frames than following the relative/deictic prime reference frames. There was a marginally greater likelihood of the participants using an intrinsic reference frame following an intrinsic prime reference frame, compared to following both the ego-referent prime reference frame and the relative/deictic prime reference frame.

This is suggestive that the reference frame that underlies descriptions such as *The car is in front of us* is the same as the reference frame that underlies descriptions such as *The mast is to the left of the train*. This finding supports the Traditional taxonomy, according to which both of these descriptions employ a deictic reference frame. Levinson's taxonomy argues that these two descriptions use different reference frames (intrinsic and relative respectively), but such a position is not congruent with the patterns of alignment shown by participants in this experiment.

It is possible that rather than participants aligning reference frames with the confederate they are aligning where the origin of the reference frame is situated. It could not be that they are aligning on the object used to define the direction parameter of the reference frame because this object changed between the prime scene and target scene in the intrinsic prime condition and alignment was still evident. However, if it were the origin of the reference frame which was being aligned then it would indicate that the origin of the reference frame was of greater importance than argument structure. Participants ignored the argument structure of an utterance and instead aligned the fact that the origin was egocentric. This may be the case; however, it still would seem to be something that does not sit well with Levinson's taxonomy which does not hold that an egocentric origin is represented with any special status over an allocentric origin.

The results of this experiment again show that interlocutors align reference frames. This is in line with the results of Chapter 3, but it also extends upon those findings in several ways. In this experiment the objects on the target scene were always different from the objects on the prime scene. This was not the case for the experiments in Chapter 3. It has been shown that the more things in common between the prime and the target there are, the larger the amount of alignment (e.g. Branigan et al, 2000). This effect extends to semantic relationships between the prime and target, (e.g. if the prime contains a sheep and the target a goat; Pickering & Cleland, 2003). Therefore as there are fewer things in common between the prime and the target in this experiment it may have been expected that the amount of alignment in this experiment would be smaller than that displayed in Chapter 3. However, the amount of alignment in Experiment 4 is comparable to the amount of alignment shown in the first three experiments. This suggests that repeating attributes of the prime may not impact on alignment of reference frame levels. If this is the case the reason may be due to some level of schematisation that is present in spatial language (Herskovitz, 1986, Talmy, 1983). Spatial language is sensitive to some attributes of the figure and reference object, for example the dynamic kinematic features (e.g., Coventry & Garrod, 2004). However, it is also the case that many features of the shape of the object can be ignored and the figure and reference object for projective relations can be thought of as points. In the present experiments (both Chapter 3 and Chapter 4) the figure and reference object shapes are largely irrelevant and for the purposes of spatial language can be treated as equivalent and this may negate any prime-target repetition effect. However, a direct comparison between the experiments in Chapter 3 and the present experiment is not possible as the tasks and materials were slightly different, but investigating the effect of figure and reference object prime target overlap is an area for future research which may yield insights into the level of schematization in spatial language.

The results showed alignment of reference frames, but the experiment was also indicative of alignment at other levels. For example, there seemed to be a great degree of lexical alignment. Some of the participants were unsure of the names of some of the objects initially, or used alternate names (e.g. tower instead of mast) if they saw one of the objects on the target phase before the confederate had named it on a previous trial. However, by the end of the experiment (and mostly quite quickly in to it) all were using the same lexical terms as the confederate. Although lexical alignment was not directly investigated here the pattern of use of lexical items is in keeping with investigations of lexical entrainment (Brennan & Clark, 1996). Further to this, it seems clear that participants were aligning the conceptual representation which determined the spatial relationship between the people at the bottom of the screen and the object in the centre of the screen. Figure 32 shows a typical scene from the experiment with the people at the bottom, a catapult in the centre, and a tree to the left of the catapult. The relationship between the catapult and the people in this experiment was always described by the confederate as *the catapult is in front of us*. This imposed a perspective on the scene as if the catapult was further away and the people were closer and looking into the screen. However, it is an equally valid description to say *the catapult is above us* (which of these would ordinarily be preferred is not clear from this experiment). Yet none of the participants ever chose to describe the central object as above us or even as above the people. This may be because the participants were aligning with the confederate in the interpretation of the central object as being in front of the people.

There are several possibilities as to what representation is being aligned to make participants choose *in front* of over *above*. At the most general level participants may have been aligning the situation model of the scene with the confederate. This would be similar to the representation investigated by Garrod and Anderson (1987) in their maze task. Participants are aligning with the confederate on the imposition of perspective on the scene and choosing to conceptualise it in that way. The parameters of the reference frame are then set based upon that conceptualisation of the scene. At a more specific level, participants may be aligning directly the setting of the parameters of the reference frame. In particular, participants are aligning the orientation parameter so that the front-back axis extends into and out of the screen. At more specific level still participants could be aligning lexically with the confederate. Specifically they could align that the region of space where the central object is located is termed *above*. It is not possible to determine which of these is true, but it seems like this is a type of alignment distinct from reference frame alignment. However, this is speculative as there is no control condition to make a valid comparison. Despite this it remains an interesting question for future research as to how reference frame alignment interacts with the alignment of other representations.

In Experiment 4 the prime sentences were held constant across conditions, whilst the target scenes were varied in the spatial relationship between the intrinsic and the non-intrinsic object. Experiment 5 reversed this arrangement and used the same spatial relationship between items for the target scene across six different lists. There was one further factor changed for Experiment 5: The match tasks in Experiment 5 returned (as in Chapter 3) to being a forced-choice decision between one scene that matched the confederate's description and one scene that did not. This was as opposed to judging whether or not the prime sentence was an accurate description of a single scene. The reason for this is that six participants in Experiment 4 had to be replaced because they would not interpret the match scene using an intrinsic reference frame and as a consequence they were unable to align with the intrinsic reference frame. Presumably when the confederate described a scene like Figure 32 as *The tree in front of the catapult* this group of participants were using a relative reference frame and searching for a tree between the people and the catapult. This is despite the participants being told that all the scenes in the experiment would be in the same spatial arrangement as shown in Figure 32.

This suggests two interesting things: First it suggests that some people have a definite preference for the relative reference frame and are unwilling to consider the use of the intrinsic reference frame (it is possible they are not aware of the possibility of the intrinsic reference frame). In Experiments 1-3 there were also participants who used one reference frame exclusively in preference to the other. This is in line with research that has shown individual difference in reference frame preference (Carlson-Radvansky & Irwin, 1993). It also shows that despite evidence that they are not interpreting their partner's descriptions correctly (i.e. that their description is not an accurate description of any scene and the explicit instruction of the arrangement of objects in scenes) people still do not always change their interpretations (Russell & Schober, 1999).

The results of Experiment 4 reproduce the alignment effect shown in Chapter 3. Interlocutors are more likely to use a reference frame if it has just been used by their partner. In addition, Experiment 4 has extended this and shown that the dialogue paradigm can be used to investigate reference frame representation. The primary purpose of this experiment was to distinguish between the Traditional taxonomy of reference frames and Levinson's taxonomy of reference frames. This was done by including the ego-referent reference frame prime: Levinson's taxonomy predicted that the participants would be more likely to use an intrinsic reference frame than a relative/deictic reference frame following hearing the ego-referent prime. The Traditional taxonomy predicted the reverse. The results showed some support for the Traditional taxonomy: Participants were used relative/deictic reference frame just as much after hearing an ego-referent reference frame as after hearing a relative/deictic reference frame, suggesting that the two utterances use the same reference frame. Experiment 5 looked to confirm the results of Experiment 4 using a within-items design for the target scenes.

4.3. Experiment 5

Experiment 5 was designed as a replication of Experiment 4 using a within-item design. In this experiment the target scenes were held constant across conditions, whereas in Experiment 4 the prime scenes were held constant. In addition, the match task was changed so that it was a forced choice rather than a yes or no task. This was to avoid the loss of data due to participants not interpreting the scene using an intrinsic reference frame. As in Experiment 4 if interlocutors align reference frames then participants should be significantly more likely to use a reference frame when they have just heard the confederate use that reference frame than when they have just heard the confederate use an alternative reference frame. In addition, if

Levinson's taxonomy of reference frames is correct then participants will be more likely to use an intrinsic reference frame after hearing the confederate use a ego-referent reference frame, than after hearing the confederate use a relative/deictic reference frame. If the Traditional taxonomy of reference frames is correct then participants will use a relative/deictic reference frame as much after hearing the confederate use an ego-referent reference frame as after hearing the confederate use a relative/deictic reference frame.

4.3.1. Participants

16 students of the University of Edinburgh were paid volunteers in the experiment, which lasted 20 minutes. All were native English speakers who had not participated in Experiments 1-4. The confederate was a female postgraduate at the University of Edinburgh. Two further participants were excluded from analysis because of a recording error.

4.3.2. Materials and design

The target scenes were formed in the same way as Experiment 4. The match scenes were almost the same as in Experiment 4, except in this experiment two scenes were presented on the screen during the match phase. Figure 35 shows one of the match scenes from Experiment 5. For half of the match scenes the left hand scene was described by the confederate, for the other half the right hand scene was described. The left hand scene and right hand scene were surrounded by boxes to separate them and a fixation cross was presented in the centre of the screen.

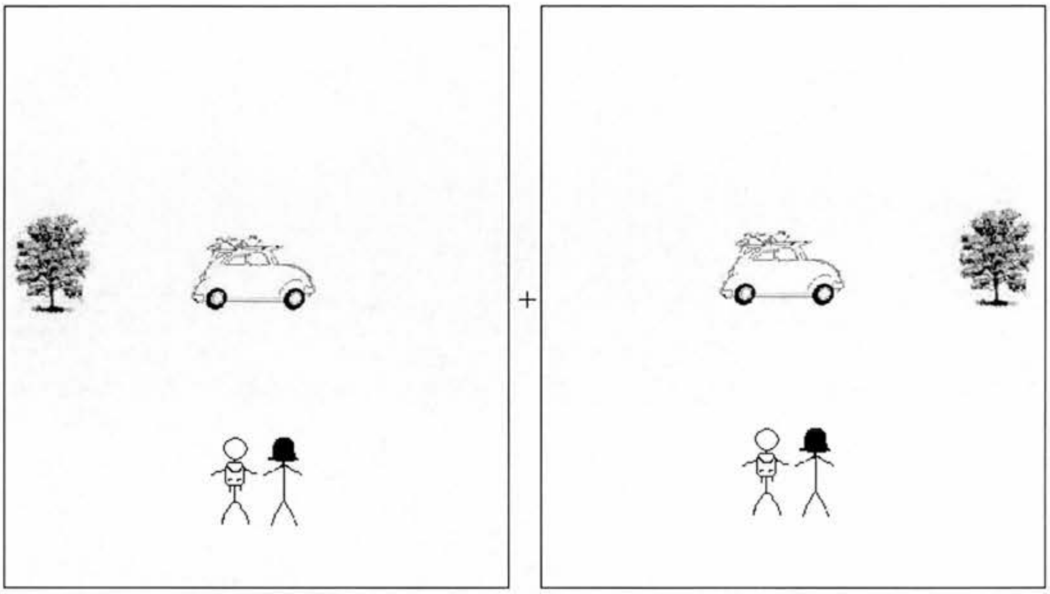


Figure 35: A match task from Experiment 5.

The two pictures that made up each match scene differed only in the spatial relation between the bi-axial and the tri-axial object, (e.g. in Figure 35 the tree is in front of and to the right of the car in the right hand scene, but behind and to the left of the car in the left hand scene).

There was one within-participants and within-items factor of prime reference frame (relative/deictic, intrinsic, and ego-referent). Six lists were created where each of the tri-axial and bi-axial pairs for the target scenes (e.g. hovercraft and the island) were paired with a match scene that had different tri-axial and bi-axial objects (e.g. the windsurfer and the buoy). Each target scene pair appeared in each of the three conditions paired with a different match scene bi-axial and tri-axial pair. Figure 36 exemplifies the three experimental conditions from Experiment 5. The three target scenes are the same. The combination of tri-axial and bi-axial items in the match scenes is different for each of the conditions. Notice that two of the conditions use the car as the tri-axial item and two conditions use the mast as the bi-axial item. This is because there were only three different types of each item within the land and sea types. However, the match scene items were counterbalanced across conditions so that each bi-axial and each tri-axial item appeared the same number of times in the match scenes throughout the experiment. Everything else was the same as Experiment 4.

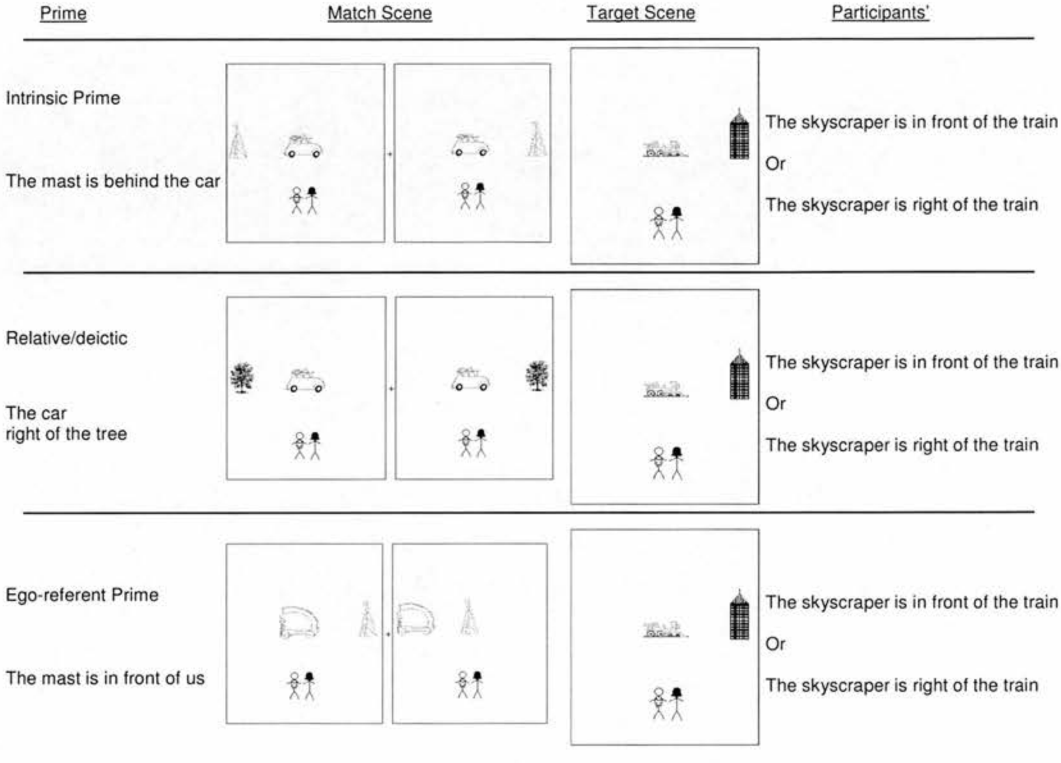


Figure 36: The experimental conditions from Experiment 5

4.3.3. Procedure

The procedure was the same as Experiment 4, except that participants were instructed to choose which of the two scenes on the match task matched the confederate's description. If they thought it was the scene on the left they pressed the Z key. If they thought it was the scene on the right they pressed the M key.

4.3.4. Results

Once again participants' first responses were used for analysis. Responses were coded as either using an *intrinsic* or *relative/deictic* reference frame. In addition some responses used an *ego-referent* response (e.g. *The tree is to the left of us*). These responses use an intrinsic reference frame according to Levinson and a deictic reference frame according to the Traditional taxonomy; these were coded separately as *ego-referent*. There were, however, relatively few *ego-referent* responses in each condition: Following a relative prime there were 4.5% ego-referent responses, following an intrinsic prime there were 4.2% ego-referent responses and following the ego-referent prime there were 2.8% ego-referent responses. The difference between the conditions was non-significant (F1<1.5; F2<2). This pattern of results indicates that participants were not using the ego-referent construction because they were aligning with the confederate when they used the ego-referent reference frame. In fact, the trend is slightly in the opposite direction to that prediction. As a consequence, these responses were omitted from further analysis. There were also a number of *other* responses. For example, when the wrong figure object was described, or the participant replied, for example, *The hovercraft is travelling towards the island.* These were also relatively rare in occurrence: following *relative/deictic*, *intrinsic* and *ego-referent* primes there were 5.6%, 2.4%, and 1.7% other responses respectively. The difference between the conditions was non-significant (F1<1.5; F2<2.6). As a consequence these were omitted from further analysis. Analysis was performed only on the relative and the intrinsic responses. This was done by creating a reference frame index (RI) as in Experiment 4.

Condition	Relative/deictic prime	Intrinsic prime	Ego-referent prime
Mean Reference frame index	0.24	0.33	0.24
Mean percentage intrinsic responses	20.5%	31.6%	23.3%
Mean percentage relative responses	69.1%	61.8%	72.2%

Table 8: The mean reference frame index for each of the conditions in Experiment 5.

Table 8 shows the mean reference frame index for each of the three conditions. There was a score difference of approximately 0.09RI between the intrinsic prime condition and the other two conditions. This indicates a preference for using the intrinsic reference frame more following hearing an utterance using intrinsic reference frame compared to hearing an utterance using another reference frame. This is an effect of a similar size to that reported in Experiment 4.

The reference frame index was analysed using two repeated measures one-way ANOVAS one for participants (F1) and one for items (F2), with prime reference frame (relative/deictic, intrinsic and ego-referent) as the condition. There was a significant main effect of Prime Reference Frame ($F1\ (2,30) = 3.46$; $p < .05$; $F2\ (2,34) = 11.47$; $p < .01$). Planned comparison t-tests showed that there was a significant difference between the *intrinsic* prime condition and the *relative/deictic* prime condition ($t1(15) = 2.54$; $p < .05$, $t2(17) = 3.76$; $p < .01$). Participants were more likely to use an intrinsic reference frame following hearing an utterance using an intrinsic reference frame, than following hearing an utterance using a relative/deictic reference frame. There was also a significant difference in the reference frame index between the *intrinsic* prime condition and the *ego-referent* prime condition ($t1(15) = 2.15$; $p < .05$, $t2(17) = 3.8$; $p < .01$). Participants were more likely to use an intrinsic reference frame following hearing an utterance using an intrinsic reference frame, than following hearing an utterance using a ego-referent reference frame. Finally, there was no significant difference in the reference frame index between the *relative/deictic* prime condition and the *ego-referent* prime condition ($t1(15) = 0.13$; $p > .05$, $t2(17) = 0.16$; $p > .05$) Participants were just as likely to use a relative reference frame following hearing an utterance using ego-referent reference frame as following hearing an utterance using a relative reference frame.

4.3.6. Discussion

The results of this experiment echo those of Experiment 4. Participants were more likely to use an intrinsic reference frame following hearing an utterance using an intrinsic reference frame, than following hearing an utterance using either of the other two reference frames. This once again confirms that interlocutors align reference frames, but importantly also supports the Traditional taxonomy of reference frames over Levinson's taxonomy.

The pattern of alignment shown in Experiment 5 mirrored the pattern shown in Experiment 4 and the size of the alignment effect was approximately the same. The fact that there is no significant difference in the proportion of reference frames used following an ego-referent prime compared to following a relative/deictic prime suggests that the two are represented as the same reference frame. This suggests that reference frames should be categorised in terms of the origin and how the directional axes are determined, rather than the argument structure.

4.4. Experiment 6

The Traditional taxonomy divides reference frames up into intrinsic, deictic and absolute. These correspond to when the reference frame is based upon the intrinsic axes of an object, when the reference frame is egocentric, and when the reference frame is defined by features of the environment, respectively. However, one of the issues with the Traditional taxonomy is what is included as part of a deictic reference frame. Miller and Johnson-Laird (1976) included addressee as well as speaker centred reference frames as being deictic reference frames. It is, however, also possible that any human centred reference frame is deictic. For example, hearing the sentence *The ball is in front of the man* may increase the likelihood of a person using the reference frame *The ball is to the left of the car* (when the ball is to the speaker's left of the car). What this would indicate is the deictic reference frame is represented as being centred on any person.

This experiment was designed to test whether or not deictic reference frames include any reference frame whose origin is situated on a person. Experiment 5 was repeated, but rather than having two people representing the confederate and the participant at the bottom of the screen, there was a single person who was referred to as *the man*. The ego-referent condition used in Experiments 4 and 5 was now replaced with a *man-referent* condition. If the deictic reference frame includes reference frames that are centred on any person then the participants will be just as likely to produce a relative/deictic utterance following hearing the confederate use the man-referent reference frame as following hearing the confederate use a relative/deictic reference frame.

4.4.1. Participants

22 students of the University of Edinburgh were paid volunteers in the experiment, which lasted 20 minutes. All were native English speakers who had not participated in Experiments 1-5. The confederate was a female postgraduate at the University of Edinburgh. Two further participants were excluded from analysis as a result of them producing responses of the form *The car is travelling away from tree* throughout the experiment. Descriptions such as this do not use a reference frame and therefore can not be coded for the analysis.

4.4.2. Materials and Design

The materials and design were exactly the same as Experiment 5, except that the two people at the bottom of each screen were replaced with a single person who was referred to as *the man.* The three reference frame prime conditions were therefore *intrinsic*, *relative/deictic*, and *man-referent* (the latter replacing the ego-referent condition). The match and target scenes in the experiment now looked like the one shown in Figure 37.

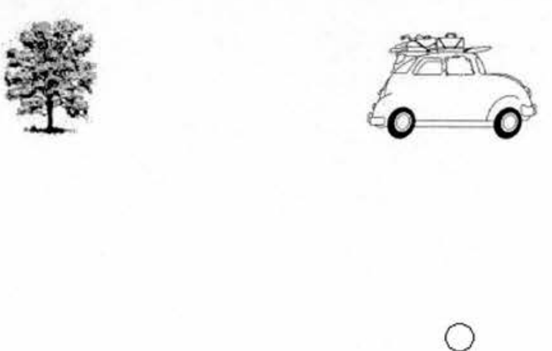


Figure 37: One of the scenes from Experiment 6.

4.4.3. Procedure

The procedure was exactly the same as Experiment 5 except that the participants were not told that the person in the scenes represented him or herself.

4.4.4. Results

Participants' first responses were used for analysis. Responses were coded as either intrinsic, relative/deictic, *man-referent* when participants used a construction like *The tree is left of the man*, or an *other* when the response could not be coded.

Most responses used either an intrinsic or a relative reference frame. There were very few *man-referent* responses, 0.5%, 0.4% and 1.4% mean percentage of responses for the relative, intrinsic and man-intrinsic conditions respectively. There was no significant difference between the conditions (F1<1; F2<1.2) and these responses are omitted from further analysis. There were also very few *other* responses 5.2%, 4.3%, and 5.1% mean percentage of other responses for each of the conditions, relative, intrinsic and man-intrinsic respectively. Again there was no significant difference between the three conditions (F<1.1; F2<1) and so these responses are omitted from further analysis.

Condition	Relative/deictic prime	Intrinsic prime	Man-referent prime
Mean Reference frame index	0.42	0.49	0.42
Mean percentage intrinsic responses	38.3	41.5	36.4
Mean percentage relative responses	56.7	49.5	56.3

Table 9: The mean reference frame index scores for each of the three conditions in Experiment 6.

Analyses were performed only on the relative and the intrinsic responses. This was done by creating a reference frame index as in Experiment 4. Table 9 shows the mean score on the reference frame index for each condition. Once again there was a preference for participants to use an intrinsic reference frame more after hearing the confederate use an intrinsic reference frame.

The reference frame index scores were analysed using two repeated measures one-way ANOVAs one for participants (F1) and one for items (F2) with prime reference frame as the condition (relative/deictic, intrinsic and man-referent). There was a marginally significant main effect of Prime Reference Frame by participants and a significant main effect of prime reference frame by items ($F1\ (2,44) = 3.065$; $p = .057$; $F2\ (2,34) = 6.8$; $p < .01$). Planned comparisons revealed that there was a significant difference between the *intrinsic* and the *relative/deictic* Prime Reference Frame conditions ($t1(21) = 2.39$: $p < .05$, $t2(17) = 5.04$; $p < .01$). Participants were more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame than after hearing the confederate use a relative/deictic reference frame. There was also a marginally significant difference in the reference frame index between the *intrinsic* and the *man-referent* Prime Reference Frame conditions by participants and a significant effect by items ($t1(21) = 2.06$; $p = .055$, $t2(17) = 2.76$; $p < .05$). Participants were slightly more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame than after hearing the confederate use a man-referent reference frame. The difference between the relative/deictic prime reference frame condition and the man-referent prime reference frame condition was non-significant ($t1(21) = 0.06$; $p > .05$, $t2(17) = 0.34$; $p > .05$) . Participants were just as likely to use an intrinsic reference frame after hearing the confederate use a relative reference frame as after hearing the confederate use a man-referent reference frame.

4.4.5. Discussion

In keeping with the findings of the rest of this thesis the results showed that interlocutors align reference frames during dialogue. Participants were significantly more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame, than after hearing the confederate use a relative/deictic reference frame.

The experiment, however, was specifically designed to investigate the representation of reference frames that are not egocentric, but are based upon the axes of another person. The results suggest that reference frames that are based upon the axes of a person are represented as deictic reference frames, rather than intrinsic reference frames. Participants were more likely to use an intrinsic reference frame after hearing the confederate use an intrinsic reference frame than after hearing the confederate use the man-referent reference frame. More importantly participants' pattern of reference frame use was the same following hearing the confederate use the man-referent reference frame as following hearing the confederate use a relative/deictic reference frame. This suggests that the two use the same reference frame.

4.5. Experiment 7

The experiments in this chapter have aimed to distinguish between the Traditional taxonomy of reference frames and Levinson's taxonomy, with the evidence pointing in favour of the former. However, the results of Experiment 6 suggest an explanation of the patterns of alignment shown that does not entail a dismissal of Levinson's taxonomy. Experiments 4-6 have all included an *ego-referent* (*man-referent* in Experiment 6) condition which, according to Levinson's taxonomy uses an *intrinsic* reference frame, but according to the Traditional taxonomy uses a *deictic* reference frame. The reference frame that the participants were more likely to use following the *ego-referent* description is argued to be the same as the reference frame underlying the *ego-referent* utterance. However, the reference object (either the two people depicting the participant and the confederate, or the man) was always located at the bottom of the screen, which meant that for the *ego-referent* and *man-referent* conditions the reference object was in a different position on the match scene to the reference object on the target scene. For the other two conditions (*relative/deictic* and *intrinsic*) the reference object was in the same position on the match scene as on the target scene although not oriented the same.

The results have shown that participants have preferred to use a relative/deictic reference frame in the *ego-referent* condition and this has been interpreted as evidence in favour of the Traditional taxonomy as opposed to Levinson's taxonomy. However, it may be that the effect was due to the different placement of the reference object between the match and the target scene in the *ego-referent* condition.

This Experiment was designed to identify whether or not the pattern of alignment in the previous three experiments was caused by the reference object being in a different position in the target scene to the match scene. This Experiment had four Prime Reference frame conditions. Three of these were the *intrinsic*, *relative*, and *ego-referent*, as used in Experiments 4 and 5. The fourth condition used an *intrinsic* prime reference frame, but the reference object on the match scene was in the position occupied by the people on the match scenes in the other three conditions: this is the *intrinsic-different* condition. Figure 38 shows the conditions used in Experiment 7. The correct match scene is underlined.

Note that in the *intrinsic-different* prime condition the reference object on the match scene is in a different position to the reference object on the target scene. The other three conditions are the same as in Experiments 5 and 6. As in previous experiments it is predicted that there will be an effect of reference frame alignment. Participants will be more likely to use a reference frame after hearing the confederate use that reference frame, than after hearing the confederate use an alternative reference frame.

The experiment also investigates whether or not participants are more likely to align with the confederate on the use of the intrinsic reference frame, when the reference object's position is held constant between match and target scene, compared to when the reference object's position is different from the match to the target scene. If interlocutors align reference frames independently of position, then participants should be more likely to use a reference frame when they have just heard the confederate use that reference frame compared to an alternative reference frame, regardless of whether the reference object is in the same position or a different position on the match scene and target scene

The third prediction is that the *ego-referent* condition uses the same reference frame as the *relative/deictic* condition. Experiments 4 and 5 have shown no difference in participants' number of intrinsic responses in these conditions. Therefore, it is predicted that participants will use an intrinsic reference frame as often following hearing the confederate use a *relative/deictic* reference frame as following hearing the confederate use an *ego-referent* reference frame.

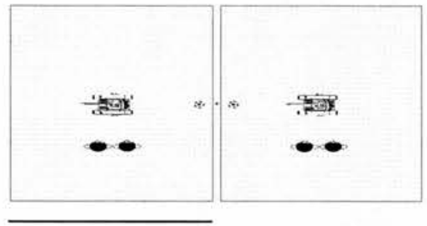
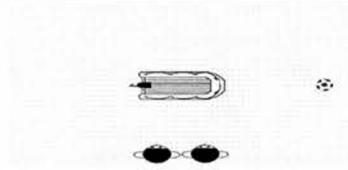
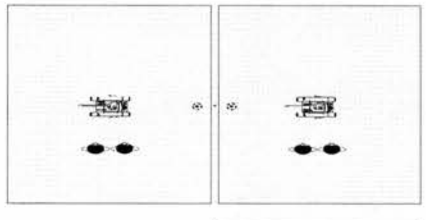
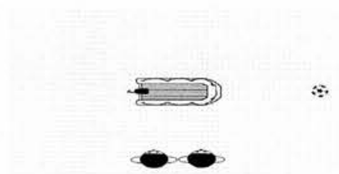
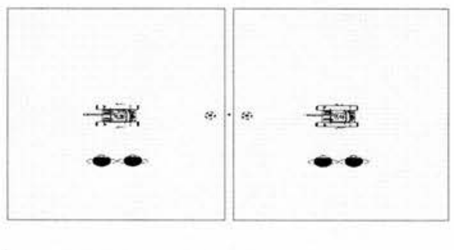
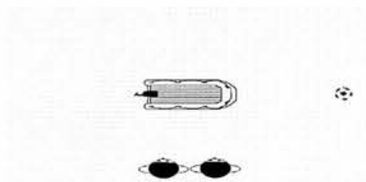
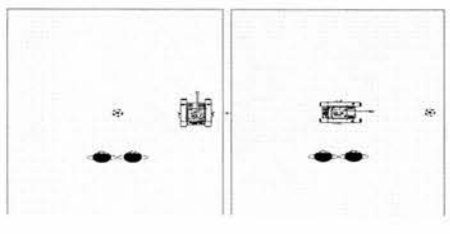
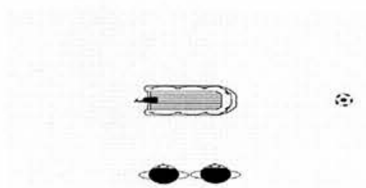
Prime Sentence	Match Scene	Target Scene	Participants' description
Intrinsic Prime The ball is behind the tank			The ball is in front of the dinghy Or The ball to the right of the dinghy
Relative/deictic Prime The ball to the left of the tank			The ball is in front of the dinghy Or The ball is right of the dinghy
Intrinsic-different Prime The ball is in front of the tank			The ball is in front of the dinghy Or The ball is left of the dinghy
Ego-referent Prime The ball is in front of us			The ball is behind the dinghy Or The ball is left of the dinghy

Figure 38: The conditions used in Experiment 7

4.5.1. Participants

18 students of the University of Edinburgh were paid volunteers in the experiment, which lasted approximately 15 minutes. All were native English speakers who had not participated in Experiments 1-6. The confederate was a male postgraduate at the University of Edinburgh.

4.5.2. Materials and Design

In this experiment each match and target scene contained three objects: a football, two people side by side, viewed from overhead, and one of 10 objects with intrinsic sides that was also viewed from overhead (see appendix E).

There was one within-participants and within-items factor of Prime Reference frame, with four levels (*intrinsic, relative/deictic, intrinsic-different,* and *ego-referent*). The four conditions are shown in Figure 38. In the top scene the confederate described the location of the ball in the left hand picture of the match scene, using an *intrinsic* reference frame (e.g. *The ball is behind the tank*). The participant then chose the left hand scene as matching the confederate's description. This was counterbalanced throughout the experiment so that half the time the matching scene was the right hand scene and half the time the left hand scene. The participant then saw the target scene and described the location of the football using an *intrinsic* reference frame (e.g. *the ball is in front of the dinghy*) or a *relative/deictic* reference frame (e.g. *the ball is right of the dinghy*). In the second scene, of Figure 38, the confederate described the location of the ball in the right hand picture of the match scene using a *relative/deictic* reference frame (e.g. *the ball is to the left of the tank*). The participant then chose the right hand scene as matching the confederate's description. The participant then saw the target scene and described the location of the football using an *intrinsic* reference frame (e.g. *the ball is in front of the dinghy*) or a *relative/deictic* reference frame (e.g. *the ball is right of the dinghy*). The same process was followed for the *ego-referent* and *intrinsic-different* conditions (details in the bottom two scenes of Figure 38).

Note that the two pictures in the match scenes differed only in the spatial relationship between the ball and the reference object. For example, in the top trial of Figure 38 the ball is *behind* and to the *right* of the tank on the left hand scene and *in front of* and to the *left* of the tank on the right hand scene.

Note also that the reference object used in the match scene was always different to the reference object used in the target scene; however, the pairings of the reference objects was different for all conditions. This was counterbalanced across nine lists, so that each reference object appeared in each condition on the match scene, paired with each of the other reference objects on the target scene. In each list all reference objects appeared once per condition in the match scene, and once per condition in the target scene. This meant that each object was seen in the match scenes four times and seen in the target scenes four times. As there were 10 objects this made a total of 40 trials (10 per condition) in the experiment. All of the trials were experimental; there were no fillers

4.5.3. Procedure

The participant was introduced to the confederate as if he was another participant in the experiment. The confederate and participant were then seated at a comfortable distance from two computers positioned as in Experiments 4-6, with a screen between the two so that they were unable to see each other. The participant and confederate were told that they would be describing the location of objects to each other so that their partner could choose which of two scenes matched the description, and that they would take it turns to be the matcher and the describer.
After pressing the space bar there were eight practice trials, two from each of the four conditions. After completion of the practice trials the experiment began. Each trial proceeded as follows:

After pressing the space bar the confederate was presented with a prime sentence in the centre of the screen which he read to the participant. After participants pressed the space bar they were presented with a match scene and were instructed to choose which of the two scenes matched the confederate's description. If participants thought it was the left hand scene they pressed the Z key; if they thought it was the right hand scene they pressed the "/" key. If they were unsure which scene matched the description they were instructed to choose which one they thought matched closest. Participants were instructed not to ask any questions of each other except to repeat the description.

When the participant had chosen one of the match scenes, by pressing either the Z or "/" key, the match scenes disappeared (no feedback was given) and a fixation cross appeared in the centre of the screen. This remained on screen for 1000ms. Following this, the fixation cross was replaced with a target scene. The participants were instructed to describe the location of the ball on the target scene without using the ball's position on the screen (e.g. participants were told not to say *The ball is on the left of the screen*). When the participant had described the location of the ball they pressed the space bar and the target scene was replaced by the next match scene.

When the confederate had heard the participant's description of the target scene they pressed the space bar and the next prime sentence appeared on their screen. Each trial progressed in the same manner until the end of the experiment.

4.5.4. Results

As in the previous experiments participants' first responses were used in the analysis. Responses were coded as either using a *relative/deictic* reference frame, if the ball was described as to the *left* or *right* of the reference object, and as using an *intrinsic* reference frame, if the ball was described as *in front* or *behind* the reference object. The majority of responses used an intrinsic or a relative/deictic reference frame, but there were a minority of responses that were of the form *the ball is to the left of us* (4.7%). There was no significant difference between the mean number of these that occurred per condition ($F1<1$; $F2<1.9$) and so these responses were omitted from the analysis. There were also a minority of responses that could not be coded as using a reference frame (e.g. *the cannon is pointing at the ball*) (11.9%). There was no significant difference between the mean number of these that occurred per condition ($F1<1$; $F2<1$) and so these are omitted from the analyses.

The mean percentage intrinsic responses and mean percentage relative/deictic responses for each participant in each condition were converted into a single score for each condition by dividing the percentage of intrinsic responses by the percentage of intrinsic responses plus the percentage of relative responses. This yielded a reference frame index (RI) score between 1 (all intrinsic) and 0 (all deictic).

Table 10 shows the mean reference frame index and the mean percentage intrinsic and relative/deictic scores for each of the four conditions. Participants were more likely to use an intrinsic reference frame following hearing the confederate use an intrinsic reference frame than following hearing the confederate use a relative/deictic reference frame. The mean reference frame index scores were analysed using two one-way, within-participants (F1), and within-items (F2) ANOVAs, with Prime Reference frame (intrinsic, relative/deictic, intrinsic-different and ego-referent) as the factor and levels.

Condition	Intrinsic prime	Relative/deictic prime	Intrinsic-different prime	Ego-referent prime
Mean Reference frame index	0.77	0.60	0.68	0.68
Mean percentage intrinsic responses	62.8	50.6	57.8	55
Mean percentage relative responses	18.9	33.9	27.8	26.1

Table 10: The mean reference frame index scores, percentage intrinsic and deictic responses for each of the four conditions in Experiment 7.

There was a significant main effect of Prime Reference frame by both participants and items ($F1(3,51) = 5.4$; $p < .01$, $F2(3,27) = 3.1$; $p < .05$). Planned comparisons showed that there was significant difference between the *relative/deictic* and *intrinsic* Prime Reference frame conditions both by participants and items ($t1(17) = 4.4$; $p < .01$, $t2(9) = 2.6$; $p < .05$). Participants used an intrinsic reference frame more following hearing an *intrinsic* reference frame (0.77 RI), than following hearing a *relative/deictic* reference frame (0.66 RI). There was also a significant difference between the *intrinsic* and the *intrinsic-different* Prime Reference Frame conditions both by participants and items ($t1(17) = 2.28$; $p < .05$, $t2(9) = 2.53$; $p < .05$).

Participants used an intrinsic reference frame more often following hearing an *intrinsic* reference frame when the reference object position was held constant from the match scene to the target scene (0.77 RI), than when the reference object position changed from match scene to target scene (0.68 RI). There was no significant difference between the *relative/deictic* and *ego-referent* Prime Reference Frame conditions ($t1(17) = 1.47$; $p > .05$, $t2(9) = 0.92$; $p > .05$). Participants used an intrinsic reference frame just as often following hearing the confederate use a *relative/deictic* reference frame, as following hearing the confederate use an *ego-referent reference* frame.

4.5.5. Questionnaire responses

Following this experiment, participants were given a short questionnaire to fill out. This was to discover what they thought the intention of the experiment was and also if they had any strategies for completing the task. 19 of the 24 participants completed the questionnaire and the results are reported here.

The questionnaire consisted of a question enquiring what the participants thought the purpose of the experiment was, a question enquiring if they had a strategies for completing the task, a question which asked if they thought their partner (the confederate) had any strategies for completing the task and finally a space for any other comments about the experiment.

There were a large number of different responses to the question regarding the purpose of the experiment. 7 of the participants put that they thought the experiment was investigating how people communicate object location, of that 7, 3 explicitly referred to the choice between object centred and person centred descriptions. These participants thought the experiment was specifically investigating the production of spatial descriptions. A further 3 participants thought the experiment was designed to investigate how people comprehend other people's spatial descriptions. Only three participants thought that the experiment was interested in how a partners descriptions affect the descriptions they produced. The remaining participants each thought the experiment was investigating something else, but each purpose was suggested only by one participant. These results show that the majority of participants were not wholly able to guess the purpose of the experiment.

There were three different kinds of strategies which most participants used: Participants said they used both front and behind to describe the location of the ball (4), or left and right to describe the location of the ball (1) or both of these strategies (3). 4 other participants responded that they did not have a strategy. A single participant said that he just described the location of the object on the screen as fast as possible and two further participants said that they used the *best* description possible each time.

Some of the participants (6) thought that confederate used the same strategy as they did to describe the location of the ball. However, the confederate had no strategy as the trials were presented randomly. 5 participants noticed that the confederate was not using a strategy and 4 more noticed that the confederate was varying between using left and right and front and behind spatial terms. However, one participant thought that the confederate used mainly left and right spatial terms and two thought that the confederate used mainly in front of and behind spatial terms; in reality the confederate used each half of the time.

4.5.6. Discussion

This experiment was designed to investigate whether or not the *ego-referent* primes used in Experiment 4-6 used an intrinsic reference frame, but did not prime use of the intrinsic reference frame in participants because the reference object was in a different position on the match scene and target scene. The results of Experiment 7 suggest that this may be the case: Participants used an intrinsic reference frame more following the confederate using an intrinsic reference frame when the position of the reference object was held constant between the match and target scenes, than when it changed position.

There was an effect of reference frame alignment. Participants were more likely to use an intrinsic reference frame following hearing the confederate use an intrinsic reference frame than following hearing the confederate use a relative/deictic reference frame. This is an effect common to all of Experiments 1-7. In this Experiment the size of the alignment effect was slightly greater (0.17 RI) than in previous experiments. This may be because the items in this experiment were presented from an overhead perspective. The overhead perspective may make the use of an intrinsic reference frame easier and so the alignment effect may be greater. This may also explain why there was a larger use of the intrinsic reference frame overall in Experiment 7 compared to the previous three experiments. An overhead perspective is not the canonical way a person experiences an environment and as such this may make using a relative/deictic reference frame, from this unusual perspective, more difficult. In contrast, in the previous experiments the scenes were viewed from a more naturalistic perspective.

In this Experiment there was no difference between the *ego-referent* prime and the *relative/deictic* prime. This effect has been found in Experiments 4 and 5 and has been interpreted as evidence for sentences such as *the ball is in front of us* using the same reference frame as sentence such as *the ball is to the left of the car*. This is in turn has been argued to be evidence in favour of the Traditional taxonomy of reference frames. However, it may be that *ego-referent* sentences use an intrinsic reference frame, but did not cause participants to do so because the reference object was in a different position on the match and target scenes.

The results suggest that interlocutors align reference frames and that the amount of alignment is influenced by whether or not the reference object maintains its position during comprehension and production. This can be seen as analogous to evidence from other confederate priming studies which have shown that the more attributes in common between the prime and the target the greater the level of alignment (e.g. Cleland & Pickering, 2003).

However, it does suggest that we cannot be completely confident in the conclusion from previous experiments that sentences such as *the ball is in front of us* use a deictic reference frame because the lack of alignment may have been caused by the changing position of the reference object. However, there is another possible reason for the difference between the *intrinsic* and *intrinsic-different* conditions. The difference may have occurred because in the *intrinsic-different* condition there was no reference frame competition. In the *intrinsic* condition the football could be described as *in front of the X* or *to the left of the X* (where X refers to one of the experimental items). Thus when using the *X* as the reference object there were two possible reference frames that could be used. However, in the *intrinsic-different* condition there was no such competition; the football could be described only as *in front of the X* (i.e. there was no description using the *relative/deictic* reference frame that would have made the *X* reference object). This is a manipulation which needs to be made in future experiments.

The results of this experiment again confirm that interlocutors align reference frames. However, it does suggest that the reason *ego-referent* primes produced more *relative* reference frame responses in Experiments 4-6, was because the reference object was in a different position on the target scene and the match scene. This conclusion is not certain and sentences such as *the ball is in front of me* may use the same reference frame as sentences such *the ball is to the left of the car*. It means that future research needs to investigate the relationship between the position of the reference object on the prime and target trials and the use of *ego-referent* primes. This will help explain which of the two taxonomies is the correct way of classifying reference frames and how different aspects of reference frames interact in processing.

Chapter 5: The Distance Parameter

5.0. Chapter overview

This chapter is a departure in investigation from the use of reference frames in dialogue, which has been the focus of the previous chapters. Instead, this chapter focuses on an investigation of the distance parameter of the reference frame. Logan and Sadler (1996) have argued that reference frames are represented by the setting of at least four different parameters: origin, orientation, direction and distance (they termed it the scale parameter, but Carlson (2003) has suggested that distance is better as the term scale implies units of discrete intervals) .

This chapter presents three experiments that investigated how the distance parameter of a reference frame is set. The first experiment (Experiment 8), used a paradigm where participants were asked to imagine a spatial relationship between two objects, one above the other, and then report the imagined distance between a figure object and a reference object, which were either functionally related or functionally not related, and either interacting or not interacting. The aim of the experiment was to investigate whether the function has an effect on setting the distance parameter and also whether the reference frame used, either intrinsic or relative, has an effect on setting the distance parameter. The results showed that when a figure and a reference object are functionally related the distance parameter is set as shorter than when the objects are not functionally related. The results also showed that the choice of reference frame did not affect the setting of the distance parameter. Experiments 9 and 10 used a visual search paradigm (Treisman & Gelade, 1980, Treisman & Gormican, 1988) to investigate whether apprehending the distance between a figure and a reference object requires the use of attention, as is the case for apprehending the spatial relationship between two objects (Logan, 1994). The results of these experiments showed that setting the distance parameter in order to apprehend a spatial relationship between two objects does require the use of attention. The results also suggested that the more parameters of the reference frame that are required to apprehend a spatial relationship, the more difficult the process of apprehension is. There was also some suggestion that the intrinsic reference frame is more sensitive to spatial relations between objects that are closer together.

5.1. Introduction

Reference frames are axial co-ordinate systems which define four parameters: direction, orientation, origin and distance (Logan & Sadler, 1996). These parameters can be set with different values e.g. the orientation of a reference frame can be set according to the axes of the reference object, as in the intrinsic reference frame, or according to the axes of a viewer, as in the relative reference frame. There has, however, been very little work investigating how the distance parameter is set. In fact, for many spatial terms Logan and Sadler (1996) have argued that the distance parameter is not set (e.g. *above*) because the distance between the figure and reference object is irrelevant and not specified by the spatial term. However, there is evidence that upon reading a spatial description the distance parameter is set, even when the distance is not explicitly conveyed by the spatial term (Carlson & Van Deman, 2004). For example, if I ask someone where the television remote control is and he or she replies *It's in front of the TV* I limit my search with regards to the distance in front of the television. This is in spite of the fact that the spatial term *in front of* does not explicitly specify how I should set the distance parameter. People presumably assume that when a people describe the location of an object they are obeying the Gricean maxims of co-operation and are providing a description that will enable them to easily locate the object. However, it is not clear how people establish what is co-operative when describing the location of an object using projective descriptions that do not explicitly define distance between the figure and reference object. It is an important research question to identify how the distance parameter is set when comprehending a spatial description that does not explicitly specify the distance between the figure and reference object.

5.2. Experiment 8

Carlson and Van Deman (2004) showed that the distance parameter is set when using directional terms (e.g. *above*), even though distance is not explicitly stated in the spatial term. In their experiment participants had to judge whether a sentence was an accurate description of the spatial relationship between two letters. Unknown to participants, the trials were paired into prime and probe trials. The probe trials either had the letters the same distance apart as the prime or a different distance apart, and the spatial relationship could either be the same as the prime trial or different. The results showed that there was significant interference in verifying the sentence on the probe trials when the distance between the letters was different to that on the prime trials, indicating that participants had encoded the distance despite the fact that it was not articulated in the spatial terms. This was the case even when the relationship between the letters was not the same, meaning that the distance parameter was set independently of the direction parameter of the reference frame.

The distance parameter has also been shown to be influenced by the size of the objects in the spatial array. Carlson and Covey (2005) showed that participants imagine the distance between two objects in a spatial description to be larger if the objects in the array are larger. They argued that this is because smaller objects have to be closer together in order for them to interact. In support of this they also found that the distance estimates were smaller when the figure object was in front of the reference object compared to when it was behind. This is presumably because a figure object is more likely to be in the reference object's zone of interaction when it is in front of the object than when it is behind the object. According to Carlson and Covey if a spatial description is of two objects that are interacting (e.g. *the umbrella is above the man*) then the distance parameter will be set as shorter than if a spatial description is of two objects that are not interacting, but are of a similar size (e.g. *the coffee table was above the man*). However, Carlson and Covey (2005) did not investigate this directly, but instead used the finding that people imagine figure objects to be closer to reference objects when they are in front of the reference object compared to when they are behind the reference object. The present experiment investigated directly the influence of function and interaction on setting the distance parameter.

In this experiment I examined the influence of interaction on the distance parameter further by asking participants to estimate the distance between two objects when the figure and reference objects are functionally related (e.g. a basketball hoop and a basket ball) compared to when the objects are of equivalent geometrical properties, but non-functionally related (e.g. a basketball hoop and a globe) and then report the distance apart they imagined the two objects to be. If the interaction between objects affects the setting of the distance parameter then participants' estimations of the distance between the figure and reference objects will be significantly smaller when the figure and the reference object are functionally related, compared to when they are functionally unrelated. This is because objects that are functionally related are more likely to be interacting than objects that are functionally unrelated. Alternatively, the distance may be set solely by the size of the figure and reference objects and there will be no difference in the estimations of distance between the figure and reference objects when they are functionally related compared to when they are functionally unrelated.

This experiment also examined the relationship between the type of reference frame used to apprehend a spatial relationship, either intrinsic or relative, and the setting of the distance parameter. Carlson and Van Deman (2004) have shown that the distance between a figure and reference object is encoded even for directional terms, but they did not investigate whether the distance parameter was set differently for the different reference frames. This is possible as when two objects are functionally related and interacting, people are more likely to use an intrinsic reference frame to describe the spatial relationship between the two objects (Carlson 2000). In addition Carlson-Radvansky and Irwin (1993) showed that participants rate a description using an intrinsic reference frame as a more acceptable way of describing a spatial relationship when the distance between the figure and the reference object is smaller. This finding suggests that using the intrinsic reference frame to interpret a spatial description results in a smaller setting of the distance parameter. Therefore in this experiment participants were also instructed to imagine the relationship between the two objects using an intrinsic or a relative reference frame. If the reference frame used to apprehend a spatial relationship affects the setting of the distance parameter, it is expected that participants' distance estimates will be significantly different depending upon the reference frame used to apprehend the spatial relationship. If the distance parameter is set independently of the reference frame used, there will be no significant difference between participants' distance estimates, regardless of the reference frame used to imagine the spatial relationship between the two objects.

5.2.1. Participants

64 University of Notre Dame undergraduates participated in the experiment in return for course credit. All were native English speakers.

5.2.2. Materials

16 colour digital photographs of reference objects were used. For 8 of these the participants had to imagine the figure object below the reference object. For 8 of these the participants had to imagine the figure object above the reference object. In

addition, each reference object could be presented with one of two colour digital photographs of figure objects; hence there were 32 figure objects. For each reference object one of the figure objects was functionally related to it (e.g. for the *hammer* reference object a *nail* was the functional figure object), and the other figure object was of comparable size and shape, but not functionally related (e.g. for the *hammer* reference object a *toothpick* was the non functional reference object). A full list of figure and reference objects used is shown in Appendix F. On each trial of the experiment the reference object was presented in the vertical centre of the screen on the left hand side, and the figure object was presented in the vertical centre of the screen on the right hand side. Above each of the pictures there was a sentence describing the picture. A scene from a trial of the experiment is shown in Figure 39.

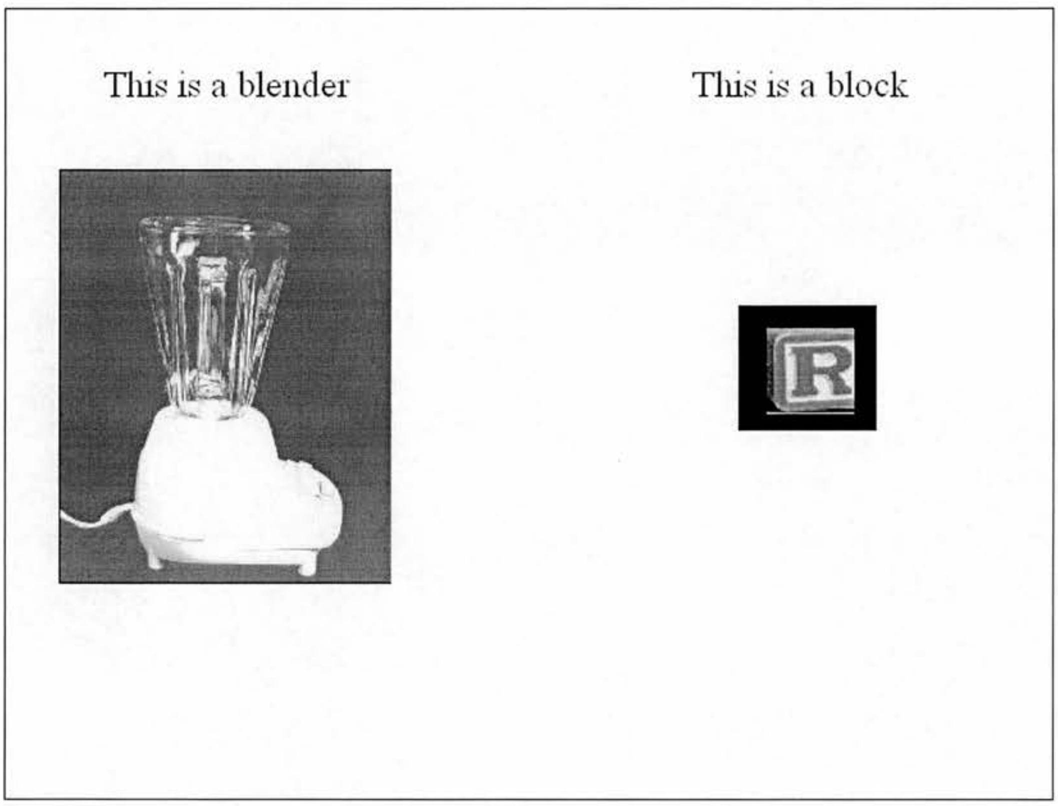


Figure 39. A display from Experiment 8.

5.2.3. Design

The experiment used a 2x2x2x2 within-participants design, with spatial relationship (above, below), reference frame (relative, intrinsic), reference-frame dissociation (conflated, dissociated) and figure-object relation (functional, non-functional) as the factors and levels. Spatial relationship referred to whether or not the participants were asked to imagine the figure object above or below the reference object (this was a between-items condition as different items were used for above and below conditions). The reference frame condition referred to whether participants were asked to imagine the spatial relationship between the figure and reference objects using an intrinsic or a relative reference frame. The reference-frame dissociation condition referred to whether or not the reference object was presented in its canonical upright position as in Figure 40, or rotated 90° clockwise as shown in Figure 41. In the canonical condition, the intrinsic and relative top-bottom axes were conflated; when the reference object was rotated, the reference frames were dissociated. Finally, the figure-object relation condition refers to whether or not the figure object was functionally related or functionally unrelated to the reference object. Eight lists were created across which each of the eight 'above' reference objects and each of eight 'below' reference objects appeared in each of the eight conditions once.

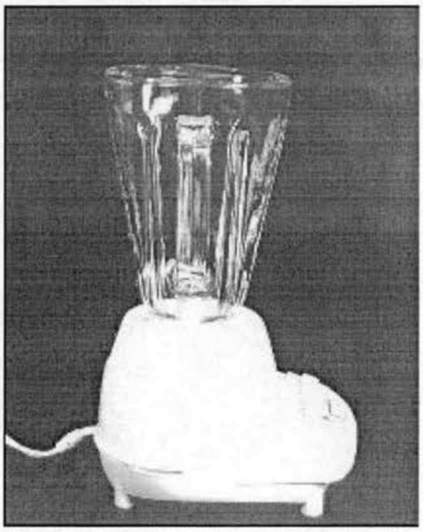


Figure 40: The presentation of the reference object in the conflated condition. The intrinsic and relative top-bottom axes are coincident.

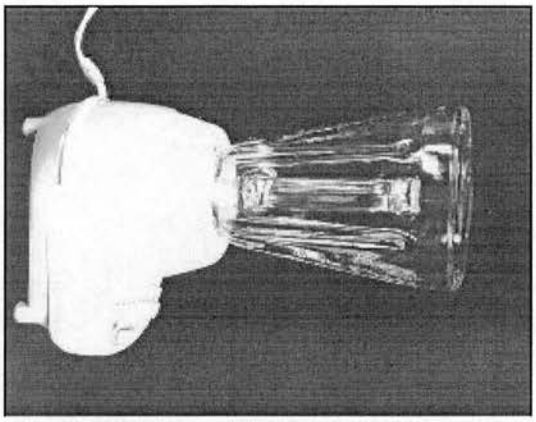


Figure 41: The presentation of the reference object in the dissociated condition. The intrinsic and the relative reference frames are dissociated.

Therefore each participant saw an example of each condition twice, once for the spatial relation above and once for the spatial relation below. In each list each reference object was presented only once. The eight conditions were presented in a random order.

5.2.4. Procedure

Participants were seated at a computer and read the experiment instructions presented on the screen. The instructions defined the difference between the intrinsic and the relative reference frames and gave examples. Participants were then told that if they did not understand what intrinsic and relative spatial relations meant then they should ask the experimenter before continuing with the experiment. All participants were clear on the difference between the two reference frames before commencing with the experiment.

Participants were told that they would be required to imagine the relationship between two objects. Their task was to type on the keyboard how far apart these two objects were in their imagination and to provide the unit of measurement.

After reading the instructions the participants began the experiment by pressing the space bar. On each trial the participant was presented with a reference object on the left hand side of the screen either in its canonical position or rotated 90° clockwise. On the right hand side of the screen a figure object was presented. Each of these had a sentence above them of the form *This is a* X. These were presented to identify the objects that the participants should be imagining. When the participant had identified both objects, they pressed the space bar and the objects were replaced with a sentence in the centre of the screen indicating the spatial relationship the participant should imagine the objects in and the reference frame they should use (e.g. *The ice cube above the blender (relative)*). When the participant had imagined the objects in the spatial relation they pressed the space bar and the sentence disappeared. Participants were then prompted to type, using the keyboard, the distance apart they imagined the two objects to be and then press the *enter* key. Participants were then prompted to enter the units (inches, feet, yards) they had used by pressing either the 'i' 'f' or 'y' key and then to press the *enter* key. After pressing the *enter* key participants were presented with the next trial. Each trial proceeded like this until the experiment was over.

5.2.5. Results

The responses that were entered in units other than inches were converted to inches and responses that were 3SD above the mean were truncated to the score 3SD above the mean. The data was then log-transformed because distance is best described by a power function (Carlson & Covey, 2005, Radvansky, Carlson-Radvansky & Irwin, 1993). Figure 42 shows the mean distances (log-transformed) for each of the eight conditions for below and above relationships.

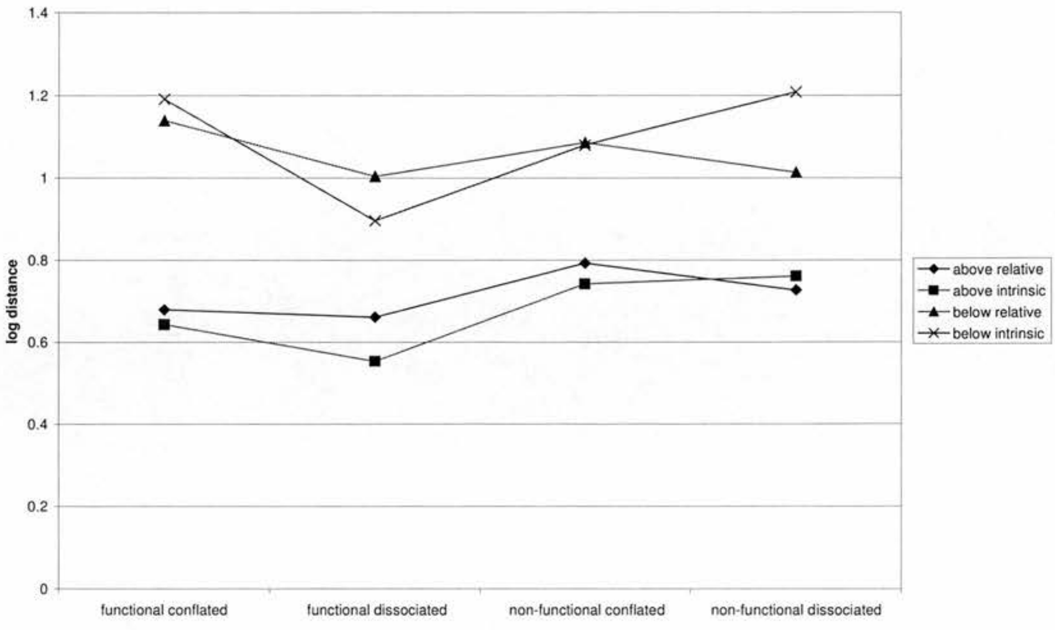


Figure 42: The mean log-transformed distances for each of the conditions in Experiment 8.

The results were analysed using a 2x2x2x2 within participants ANOVA (F1) and a mixed 2x2x2x2 ANOVA for items (F2) with spatial relation (above vs. below) as a within-participants, but between-items factor and reference frame (relative vs. intrinsic), dissociation (conflated vs. dissociated) and function (functional vs. non functional) as within-participants and -items factors.

The main effect of function approached significance by participants and was significant by items ($F1(1,63) = 3.29$; $p = .07$; $F2(1,14) = 8.76$; $p < .05$). If this effect is valid then participants estimated the distance between the figure and reference objects to be smaller when the two were functionally related than when they were not functionally related. The main effect of reference frame was not-significant ($F1(1,63) = 0.01$; $p > .05$; $F2(1,14) = 0.03$; $p > .05$); there was no difference between the distance estimate scores when participants were told to use an intrinsic reference frame and when they were told to use a relative reference frame. There was a main effect of spatial relation ($F1(1,63) = 93.3$; $p < .01$; $F2(1,14) = 114.4$; $p < .01$); participants estimated a greater distance between the reference and figure object when asked to imagine the figure object below the reference object than when asked to imagine the figure object above the reference object. As this was a between-items factor, this is probably due to the different sizes of the objects used in the different conditions. There was no main effect of dissociation by participants, but there was a significant difference by items ($F1(1,63) = 1.97$; $p > .05$; $F2(1,14) = 4.97$; $p < .05$). There was no significant interaction between function and dissociation for participants, but this was significant by-items ($F1(1,63) = 2.627$; $p > .05$; $F2(1,14) = 6.06$; $p < .05$). Figure 43 shows the distance estimations of participants for functionally related and non-functionally related items when the reference frames were conflated and dissociated. The means suggest that function only has an effect on setting the distance parameter when the reference frames were dissociated. The three-way interaction of reference frame, function and dissociation approached significance by participants, but not by items ($F1(1,63) = 3.57$; $p = .06$; $F2(1,14) = 1.01$; $p > .05$), the means for this analysis are shown in Figure 35. All other interactions were non significant by participants and items (all F1s < 3.1; F2s < 2.7).

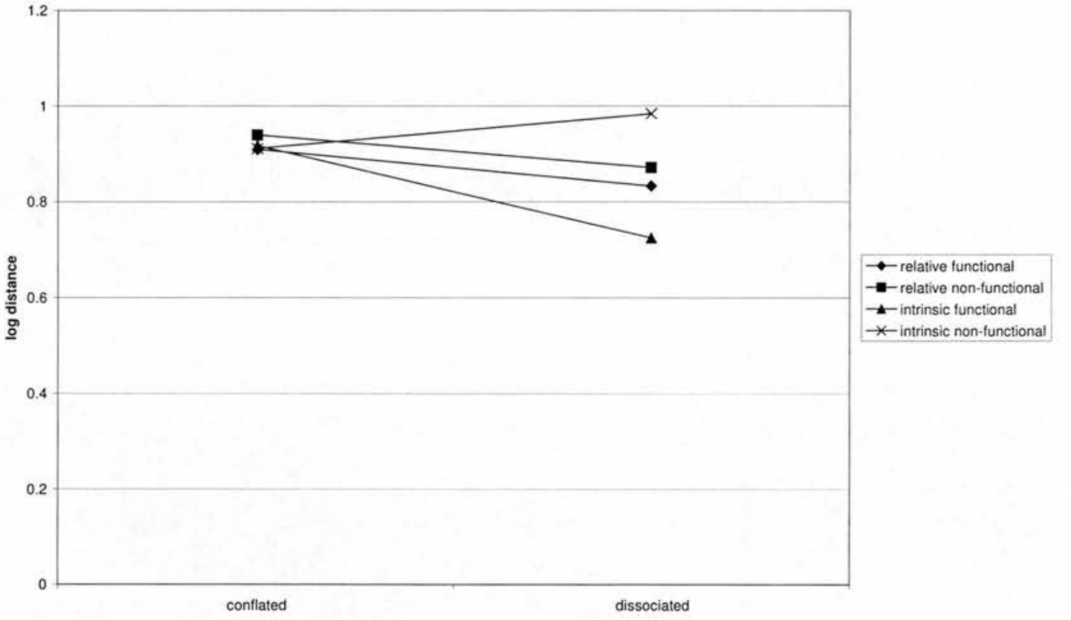


Figure 43. The mean distance estimation when the objects were functionally related and unrelated and the reference frames conflated and dissociated.

5.3. Experiment 8 Discussion

Experiment 8 set out to test whether or not functionally related figure objects and reference objects had influence on setting the distance parameter. It also set out to identify whether or not the reference frame used had an influence on setting the distance parameter.

5.3.1. Functional effects on distance

The main effect of function was close to significance by participants and was significant by items. Participants estimated the imagined distance between the figure and reference objects to be closer when they were functionally related than when they were not functionally related. This effect cannot be due to geometric differences between the functional figure objects and the non-functional figure objects because they were matched for size and shape. This finding supports the claims of Carlson and Covey (2005), that the distance parameter is set partly based upon the interaction between the figure and the reference object. In their study they found that participants estimated distances between figure and reference objects to be closer when the figure object was in front of the reference object compared to when it was behind, and argued that this was because there was more likely to be an interaction between the two objects in the former situation. The findings of this experiment support their argument as the functional relationship between the figure and reference object implies an interaction between the two objects and thus leads to setting the distance parameter as shorter than when the figure and reference object were not functionally related, and therefore unlikely to interact. It would be interesting to see if the functional effect on distance was due to the functional association between the figure and reference object. That is, whether the effect arose because the two objects often co-occur (e.g. a basketball and a basketball hoop) or whether the effect was due to the perceived interaction between the two objects. If the latter is the case then the functional effect should persist when the figure object has affordances (Barsalou, Sloman, & Chaigneau, 2005) that enable it to fulfil a function normally associated with a reference object (e.g. using a newspaper as an umbrella ;Coventry, Prat-Sala & Richards, 2001), but not normally associated with the figure object.

5.3.2. Reference frame effects on distance

One of the goals of the experiment was to identify whether or not using different reference frames led to setting the distance parameter in different ways. There was, however, no effect of the reference-frame condition on participants' imagined distances between the figure and reference objects. This suggests that setting the distance parameter is independent of the reference frame used to apprehend the spatial relation. However, it is possible that participants were not truly interpreting the spatial relations using an intrinsic and relative reference frame. Although participants understood the difference between the two frames before commencing the experiment, it is impossible to tell whether or not they were actually imagining the relationships in the correct manner. One indication that they may not have been imagining the objects in the correct way is that there was no three-way interaction between reference frame, function and dissociation, although it did approach significance. If participants were imagining the objects using the correct reference frames then when the reference object was in a non-canonical rotation and the reference frames were dissociated there should only be functional effects when the reference frame the participants were using was intrinsic (the means distance estimates were in the predicted direction however, see Figure 42). This is because if the participant imagined the figure object and reference object using a relative reference frame when the object was in a non-canonical orientation then there can be no interaction between the two objects because the functional part of the reference object is not aligned with the figure object and therefore whether the object is functionally related to the object or not should be irrelevant (although it is possible the functional effect is due to the functional association between the two objects as mentioned above).

5.3.3. Effects of the spatial relation

The results showed an effect of spatial relation on setting the distance parameter. Participants estimated the distance between the figure and the reference object as smaller when asked to imagine the figure object above the reference object than when asked to imagine the figure object below the reference object. This is probably due to the different items used for the different spatial relation conditions. In particular, the crane, fishing rod and basketball hoop items. These three items produced larger distance estimates than the items used in the *above* relation condition. Carlson and Covey (2005) showed that when the figure and reference object are larger the distance parameter is set larger: The larger reference objects used in the *below* condition of this experiment may therefore be the cause of the spatial relation effect. If this is the explanation then the effect does confirm the finding that the size of the reference object affects the setting of the distance parameter (Carlson & Covey, 2005). It is, however, possible that the distance parameter is set smaller for *above* relations than for *below* relations, in an analogous fashion to *in front* of and *behind* (Carlson & Covey, 2005), but to test this hypothesis directly would require using a within-items design, but is an interesting issue for future research.

5.3.4. Conclusion

Because of the problem of not knowing whether participants used the reference frames as intended in this experiment, the conclusion that the distance parameter is not affected by choice of reference frame cannot be wholly accepted. However, the results do suggest that the functional relationship between the figure and the reference object does affect the setting of the distance parameter. Future research should focus upon whether the functional factor that affects the distance parameter is the association between the two objects or the interaction between the two objects.

5.4. Experiment 9

Experiment 9 investigated the apprehension of spatial relations when the figure and reference object are different distances apart. It also investigated the use of attention in the setting of the distance parameter. This experiment investigated whether or not distance does affect the time taken to apprehend the spatial relationship between two letters. Research on apprehending spatial relations is inconclusive regarding effects of distance on the time taken to apprehend a spatial relation. Logan and Compton (1996) used a sentence picture verification, task and found that the time taken to verify a spatial relation between two letters was not affected by the distance between the two letters (see also Carlson & Logan, 2001). However, Carlson and Van Deman (2004) did find an effect of distance on the time taken to apprehend a spatial relation between two letters. In their experiments, participants were faster to verify a spatial relation when the distance between the two letters was shorter. They argued that the reason they found an effect of distance and Logan (1994) did not, was that the difference between Logan's short and long distance displays was not enough to require the distance parameter to be set differently.

Logan (1994) showed that apprehending a spatial relation requires the use of attention. When participants were asked to look for a '+' *above* a '–' the time to find the target increased with the number of distracter items presented with the target (i.e. the more '–' *above* '+' distracters there were). The present experiment used the same visual search paradigm to examine whether or not setting the distance parameter requires the use of attention. In this experiment the participants were asked to decide whether a target (e.g. an S *above a* T) was present in a display containing a number of distracters. The distance between the two letters could be one of two distances: *short* or *long*; and the distracters were the same letters presented at different distances. If setting the distance parameter requires the use of attention, then the time to identify whether or not the target is present will increase as the number of distracters, with the letters at different distances apart, in the display increases. If, however, distance can be apprehended pre-attentively then the number of distracters present in the display will have no influence on the time taken to identify the presence of the target.

5.4.1. Participants

20 University of Notre Dame psychology undergraduates participated in return for course credit.

5.4.2. Materials

Each display consisted of 1, 2, 4 or 8 pairs of letters, one above the other. The letters used were an upper case S and an upper case T written in 18pt Times New Roman font. There was either a *short* distance between the two letters (250mm) or a *long* distance between the two letters (500mm). Each pair of letters was surrounded by a box that was 1 pixel from the top, bottom and sides of the edges of the letters. The pairs of letters were each presented in one of eight positions on the screen, so that the central point between the two letters was an equal distance from the central point on the screen. Figure 44 shows one of the experimental trial displays, containing eight pairs of letters. When the display contained less than eight pairs of letters the distracter items were positioned randomly. The target appeared in all eight positions for each condition, across eight trials.

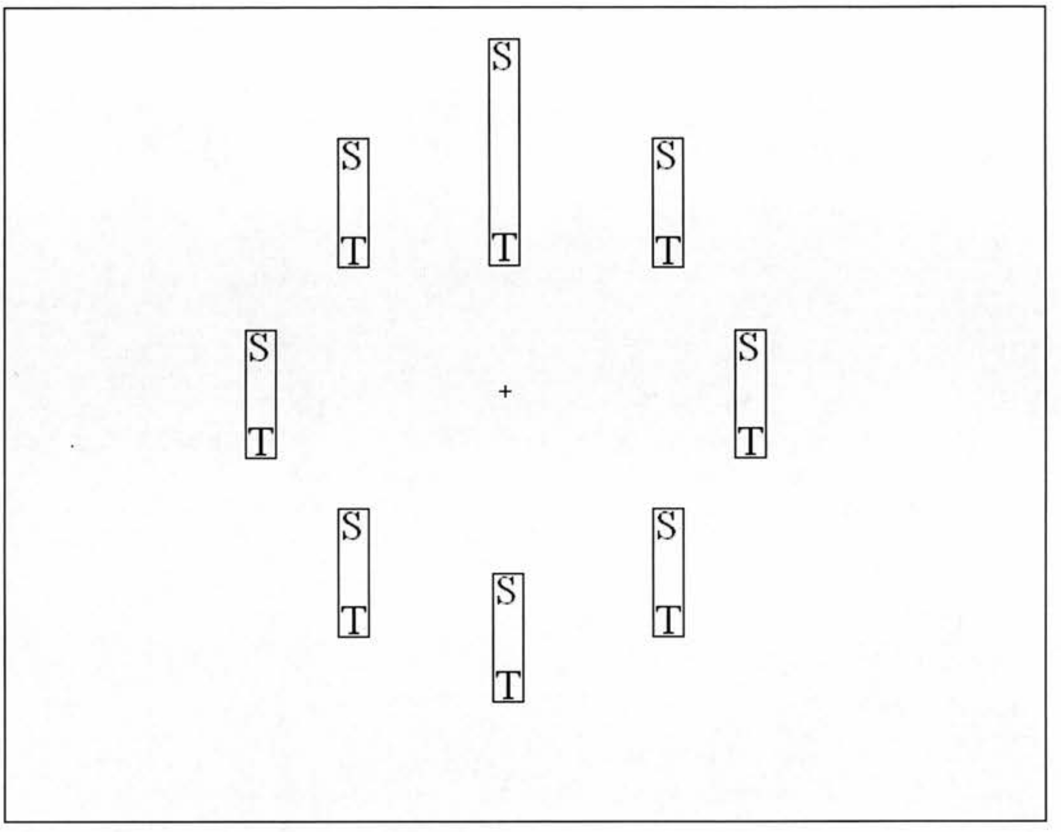


Figure 44: A display from Experiment 9 with 8 pairs of letters: the top most one is the target, given the prompt *S above T long*.

5.4.3. Design

The experiment used a 2x2x3x4 within-participants design, with the factors target-presence (present vs. absent), target-distance (short vs. long), distracter-type (distance, relation or both) and display-size (1, 2, 4, or 8). The Target-presence condition referred to whether or not the target was present on the scene or not. The target-distance condition referred to whether or not the target letters were the *short* distance apart or the *long* distance apart. The distracter-type condition referred to whether or not the distracters differed from the target by the *distance* between the letters, the *spatial relationship* between the letters or *both* of these factors. Table 11 shows the targets and the different types of distracter for each condition. Distracters for a trial were always homogenous in their type.

CONDITION	TARGET TYPE	DISTRACTER TYPE
Short target Distance distracters	S T	S T
Short target Relation distracters	S T	T S
Short target Both distracters	S T	T S
Long target Distance distracter	S T	S T
Long target Relation distracter	S T	T S
Long target Both distracter	S T	T S

Table 11: The target and distracter types for each of the conditions in Experiment 9.

The display size referred to the number of pairs of letters in the display. For example, if the target was present and the display size was 1, there was 1 pair of letters on the screen, the target. If the display size was 1 and the target was absent, there was one pair of letters on the screen, a distracter. The total combination of factors yielded 48 conditions.

Two lists were created. In each list, half of the targets had the S above the T and half of the targets had the T above the S. This was counterbalanced across lists. Each experiment contained 384 trials, made up of each of the 48 trials repeated 8 times with the target in each of the eight positions on the screen. All trials were experimental and were presented randomly in a different order for each participant.

5.4.4. Procedure

Participants were instructed that they would be asked to identify whether or not an S *above* a T or T *above* an S were present on the screen, and that the letters could be different distances apart. Participants were then shown an example of what each type of target would look like. Before the experiment began participants performed 14 practice trials, two of each display size and three of each distracter type, half of the 14 had the target present and half did not. For the practice trial feedback was given, but not in the main experiment. When the participants had completed the practice they began the main experiment. Each trial proceeded as follows: Participants were presented with an expression that described the type of target the participant was to look for on the following display, (e.g. *S above T short?*) The first part described the figure and reference object and relation of the target, and the second part described the distance between the figure and reference object. The sentence remained on the screen until the participant pressed the space bar. It was then replaced with the experimental display. The participant then had to decide as fast and as accurately as possible whether or not the target was present in the display. If they thought it was present they pressed the '/', if they thought it was not present they pressed the 'Z' key (this reversed was for half the participants). No feedback was given. The experiment then proceeded to the next trial.

5.4.5. Results

Analyses were carried out on both the response latency and response accuracy.

Response Latency: The analyses were performed only on the response latencies of correct trials. The mean percentage of correct responses for participants was 93.2% (min 80.5% - max 99%). Table 12 shows the mean response latencies for each of the conditions in Experiment 9. Response latencies were analysed using a 2x2x3x4 within-participants ANOVA, with target presence (present vs. absent), target distance (short vs. long), distracter type (distance, relation, both) and display size (1, 2, 4, 8) as the factors.

	Target Distance	Short				Long			
	Display size	1	2	4	8	1	2	4	8
Target Presence	Distracter Type								
Present	Distance	849	1013	1129	1265	825	968	1009	1113
	Relation	880	986	1283	1336	849	1031	1242	1344
	Both	867	1026	1137	1311	834	938	1085	1101
Absent	Distance	908	929	1084	1230	844	903	1015	1189
	Relation	967	976	1199	1496	903	956	1233	1548
	Both	833	965	1045	1107	809	872	1025	1065

Table12: The mean response latencies in milliseconds for each of the conditions in Experiment 9.

There was a main effect of target presence ($F(1,19) = 4.4$; $p < .05$); participants were faster to respond when the display contained the target than when it did not (1060ms vs. 1096ms). There was a main effect of display size ($F(3,57) = 220.5$; $p < .01$). Figure 45 shows the mean response latency, collapsed across the distracter types and distances, for the different display sizes when the target is present and absent. As the display size increased participants took longer to respond.

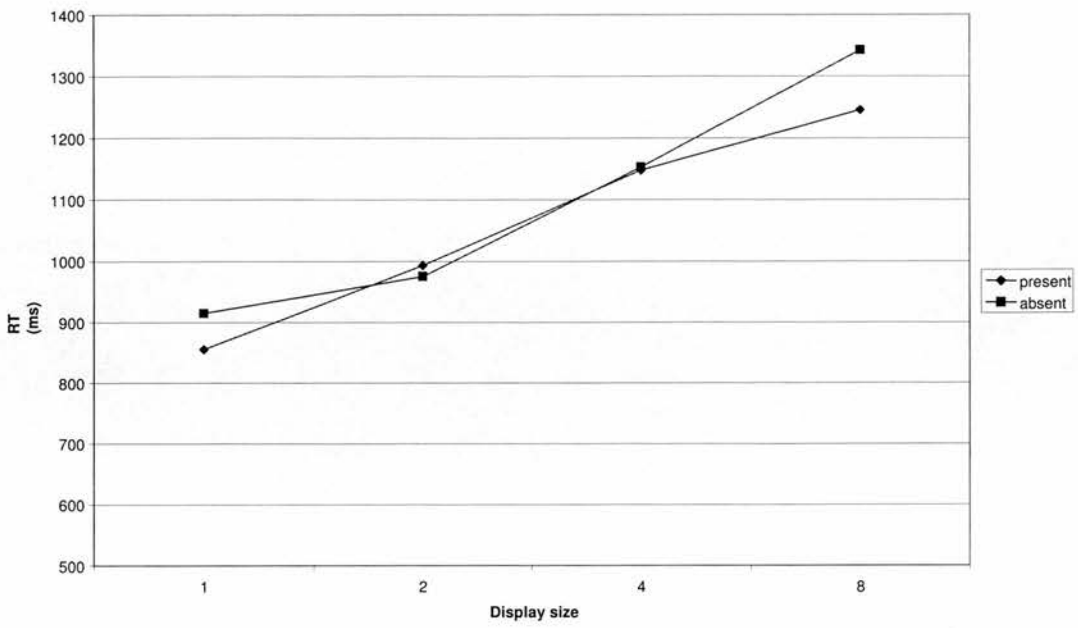


Figure 45: The mean response latency for display sizes when the target was present and absent

There was a significant main effect of target distance ($F(1,19) = 14.49$; $p < .01$). Participants were faster to respond when the distance between the distracters was long compared to when it was short (1055ms vs. 1101ms). There was also a main effect of distracter type ($F(2,38) = 66.03$; $p < .01$). Participants were significantly faster to respond when the distracters differed from the target in the spatial relationship (1168ms) than when the distracters differed from the target in the distance (1042ms) between the two letters ($t(19) = 7.5$; $p < .01$), or when both the relationship and the distance were different (1024ms) ($t(19) = 9.89$; $p < .01$). There was no difference between the response latency when the distracters differed in terms of the distance and when the distracters differed in terms of the both the distance and the relationship ($t(19) = 2.28$; $p > .05$).

There was also a significant two-way interaction between display size and distracter type ($F(6,114) = 22.79$; $p < .01$). Figure 46 shows the mean response latencies for each of the distracter types at each of the display sizes when the distracter was present and absent. The interaction between target presence and distracter type was significant ($F(2,38) = 12.23$; $p < .01$).

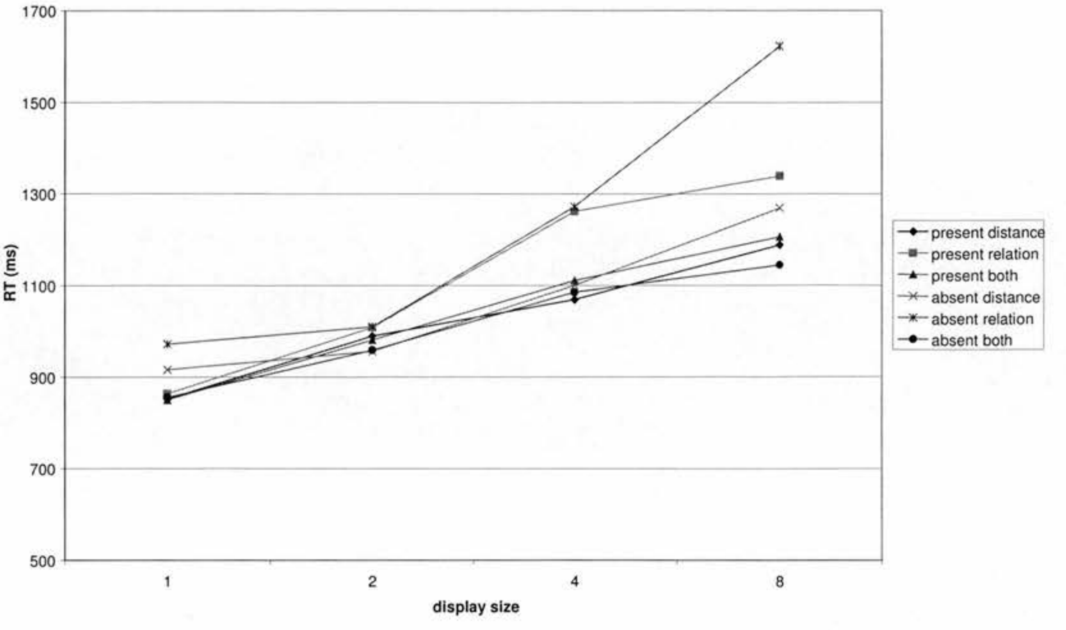


Figure 46: The mean response latencies for each of the distracter types with different display sizes.

There was a significant interaction between distance and distracter type ($F(2,38) = 5.16$; $p < .05$). Participants were faster to respond to long distance targets than short distance targets when the distracters differed in distance (difference = 64ms) and differed in distance and relation (92ms), but not when the distracters differed in relation alone (3ms).There was a significant interaction between target presence and display size ($F(3,57) = 6.15$; $p < .01$).

There was a significant three-way interaction between target presence, distracter type and display size ($F(6,114) = 6.48$; $p < .01$); the mean response latencies are shown in Figure 46. All other interactions were non-significant (all Fs < 2.3).

Response accuracy: There were on average very few errors per participant (7%). The mean numbers of errors per condition are shown in Table 13. The mean numbers of errors were analysed using a 2x2x3x4 within-participants ANOVA which had the same factors and levels as used in the response latency analysis.

There was a significant main effect of presence ($F(1,19) = 10.2$; $p < .01$). Participants made significantly more errors when the target was present than when the target was absent (0.66 vs. 0.44). There was a significant main effect of target distance (F(1,19)

= 9.8; p < .01). Participants made more errors when the distance between the letters was short than when the distance between the letters was long (0.48 vs. 0.62). There was a significant main effect of distracter type (F(2,38) = 7.7; p < .01). Post-hoc tests showed that there was no significant difference between the number of errors made when the distracters differed in terms of distance from the target and when distracters differed in terms of relation compared to the target (0.58 vs. 0.61) (t(19) = 0.8; p > .05). There were significantly fewer errors made when the distracters differed from the target in terms of both distance and relation, than when the distracters differed by distance alone (0.58 vs. 0.46) (t(19) = 3.1; p < .05), and by relation alone (0.61 vs. 0.46) (t(19) = 3.7; p < .01).

	Target Distance	Short				Long			
	Display size	1	2	4	8	1	2	4	8
Target Presence	Distracter Type								
Present	Distance	0.5	0.5	0.9	1.55	0.3	0.45	0.4	0.55
	Relation	0.3	0.65	0.95	1.15	0.25	0.6	0.85	0.9
	Both	0.35	0.55	0.65	1.4	0.6	0.4	0.55	0.45
Absent	Distance	0.6	0.4	0.35	0.65	0.5	0.35	0.7	0.55
	Relation	0.75	0.45	0.35	0.55	0.45	0.3	0.6	0.7
	Both	0.45	0.3	0.4	0.25	0.1	0.2	0.25	0.4

Table13: The mean number of errors (out of 8) per condition in Experiment 9.

There was a main effect of display size (F(3,57) = 6.6; p < .01). The mean number of errors for each of the distracter types and each of the display sizes when the targets are present and absent are shown in Figure 40. Post-hoc tests show that there was a significant difference between display sizes of 1 and 8 (0.43 vs. 0.76) (t(19) = 4.1; p < .01), and 2 and 8 (0.43 vs. 0.76) (t(19) = 3.2; p < .05). There were no significant differences in the number of errors made between any other display sizes.

There was a significant interaction between target presence and target distance (F(1,19) = 5.8; p < .05). There was no effect of distance on the number of errors made when the target was absent (0.42 vs. 0.46). When the target was present, however, participants made more errors when the target distance was short, than when the target distance was long (0.79 vs. 0.53.). There was a significant interaction between target presence and display size ($F(3,57) = 6.7$; $p < .01$). When the target was absent, the display size had a lesser effect on the number of errors made, than when the target was present (see Figure 47). There was a significant interaction between distance and size ($F(3,57) = 3.4$; $p < .05$). The mean numbers of errors made for the long and short target distances at each display size are shown in figure 48.

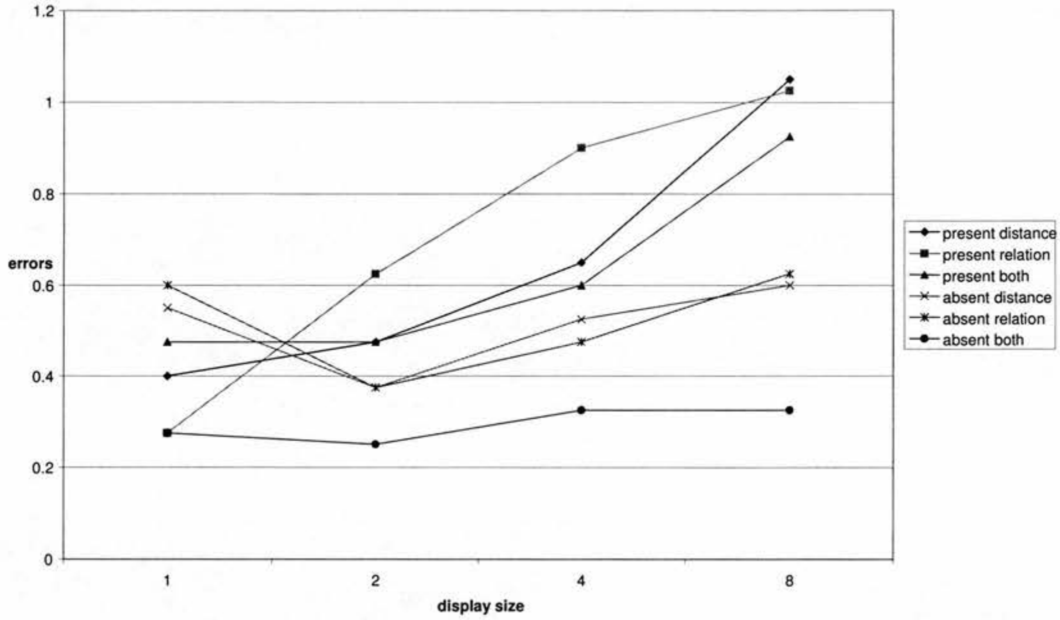


Figure 47: The mean number of errors for each of the distracter types with different display sizes.

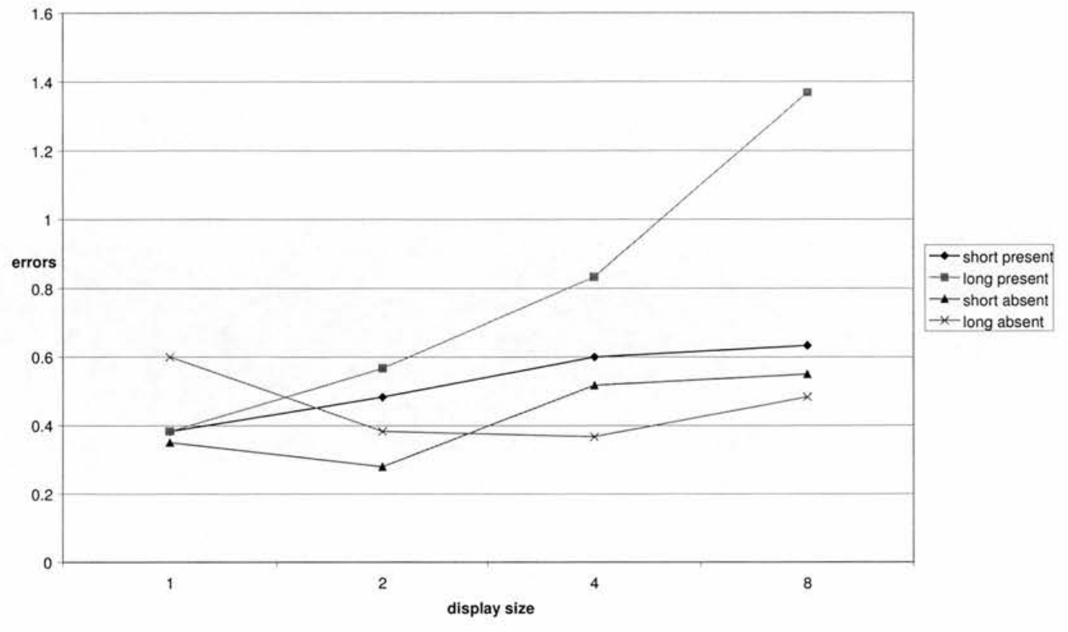


Figure 48: The mean number of errors for long and short distances at each of the display sizes when the targets are present and absent.

There was a significant three-way interaction between target presence, target distance and display size ($F(3,57) = 5.97$; $p < .01$). The means for these conditions are shown in Figure 48. All other interactions were non-significant (all Fs < 1.7).

5.4.6. Discussion

The results were largely akin to the pattern of results reported by other investigations into spatial relationships using the visual search paradigm (Logan, 1994). Overall, participants were faster to respond when the target was present, than when the target was absent. Presumably because locating the target allowed an end to searching. There was also an effect of target presence on the error rate; there were more errors made when the target was present, than when the target was absent. Participants were more likely to end their search for the target before they had actually found the target, than they were to report a target as present when it was not; there were more false negatives than false positives.

There was also an effect of display size on the response time and on the error rate. As the display size increased the response time increased and the error rate increased. This indicates that searching for the targets required the use of attention. If the targets could be found without the use of attention, then the display size would have no effect on the time to find the targets (Logan, 1994, Treisman & Gelade, 1980, Triesman & Gormican, 1988).

This experiment set out to investigate whether or not searching for a target spatial relationship based upon the distance between the figure and reference object required the use of attention. If attention was required to set the distance parameter then there should be an increase in the response latency to detect the target, as the number of distracters that differ from the target, based on distance between the letters, increases. There was an effect of display size on the time to find targets amongst all distracter types, although the effect was larger when the distracters differed from the target in relation alone. This means that setting the distance parameter in order to search for a figure object at a specific distance from the reference object, requires the use of attention. However, the results showed that participants took longer to respond when the distracters differed from the target in terms of the relation between the two letters, than when the distracters differed from the target in terms of distance or in both relation and distance. This finding suggests two things: Firstly, the task demand of setting the distance parameter as a search criterion is easier than setting both the orientation and direction parameters, which are required to apprehend a relation between two objects. This is intuitive, as to use a relation between two objects as a search criterion requires setting two parameters, both orientation and distance, but using distance requires setting a single parameter. If the experiment had asked participants to identify whether or not the target was an S *above or below* a T and the distracters were S's left or right of T's, the search criterion could be established by setting the orientation parameter alone, because only the axis on which the figure object was located would need to be identified. Such a task should yield an effect of display size comparable to that shown for the distance distracter types in this experiment, because searching based upon orientation or distance each require the setting of a single parameter.

Second, the fact that the results showed no difference in response time when the distracters differed from the target in terms of distance and in terms of both distance and relation means that participants were able to use a strategy that when the distracters differed in distance and relation, they used only the distance parameter as a search criterion and ignored the relation. If one of the conditions of the experiment had had half of the distracters differing from the target in terms of distance only, and half differing in terms of relation only, then participants would have to set the distance, orientation and direction parameters resulting in an even greater response time. In short, the results suggest the hypothesis that, the more parameters that a person has to set in order to apprehend a spatial relationship, the greater the influence of distracters on response latency and error rate. This remains an issue for future research and may show how the different parameters of a reference frame are set in relation to each other. If this hypothesis is correct then it would explain why there was a main effect of the type of distracter on response time, with those differing from the target in relation being slower than the other two distracter types, because the former requires setting two parameters (orientation and direction), whilst the latter two require setting only the distance parameter. Interestingly, for the accuracy data there were fewer errors when the distracters differed from the target in both relationship and distance. This suggests that although participants used a general strategy of using distance to search for the target, they were still sensitive to the relationship between the letters when the distracters were different from the target in terms of both distance and relationship.

There was an effect of distance, with long-distance targets being responded to faster than short-distance targets, although there were more errors for the long-distance targets than short-distance targets; this suggests a potential speed-accuracy trade-off. The reaction time data is at odds with previous research on the distance parameter. Logan and Compton (1996) found no difference in response time to apprehend spatial relations at differing distances (see also Carlson & Logan, 2001), and Carlson and Van Deman (2004) found that participants were faster to respond when the figure and reference object were a shorter distance apart. The likely explanation for this difference is that in this experiment the long-distance targets were more salient than the shorter-distance targets, because they were larger than the distracters when the distracters differed in distance. When the targets were *short* and the distracters were *long*, the target was possibly not as noticeable because it was smaller, leading to increased response times. This explanation is supported by the interaction between target distance and distracter type, which shows that the effect of distance disappears when the distracters are the same distance as the target (i.e. for distracters that differed only in relation). Therefore the distance effect can probably be attributed to salience rather than a 'true' effect of setting the distance parameter, suggesting that there is no effect of distance on time to apprehend a spatial relation. However, this interpretation must be taken with caution as it is possible that the difference between the long and the short distances was insufficient to elicit a distance effect (Carlson & Van Deman 2004).

The results show that setting the distance parameter in order to apprehend a spatial relationship requires the use of attention. They also suggest that the more parameters of the reference frame that need to be set in order to apprehend a spatial relation, the greater the difficulty of apprehending the relation in the presence of distracters. Although distance does require the use of attention, it appears that setting the distance parameter requires the same amount of cognitive effort regardless of the distance between the figure and reference object. This is demonstrated as the time to apprehend a target in this experiment was the same regardless of the distance between the letters of the target.

5.5. Experiment 10

Experiment 9 showed that setting the distance parameter requires the modulation of attention. Experiment 10 used this finding to investigate whether or not the distance parameter is set independently of the reference frame that is used. This experiment again used a visual search paradigm where the participants were asked to search for a target (as in Experiment 9, two letters) in a specified spatial relation. However, in this experiment the target could be in a spatial relationship using an intrinsic reference frame or a relative reference frame. This therefore investigated two issues: First, whether or not apprehending a spatial relationship in one reference frame is faster than another reference frame. If it is easier to apprehend a spatial relationship using one reference frame than another, then participants will be faster to verify the presence or absence of the target using one reference frame, compared to using another reference frame. Alternatively, the reference frame used to apprehend the spatial relationship may have no effect on the response latency. Second, the experiment examined whether or not the distance parameter is set independently of the reference frame used to apprehend the spatial relationship. If a specific setting of the distance parameter (e.g. shorter) is associated with a specific reference frame then participants will be faster to apprehend a spatial relationship when the reference frame is congruent with the setting of the distance parameter, than when the distance parameter and reference frame are incongruent. For example, if the intrinsic reference frame is associated with a shorter setting of the distance parameter, then people should be faster to apprehend a spatial relationship using an intrinsic reference frame if the distance between the figure and reference object is short. Alternatively, if distance is set independently of the reference frame, then the setting of the distance parameter will have no influence on the time taken for participants to find the target using a given reference frame.

5.5.1. Participants

10 students from the University of Edinburgh participated in the experiment.

5.5.2. Materials

The targets and distracters and experimental displays were formed in the same way as in Experiment 9 with the following differences. The letters used were now a T and an O, and the T was always rotated 90° clockwise so as to dissociate the intrinsic and the relative reference frame. In addition there was no box surrounding the pairs of letters. There were 8 target types in the experiment. The O could be above the T or below the T, this could either be according to an intrinsic or a relative reference frame, and the distance between the O and the T could either be long or short.

5.5.3. Design

The experiment used a 2x2x2x3x3 within-participants design, with the factors, target-presence (present vs. absent), target-reference frame (relative vs. intrinsic), target-distance (short vs. long), distracter-type (distance, reference frame, both) and display-size (1,4,or 8), making a total of 72 conditions. There were a total of 288 trials (all experimental). Half of the trials had the O *above* the T and half had the O *below* the T; this was counterbalanced across two lists and across conditions in each list. The 8 positions for the target were counterbalanced across conditions and lists so that there were 4 trials per condition with the target appearing in 4 of the 8 positions per condition on list 1 and the other 4 on list 2. The rest of the design was the same as Experiment 9. Table 14 shows the target types and the distracter types for the intrinsic reference frame in each condition in the experiment. Table 15 shows the target types and the distracter types for the relative reference frame in each condition in the experiment (only the O above T relation is shown in each table).

CONDITION	TARGET TYPE	DISTRACTER TYPE
Short target Distance distracters intrinsic	⊣ O	⊣ O
Short target Reference frame distracters intrinsic	⊣ O	O ⊣
Short target Both distracters intrinsic	⊣ O	O ⊣
Long target Distance distracter intrinsic	⊣ O	⊣ O
Long target Reference frame distracter intrinsic	⊣ O	O ⊣
Long target Both distracter intrinsic	⊣ O	O ⊣

Table 14: The intrinsic reference frame targets and distracters for each condition.

CONDITION	TARGET TYPE	DISTRACTER TYPE
Short target relative Distance distracters	O ⊣	O ⊣
Short target relative Reference frame distracters	O ⊣	⊣ O
Short target relative Both distracters	O ⊣	⊣ O
Long target relative Distance distracter	O ⊣	O ⊣
Long target relative Reference frame distracter	O ⊣	⊣ O
Long target relative Both distracter	O ⊣	⊣ O

Table15: The relative reference frame targets and distracters for each condition.

5.5.4. Results

Analyses were carried out on both the response latency and response accuracy.

Response Latency: The analysis was performed only on response latencies for correct trials. The mean percentage of correct responses for participants was 91.2% (min 81.3% - max 98.3%). Table 16 shows the mean response latencies for each of the conditions in Experiment 10.

		Target Distance	Short			Long		
		Display size	1	4	8	1	4	8
Reference frame	Target Presence	Distracter Type						
Relative	Present	Distance	1328	2636	2659	1780	2260	2478
		Relation	1383	2072	2443	1233	2080	2330
		Both	1710	2188	2191	1804	1989	1955
	Absent	Distance	1464	1721	2041	1428	1741	2025
		Relation	2124	2059	2497	1717	2034	2581
		Both	1376	1843	2029	1179	1437	1686
Intrinsic	Present	Distance	1569	2153	2966	1578	2229	2305
		Relation	1691	2106	2745	1306	2376	2454
		Both	1636	1760	2306	1529	1839	2527
	Absent	Distance	1441	1815	2802	1202	1641	2041
		Relation	1946	1931	2225	1729	1900	2911
		both	1393	1623	2109	1320	1883	1974

Table 16: The mean response latency (ms) for each of the conditions in Experiment 10.

The response latencies were analysed using a 2x2x2x3x3 within-participants ANOVA, with target-presence (absent vs. present), target-reference frame (relative vs. intrinsic), target-distance (short vs. long), distracter-type (distance vs. frame vs. both) and display-size (1 vs. 4 vs. 8) as the factors and levels.

There was a main effect of distracter-type ($F(2,18) = 7.4$; $p < .01$). Post-hoc tests showed that participants were faster to respond when the distracters differed from the target in terms of reference frame and distance (1731ms) compared to when the distracters differed only in distance (1895ms) ($t(9) = 2.9$; $p = .05$), and compared to when the distracters differed only in terms of the reference frame (1983ms) ($t(9) = 4.4$; $p < .01$). There was no difference in response time when the distracters differed in terms of reference frame only and when distracters differed in terms of distance only ($t(9) = 1.5$; $p > .05$). There was also a main effect of display-size ($F(2,18) = 41.8$; $p < .01$). Participants were significantly faster to respond when the display size was 1(1449ms), compared to when the display size was 4 (1876ms) ($t(9) = 7.3$; $p < .01$), and compared to when the display size was 8 (2285ms) ($t(9) = 6.6$; $p < .01$). Participants were also significantly faster to respond when the display size was 4 compared to 8 ($t(9) = 5.3$; $p < .01$).

The main effect of target-presence was non-significant ($F(1,9) = 1.9$; $p > .05$). Participants were not significantly faster to respond when the target was present (1930ms) compared to when the target was absent (1809ms). There was no main effect of target-distance, although it did approach significance ($F(1,9) = 4.01$; $p = .076$). Participants were marginally faster to respond when the target distance was long (1830ms) compared to when the target distance was short (1909ms). Finally there was no main effect of reference-frame ($F(1,9) = 2.1$; $p > .05$). The reference frame used to apprehend the spatial relationship had no significant effect on the time taken to respond (1843ms vs. 1897ms).

There was a significant interaction between target-presence and distracter-type ($F(2,14) = 8.34$; $p < .01$). When the target was present the mean response latency was longer when the distracters differed from the target in terms of distance or both distance and reference frame compared to reference frame alone. When the target was absent the mean response latency was longer when the distracters differed from the target by reference frame alone. There was a significant interaction between target-presence and display-size ($F(2,18) = 3.5$; $p = .05$).

There was a marginal three-way effect of target presence, distracter type and display size ($F(4,36) = 2.5$; $p = .06$). The mean response latencies for each of the distracter types using each reference frame at each of the display sizes are shown in Figure 49. There was a significant three-way interaction between target distance, distracter type and display size ($F(4,28) = 4.95$; $p < .05$). All other interactions were non significant, all Fs < 2.5, except the three-way interaction between reference frame, target distance and distracter type ($F(2,18) = 2.8$; $p = .087$).

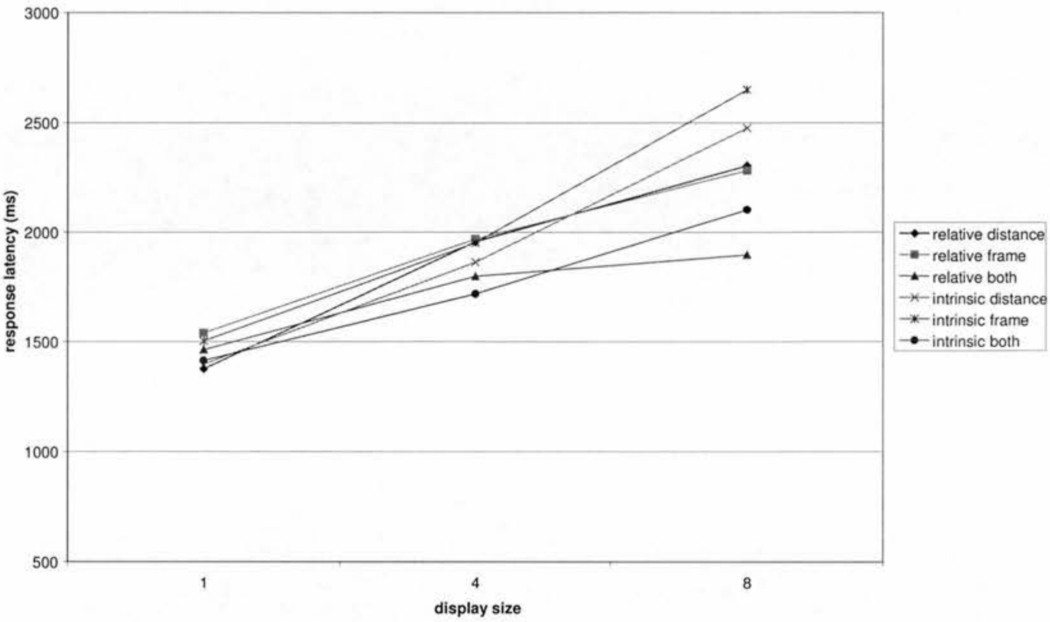


Figure 49: The mean response latency for different reference frames, with different distracter types at each of the display sizes.

Response accuracy: Each participant produced a mean of 9% errors. The mean number of errors for each condition are reported in Table 17. The mean number of errors was analysed using a 2x2x2x3x3 within-participants ANOVA which had the same factors and levels as used in the response latency analysis. There was a main effect of target-distance ($F(1,9) = 5.48$; $p < .05$). Participants made fewer errors when the distance between the letters was short compared to when it was long. There was also a main effect of distracter-type ($F(1,9) = 7.33$; $p < .01$). Post hoc tests showed that participants made fewer errors when the distracter differed from the target by both distance and reference frame, compared to when the distracters differed from the target by distance alone ($t(9) = 3.03$; $p < .05$), and when the distracters differed from the target by reference frame alone ($t(9) = 2.9$; $p < .05$). There was no difference in the number of errors made when the distracters differed from the target by distance alone and when the distracters differed from the target by reference frame alone.

Reference frame	Target Presence	Target Distance	Short			Long		
		Display size	1	4	8	1	4	8
		Distracter Type						
Relative	Present	Distance	0.4	0.4	0.4	0.3	0.2	0.9
		Relation	0	0.3	0.3	0.4	0.3	1
		Both	0.3	0.4	0.4	0.2	0.3	0.8
	Absent	Distance	0.5	0.2	0.3	0.3	0.1	0.2
		Relation	1.1	0.4	0.3	0.6	0.4	0.5
		Both	0.1	0.2	0.1	0.1	0.3	0.1
Intrinsic	Present	Distance	0.1	0.3	0.6	0.4	0.5	0.7
		Relation	0.1	0.1	0.5	0.3	0.2	0.6
		Both	0.2	0.2	0.1	0.4	0.2	0.9
	Absent	Distance	0.2	0	0.5	0.9	0.3	0.3
		Relation	0.7	0.1	0.3	0.8	0.3	0.6
		both	0.1	0.1	0.1	0.2	0.2	0.1

Table 17: The mean number of errors (out of 4) per condition in Experiment 10.

There was a main effect of display-size ($F(2,18) = 4.9$; $p < .05$). There were no significant pair-wise differences between the three display sizes. There was a significant interaction between target-presence and display-size (F2,18) = 11.3; $p < .01$). The effect of display-size on the number of errors made by participants was greater when the target was absent than when the target was present. There was a significant interaction between target-presence and distracter-type ($F(2,18) = 5.71$; $p < .05$). When the target was present and the distracters differed from the target in both reference frame and distance there were more errors (0.60), than when the target was absent (0.28). The effect of target presence was reversed when the distracters differed from the target in terms of the distance, 0.26 when present and 0.44 when the target was absent. The interaction between target reference frame and target distance approached significance ($F(1,9) = 3.7$; $p = .087$). Figure 50 shows the mean number of errors for each reference frame when the targets were short and long. There was no effect of distance when the participants used a relative reference frame to apprehend the spatial relationship. However, there was a tendency to make fewer errors when the target was at a short distance, compared to a long distance when participants used the intrinsic reference frame. There was also a significant three-way interaction between target presence, target distance and display size ($F(2,18) = 4.3$; $p < .05$). All other interactions were non-significant, all Fs < 2.1.

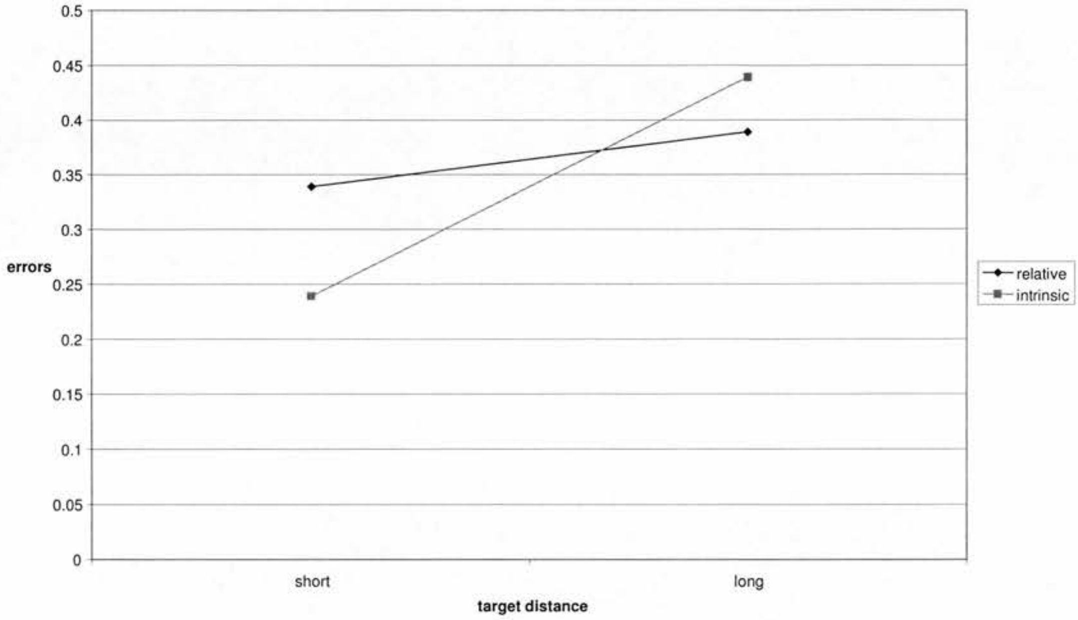


Figure 50. The mean proportion of errors for targets at each distance using each reference frame.

5.5.5. Discussion

The goal of the experiment was to examine firstly, whether there was a time difference in apprehension of spatial relations using different reference frames. There was no effect of reference frame on response latency or accuracy. This suggests that people are equally capable of using an intrinsic or a relative reference frame. The second goal of the experiment was to investigate whether different reference frames were associated with different settings of the distance parameter. We found that this is not the case for response latency as there was no interaction between the target reference frame and the target distance. However, in the response accuracy analysis the interaction between target reference frame and target distance did approach significance. There was no influence of distance upon response accuracy when the reference frame used was relative. This suggests that either the relative reference frame is not associated with a specific setting of the distance parameter, or that the differences in distance used in this experiment were not sufficient to elicit a distance effect. When the participants used an intrinsic reference frame they were more accurate when the target was short than when the target was long. This is tentative evidence in favour of the intrinsic reference frame being associated with a shorter setting of the distance parameter.

There was a significant effect of distracter type on both the response latencies and the response accuracy. In this experiment, participants were fastest and more accurate at responding when the distracters differed from the target in terms of reference frame and distance. This suggests that the target was more easily discriminated from the distracters in this condition, possibly because there were fewer things in common between the target and distracters. As in other visual search experiments there was an effect of display size on response latency and response accuracy (e.g. Logan, 1994). This confirms the results of Experiment 9 that setting the distance parameter does require the modulation of attention.

There was no effect of distance on response latency, although it did approach significance. The trend for response latency was for shorter target distances to result in increased response latency than longer target distances. This is the same pattern of results as reported in Experiment 9, and the lack of significant result in the present experiment may be because of the small sample size used. However, as in Experiment 9 any effect of distance is probably due to salience rather than distance per se. The effect of distance on response accuracy did reach significance, and the pattern of results was the same as in Experiment 9. Participants made fewer errors for short-distance targets than long-distance targets and this suggests a speed-accuracy trade-off; however, it is not clear why participants should be faster but less accurate for long distance targets and slower but more accurate for short distance targets. Possibly it is because participants perceived the short distance targets as being more difficult to find and so took more care on these trials.

This experiment largely supported the findings of Experiment 9; using the distance parameter to apprehend a spatial relation requires the modulation of attention. There was no effect of the reference frame on response latency or accuracy and this is evidence that both an intrinsic and relative reference frame can be used to apprehend spatial relationships equally easily. Finally, there was some suggestive evidence that the intrinsic reference frame is associated with a shorter setting of the distance parameter and that the relative reference frame has a more flexible range of settings for the distance parameter. However, this effect was marginal and appeared only in the accuracy data. In addition the small sample size means that the evidence for reference frame associated distance settings is tentative, but warrants further investigation.

5.6. Conclusion

This chapter has presented three experiments that have investigated how people set the distance parameter of a reference frame when comprehending spatial descriptions. Experiment 8 tested the influence of functional relatedness between the figure and the reference object on the distance parameter. The results showed that when the figure and reference object are functionally related the distance parameter is set as shorter than when the two objects are unrelated. This supports the argument of Carlson and Covey (2005) that the interaction between the figure and reference objects is important for setting the distance parameter.

Experiments 9 and 10 investigated the use of attention in apprehending the distance between a figure and a reference object. Both experiments showed that apprehending a spatial relation using distance requires the use of attention. This is an important finding as there is no apriori reason to suggest a search based upon distance should require attention. Both experiments also showed that when the distance between the figure and reference objects was long the participants were faster than when the distance was short. This is contrary to other research which has showed either no effect of distance on speed of apprehension (Carlson & Logan, 2001, Logan & Compton, 1996), or that shorter distances are apprehended faster (Carlson & Van Deman, 2004). However, the distance effect in these experiments is probably due to target salience rather than setting the distance parameter.

Experiment 10 was designed to investigate whether or not the distance parameter was defined independently of the reference frame used. There was no evidence that some distance settings are associated with the use of specific reference frames in the response latency analysis. However, there was a marginal effect in the accuracy analysis; suggesting that the intrinsic reference frame is associated with shorter distance settings and that the relative reference frame is independent of specific distance settings. This, however, is a tentative conclusion because the effect was marginal and was not evident in the response latency data.

Chapter 6: Conclusion

6.0. Chapter Overview

This thesis has presented the results of 10 experiments that have investigated the use and representation of reference frames, in both dialogue and monologue contexts. This chapter describes the major findings of the experiments and discusses directions for future research. The thesis contains three experimental chapters, and as such this chapter is broken into three sections, each of which discusses findings and issues from each of the experimental chapters.

6.1. Chapter 3: Alignment Experiments

6.1.1 Summary of Chapter 3 findings.

The experiments in Chapter 3 investigated whether or not interlocutors align reference frames during dialogue. The results showed that when an interlocutor uses a reference frame to describe the location of an object, their partner is more likely to use the same reference frame on an immediately subsequent occasion. This is strong evidence that interlocutors align reference frames. Importantly, reference frame alignment was shown to be independent of lexical alignment and that when interlocutors do align reference frames, they align the entire reference frame. This last finding is an important one, as it is possible that interlocutors only align a single axis or end point of an axis; for example they could just align on the use of a the intrinsic top-bottom axis. However, if an interlocutor describes a scene using an intrinsic reference frame, for example, *the ball is right of the car*, then their partner is more likely to use an intrinsic reference frame and describe something as *the ball is above the car* on an immediately subsequent occasion. This shows alignment from the left-right axis to the top-bottom axis.

The reference frame alignment effect was interpreted with respect to the interactive alignment model of dialogue (Pickering & Garrod, 2004). Interlocutors align at multiple levels of representation (e.g. lexical, syntactic, and conceptual). Reference frames can now be added to the list of representations that interlocutors align.

6.1.2. Mechanisms of Alignment

That alignment of reference frames occurs between interlocutors is evident. However, the results of this thesis do not clearly establish the mechanism by which alignment occurs. Pickering and Garrod (2004) have argued that alignment, at other levels of representation, between interlocutors is achieved by a resource free priming mechanism. The alternative to this is that interlocutors make a conscious decision to use the same reference frame as their partner. In the experiments presented here, interlocutors did not negotiate about which reference frame to use, which is typical of alignment effects (e.g. Garrod & Anderson, 1987). Certainly overt negotiation would suggest a level of conscious decision, but the lack of it does not necessarily indicate an automatic mechanism. However, the distinction between more automatic and more conscious mechanisms of alignment may be a false one.

Underlying all conscious decisions there are many unconscious processes. Therefore, in the present case the interlocutors may 'choose' to use the same reference frame as their partner, which would seem to be conscious, but the reason for their 'choice' is that, in comprehending the reference frame immediately previously, that reference frame becomes easier to use, but this is an unconscious effect of processing. Under the assumption that people are aiming to make language use efficient and quick, people will 'choose' the reference frame that is easier to produce. In this situation, the choice appears conscious, but the motivation for the choice is unconscious (i.e. the reference frame is easier to use, because of a priming mechanism). Indeed, the motivation for the choice (the priming mechanism) is probably not available to consciousness. However, people will be able to construct a motivation for their choice of reference frame, even if the construction is motivated by an after the fact realisation that they are using the same reference frame as their partner. For example, upon using the same reference frame as their partner, the speaker realises this and then decides that the reason they did this was because their partner also used that reference frame. Note that at this point the repetition has happened and the constructed reason for it is the speaker's interpretation, after the fact. However, consciousness is a constructive process and, because causes occur before their effects, the experience of the motivation seems, to the individual, as if it occurred prior to reference frame selection. In this way reference frame alignment can be conscious, but caused by automatic processes, and can even be automatic, but seem like a conscious decision.

6.1.3. Future Research

There was no increase in the amount of alignment of reference frames due to lexical overlap. This may be because prepositions form a closed linguistic class, which have been shown not increase the amount of syntactic priming when repeated from prime to target (Bock, 1989). However, it may also be because the meaning of the prepositions was dependent upon whether it was interpreted using an intrinsic or a relative reference frame. Therefore, there was no advantage, in terms of comprehension or production, in using the same lexical term as the confederate. If this latter hypothesis is the case then a lexical boost effect should be evident if different lexical terms are used for the intrinsic reference frame to the relative reference frame. For the top-bottom axes of a reference frame this situation is possible. The confederate can describe spatial relations with one reference frame using the terms *over* and *under* (e.g. *the dot is under the camera*) and use the terms *above* and *below* in the alternative reference frame. In such a scenario the spatial term used by the confederate is indicative of the location of the figure object, independent of the reference frame, and so a 'lexical boost' to reference frame alignment may occur. That is interlocutors may be more likely to use the same reference frame as their partner when doing so requires using the same lexical term as their partner, but only if the spatial terms are used in a reference frame specific way.

6.2. Chapter 4: Reference frame taxonomies

6.2.1. Summary of Chapter 4 findings

Chapter 4 presented four experiments that used a confederate priming paradigm to investigate the representation of reference frames. The aim of these experiments was to provide empirical evidence in favour of either Levinson's taxonomy of reference frames or the Traditional taxonomy of reference frames. The results of Experiments 4 and 5 provided evidence in favour of the Traditional taxonomy; that reference frames should be categorised based upon the object that determines the direction of the axes.

The deictic reference frame is an egocentric reference frame; however, some researchers have also argued that a reference frame is deictic if it is an addressee centred reference frame (Levelt, 1989, Miller & Johnson-Laird, 1976). Experiment 6 investigated whether or not a reference frame should be defined as deictic if it was a person-centric reference frame. The person could be the speaker, addressee or a third person. The results of Experiment 6 suggested that any person-centric reference frame was represented as deictic.

The results of Experiment 7 suggested that the results of Experiments 4-6 may have been due to the reference object being in a different position on the match scenes as on the target scenes. When the reference object was in a different position on the match scene to the target scene the participants were less likely to use an intrinsic reference frame following hearing the confederate use an intrinsic reference frame compared to when the reference object was in the same position and the participants heard the confederate use an intrinsic reference frame.

6.2.2. Reference Frames

Chapter 3 argued that interlocutors align reference frames. Chapter 4 then went on to use this as a way of investigating the representation of reference frames. However, it seems that reference frames may be more difficult to define than previously thought. For Levinson (1996, 2003), they are a linguistic representation that is defined according to the argument structure, and for the Traditional theory, reference frames are based upon the origin and orientation of the reference frame. However, in opposition to Levinson's idea, the results of this thesis have shown that interlocutors align the origin and orientation of a reference frame in dialogue. This means that in order to accept Levinson's classification system then it has to be believed that interlocutors align origin and orientation across reference frames. This may well be the case, but an effect of the argument structure of reference frames on cognitive processing has yet to be demonstrated. If it cannot be demonstrated, then it is plausible that argument structure differences between reference frames do not have psychological reality.

However, it is possible that reference frames, as the discrete representations in which they are presented, have no psychological reality either. Rather, a reference frame can be thought of as the activation and setting of a number of representations and parameters that are spatial, conceptual and linguistic in nature: a reference frame is the sum of representations that are active in the process of translating a perceptual representation into a linguistic representation. Jackendoff (2002) views a lexical entry in an analogous fashion. According to Jackendoff, a lexical entry is a combination of a word's phonological representation, grammatical representation, conceptual representation and even its plan for articulation. A reference frame can be viewed as the collection of representations that allow the translation of a perceptual representation to a linguistic representation. Part of a reference frame would be the argument structure of the sentence and part would be the origin, orientation, direction and scale parameters. Viewed this way, arguing two reference frames are the same because they have a common origin would be akin to saying two lexical entries are the same because they share the same phonological representation, which is obviously not the case. If reference frames are like this then under different contextual circumstances interlocutors should align different aspects of a reference frame, but alignment will occur between reference frames if they are defined as separate, discrete representations, as is the case in the results shown in Chapter 4 of this thesis.

6.2.3. Future Research

Experiments 4-7 were designed to provide support in favour of either Levinson's or the Traditional taxonomy of reference frames. The results so far have been inconclusive and therefore the goal of immediate future research is to extend these experiments to investigate how different aspects of reference frames are aligned between interlocutors, and ultimately provide support for on or the other taxonomies.

The amounts of alignment observed in Experiments 4-6 was comparative to the amount of alignment observed in Experiments 1-3. This was despite the reference objects being repeated from prime to target in the Chapter 3 experiments, but not in the Chapter 4 experiments. This suggests that the reference object identity is irrelevant for alignment, possibly because of schematization. However, a direct comparison is not possible, because the effect is between experiments. This, therefore, becomes an issue for future experimental work. Talmy (1983) has suggested that when using spatial language very little information about the figure and reference object is represented (e.g. in the sentence *the ball is in front of the car* the ball and car are represented as points). However, some features of the reference and figure object do seem to be represented (e.g. functional features, and dynamic-kinematic features; Coventry & Garrod, 2004). This presents the question as to which features of a reference object are represented fully and which as schematisations? An answer to this question may be possible using a confederate priming paradigm: If a feature of the reference object is explicitly represented, then repeating this from the prime to the target should yield a larger amount of reference frame alignment. When a feature of the reference object is not explicitly represented then repeating it from the prime to the target should have no effect on the amount of reference frame alignment.

6.3. Chapter 5: The Distance Parameter

6.3.1. Summary of Chapter 5 findings

Chapter 5 presented three experiments that investigated how the distance parameter of a reference frame is set. Experiment 8 asked participants to imagine two objects, which could either be functionally related or not related, in a spatial relationship, and then report the imagined distance between two objects. The results showed that participants imagined the distance between the figure and the reference object as smaller when they were functionally related, compared to when they were functionally unrelated. This is argued to be evidence in support of the claim that setting the distance parameter is influenced by the interaction between the figure and the reference object (Carlson & Covey, 2005).

Experiments 9 and 10 investigated the use of attention in setting the distance parameter. These experiments used a visual search paradigm, where participants had to verify whether or not a target was present amongst a number of distracters. The results showed that setting the distance parameter in order to search for the target, requires the use of attention. The experiments also investigated whether the distance parameter is set independently of the reference frame used. There was no evidence of reference frame dependent distance settings in the response latency analyses. However, there was a trend, in the accuracy data, which suggested that shorter settings of the distance parameter are associated with the intrinsic reference frame.

6.3.2. Future Research

There was tentative evidence showing that the intrinsic reference frame may be associated with shorter settings of the distance parameter. However, this effect was marginal and was only present in the accuracy data. However, it does warrant further research. One reason an effect of reference frame on setting the distance parameter may not have been evident is that participants were told which reference frame they should use to search for the target. If participants are told to search for a relationship between two letters, at a specified distance, according to any reference frame then distance effects may emerge. This is because participants would only be able to set the distance parameter and would not be able to set the direction parameter of the reference frame prior to seeing the experimental display.

The findings also suggested that the more parameters of a reference frame that require setting to apprehend a spatial relationship, the more difficult the task is. Participants were faster when the target differed from the distracters in terms of distance (which required setting a single parameter) than when the target differed in terms of the spatial relationship between the two objects (which required setting two parameters; both the orientation and direction). An issue for future research is to investigate whether all parameters of a reference require the same amount of cognitive effort to set.

References

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions. *Journal of Memory and Language*, 42, 1-22.

Bard, E. G., & Aylett, M. P. (2004). Referential form, word duration, and modelling the listener in spoken dialogue. In J. Trueswell, & M. Tanenhaus, (Eds.), *Processing World Situated Language: Bridging the Language-as-product and Language-as-action Traditions*. Cambridge, MA: MIT Press.

Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28, 937-962.

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46, 391-418.

Barsalou, L., Sloman, S., & Chaigneau, S. (2004). The HIPE theory of function. In L. A. Carlson, & E. van der Zee (Eds.), *Functional Features in Language and Space: Insights from Perception, Categorization, and Development* (pp131-147). Oxford: Oxford University Press.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79, 941-952.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355-387.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, 31, 163-186.

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin and Review*, 3, 395-421.

Boroditsky, L., (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75, 1-28.

Bowerman, M. (1996). Learning how to structure language for space: A crosslinguistic perspective. The architecture of the linguistic-spatial interface. In P.Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space,* (pp. 385-436).Cambridge, MA: MIT Press.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition,* 75, B13-B25.

Brennan, S. E., 7 Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 1482-1493.

Bryant, D. J. (1997). Representing space in language and perception. *Mind and Language*, 12, 239-264.

Bryant, D. J., & Wright, W.G. (1999). How bodily asymmetries determine accessibility in spatial frameworks. *Quarterly Journal of Experimental Psychology (A)*, 52, 487-508.

Buhl, H. M. (2001). Partner orientation and speaker's knowledge as conflicting parameters in language production. *Journal of Psycholinguistic Research*, 30, 549-567.

Burigo, M., & Coventry, K. (2004). Spatial language and reference frame assignment: The role of the located object. *Proceedings of the 25th Annual Conference of the Cognitive Science Society.*

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 35-69.

Carlson, L. A. (2000). Object use and object location: The effect of function on spatial relations. In E. van der Zee and U. Nikanne (Eds.), *Cognitive Interfaces: Constraints on Linking Cognitive Information* (pp. 94-115). Oxford: Oxford University Press.

Carlson, L. A. (2003). Using spatial language. *The Psychology of Learning and Motivation*, 43, 127-161.

Carlson, L. A, & Covey, E. S. (2005). How far is near? Inferring distance from spatial relations. *Language and Cognitive Processes*, 20, 627-631.

Carlson, L. A., & Logan, G. A. (2001). Using spatial relations to select an object. *Memory and Cognition*, 29, 883-892.

Carlson, L. A., & Van Deman, S. R. (2004). The space in spatial language. *Journal of Memory and Language*, 51, 418-436.

Carlson, L. A., & Van Deman, S. R. (in press) What is active when a reference frame is activated?

Carlson, L. A., West, R., Taylor, H. A., & Herndon, R. W. (2003). Neural Coordinates of spatial term use. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1391-1408.

Carlson, L. A., & van der Zee, E. (2004). *Functional Features in Language and Space*. Oxford: Oxford University Press.

Carlson-Radvansky, L. A., Covey, E. S., & Lattanazi, K. L. (1999). "What" effects on "where": Functional influences on spatial relations. *Psychological Science*,10, 516-521.

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, 46, 223-244.

Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language*, 33, 646-671.

Carlson-Radvansky, L. A., & Jiang, Y. (1998). Inhibition accompanies reference frame selection. *Psychological Science*, 9, 386-391.

Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, 37, 411-437.

Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, 7, 56-60.

Carlson-Radvansky, L. A., & Tang, Z. (2000). Functional influences on orienting a reference frame. *Memory and Cognition*, 28, 812-820.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology*, 76, 893-910.

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, 214-230.

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

Clark, H. H., Carpenter, P. A., & Just, M. A. (1973). On the meeting of semantics and perception. In W. G. Chase, (Ed.), *Visual Information Processing*. New York: Academic Press.

Clark, H. H., & Krych. M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62-68.

Clarke, H.H., & Wlkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

Coventry, K. R. & Garrod, S.C, (2004). *Saying, Seeing and Acting. The Psychological Semantics of Spatial Prepositions*. Psychology Press. Hove and New York.

Coventry, K. R., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of 'over', 'under', 'above', and 'below'. *Journal of Memory and Language*, 44, 376-398.

Crawford, L. E., Regier, T., & Huttenlocher, J. (2000). Linguistic and non-linguistic spatial categorization. *Cognition*, 75, 209-235.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-234.

Fowler, C., & Housum, J. (1987). Talkers signaling 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489-504.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, 119, 63-76.

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models. *Memory and Cognition*, 20, 507-518.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.

Friederici, A. D., & Levelt, W. J. M. (1990). Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception and Psychophysics*, 47, 253-266.

Garnham, A. (1989). A unified theory of meaning of some spatial relational terms. *Cognition*, 31, 45-60.

Garrod, S. G., & Anderson, A, H. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181-218.

Garrod, S. G., & Clark, A. (1993). The development of dialogue co-ordination skills in schoolchildren. *Language and Cognitive Processes*, 8, 101-126.

Garrod, S. G., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*,53, 181-215.

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

Giles, H., & Powesland, P. F. (1975). *Speech Styles and Social Evaluation*. New York: Academic Press.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 8, 875-894.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and Semantics 9: Speech Acts*, pp. 41-58. New York: Seminar Press.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential communication. *Journal of Memory and Language*, 49, 43-61.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15, 409-414.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, 55, 39-84.

Haywood, S. L., & Pickering, M. J., & Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science*, 16, 362-366.

Herskovitz, A. (1986). *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions of English*. Cambridge: Cambridge University Press.

Horton, W. S., & Gerrig, R. J. (2005a). The impact of memory demands on audience design during language production. *Cognition*, 96, 127-142.

Horton, W. S., & Gerrig, R. J. (2005b). Conversational common ground and memory processes in language production. *Discourse Processes,*

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91-117.

Huttenlocher, J., Hedges, L. V., Corrigan, B., & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition*, 93, 75-97.

Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116, 26-37.

Jackendoff, R. (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P.Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space,* (pp. 1-30).Cambridge, MA: MIT Press.

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

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, 13, 289-303.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychological Science*, 11, 32-38.

Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language*, 39, 1-20.

Klabunde, R., & Porzel, R. (1998). Tailoring spatial descriptions to the addressee: A constraint based approach. *Linguistics*, 36, 551-577.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for 2 types of spatial representations – Hemispheric specializations for categorical and co-ordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 723-735.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1993). Cetegorical versus co-ordinate spatial relations – computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 562-577.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioural and Brain Sciences*, 16, 217-265.

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

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14, 78-106.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind, and J. J. Koenderink (Eds.), *Limits in Perception* (pp323-358). Utrecht. VNU Science Press.

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P.Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space,* (pp. 77-108).Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.

Levelt, W. J. M., Schriefers, H., Norberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production. A study of picture naming. *Psychological Review,* 98, 122-142.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* Chicago: University of Chicago Press

Levinson, S. C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Levinson, S. C. (1996). Frames of reference and Molyneux's questions: Cross linguistic evidence. In P.Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space,* (pp. 109-169).Cambridge, MA: MIT Press.

Levinson, S. C. (2003). Space in language and cognition: Explorations in cognitive diversity. Cambridge: Cambridge University Press.

Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: Language affects spatial reasoning. *Cognition*, 84, 155-188.

Logan, G. D. (1994). Spatial attention and apprehension of spatial relations. *Journal of Experimental Psychology: Human, Perception and Performance*, 20, 1015-1036.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, 28, 103-174.

Logan, G. D., & Compton, B. J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 159-172.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, and M. Garrett (Eds.), *Language and Space* (pp. 493-529). Cambridge, MA: MIT Press.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49, 201-213.

Maki, R., Grandy, C. A., & Hauge, G. (1979). Why is telling left and right more difficult than telling above from below. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 52-67.

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology*, 46, 229-259.

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *Journal of Speech, Language, and Hearing Research*, 44, 128-143.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge, MA: MIT Press.

Munnich, E., Landau, B., & Dosher, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, 81, 171-207.

Pickering, M. J. & Garrod, S. C. (2004). Towards a mechanistic theory of dialogue. *Behavioral and Brain Sciences, 27,* 169-226.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 940-961.

Radvansky, G. A., & Copeland, D. E. (2000). Functionality and spatial relations in memory and language. *Memory and Cognition*, 28, 987-992.

Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, 130, 273-298.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Bayes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.

Ross, L. D., Greene, O., & House, P. (1977). The false consensus phenomenon: An attributional bias in self perception and social perception processes. *Journal of Experimental Social Psychology*, 13, 279-301.

Russell, A. W., & Schober, M. F. (1999). How beliefs about a partner's goals affect referring in goal-discrepant conversation. *Discourse Processes*, 27, 1-33.

Schober, M. F. (1993). Spatial perspective taking in conversation. *Cognition*, 47, 1-24.

Schober, M. F. (1995). Speakers, addressees and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, 20, 219-247.

Schober, M. F. (1998). How addressees affect spatial perspective choice in dialogue. In P. L. Olivier, & K. P. Gapp (Eds.), *Representation and Processing of Spatial expressions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture word interference studies. *Journal of Memory and Language*, 29, 86-102.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin and Review*, 4, 102-106.

Shockley, K., Santana, M. V. and Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance,* 29, 326-332

Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory and Cognition*, 20, 483-496.

Talmy, L. (1983). How language structures space. In H. L. Pick and L. P. Acredolo (Eds.), *Spatial Orientation: Theory, research and Application* (pp. 225-282). New York: Plenum Press.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention.*Cognitive Psychology*, 12, 97-136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15-48.

Van der Zee, E., & Eshuis, R. (2003). Directions from shape: How spatial features determine reference axis categorization. In E. van der Zee and J. Slack (Eds.), *Representing direction in Language and Space*. Oxford: Oxford University Press.

Watson, M.E., Pickering, M. J. & Branigan, H. P. (2004). Alignment of reference frames in dialogue. *Proceedings of the 26th Cognitive Science Meeting*

West, R. & Alain, C. (1999). Event-related neural activity associated with the Stroop task. *Cognitive Brain Research*, 8, 157-164.

Appendix A

The Reference objects used in Experiment 1; shown here in their canonical orientations. In the experiment they were presented rotated 90º left or right:

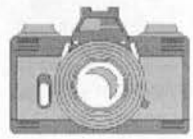
Anchor

Apple

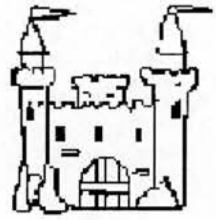
Bell

Bottle

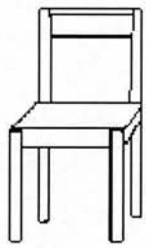
Bowl



Camera



Castle



Chair

Crown

Glass

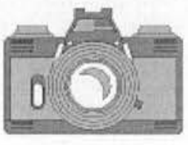
Mushroom

Telephone

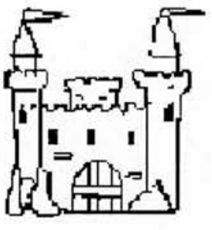
Well

Appendix B

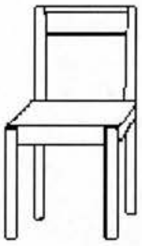
The Reference objects used in Experiment 2 and 3. Here shown in their canonical orientations. In the experiment they were presented rotated 90º left or right:



Camera



Castle



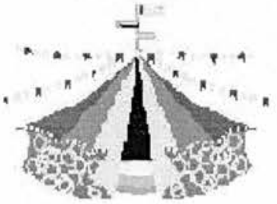
Chair

Lock

Telephone

Cupboard

Cello



Tent

Clock

Medal

Stereo

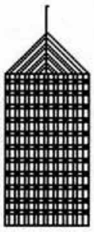
Television

Appendix C

The biaxial objects used in Experiments 4-6. The first three are the *land* objects and the second three are the *sea* objects.

Mast

Lighthouse



Skyscraper

Tree

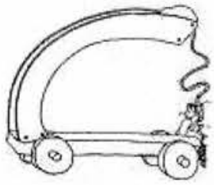
Buoy

Island

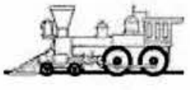
Appendix D

The tri-axial objects used in Experiments 4-6. The first three are the land items and the second three are the sea items.

Car



Catapult



Train

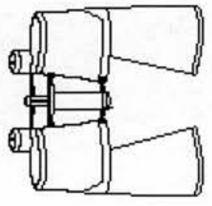
Hovercraft

Ship

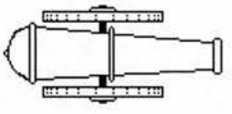
Windsurfer

Appendix E

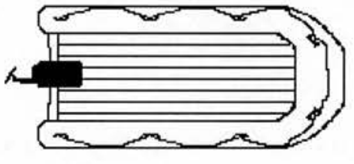
The items and names used to refer to them, used in Experiment 7.



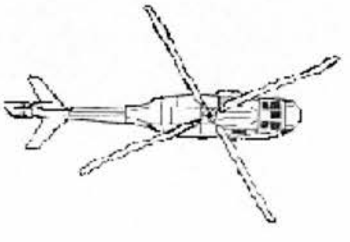
Binoculars



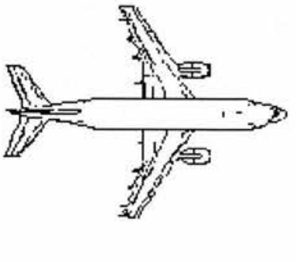
Cannon



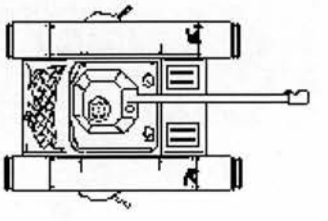
Dinghy



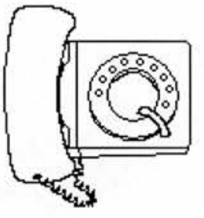
Helicopter



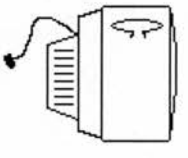
Plane



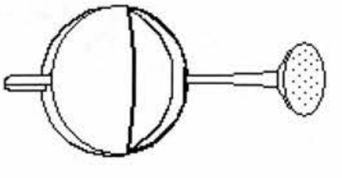
Tank



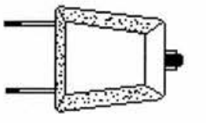
Phone



TV



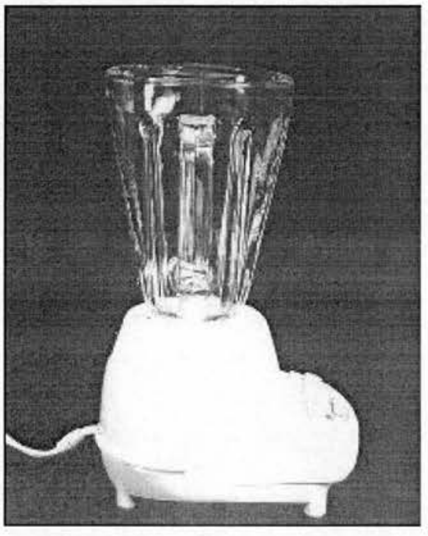
Watering Can



Wheelbarrow

Appendix F

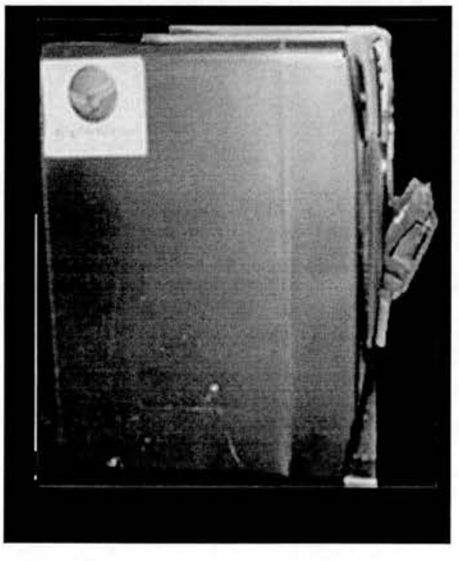
The 'above' reference objects and the functionally related and functionally unrelated figure objects used in Experiment 8.



Blender

Ice cube

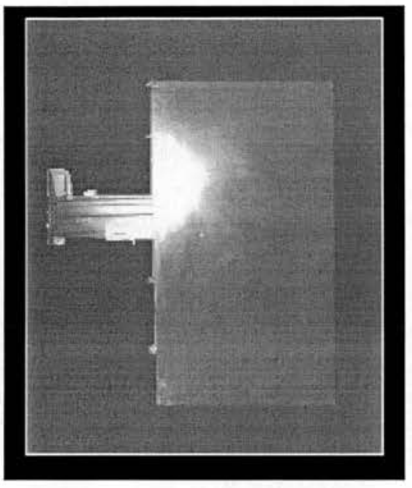
Block



Cassette Player

Cassette

Playing Cards



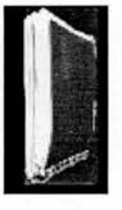
Filing cabinet

File

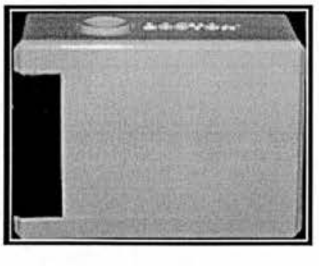
Clipboard

Newspaper dispenser

Newspaper



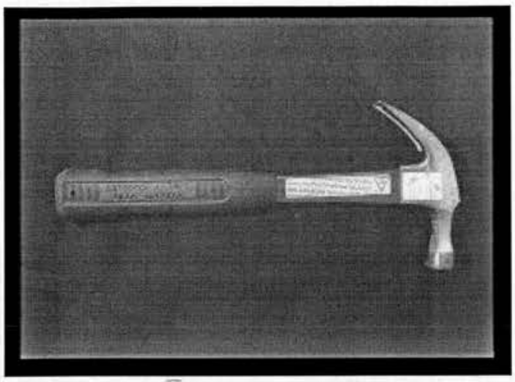
Notebook



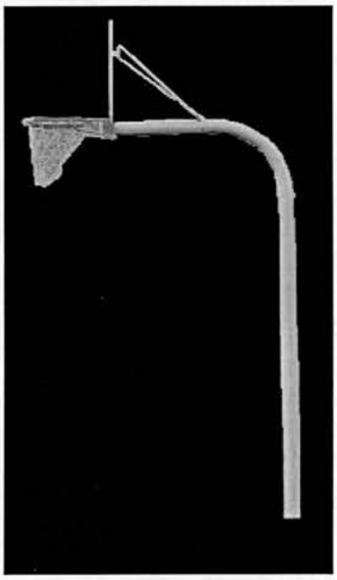
Pencil sharpener

Pencil

Paint brush



Hammer Nail Toothpick



Basketball hoop Basketball Globe