“Near” and “far” in language and perception

David Kemmerer*

Division of Behavioral Neurology and Cognitive Neuroscience, Department of Neurology, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242, USA

Received 5 March 1999; received in revised form 4 May 1999; accepted 11 June 1999

Abstract

A major interest in cognitive science is the relationship between linguistic and perceptual representations of space. One approach to exploring this relationship has been to investigate aspects of the linguistic encoding of space that correspond closely to aspects of the visual system. Another approach, which does not contradict the first but rather complements it, is to investigate ways in which linguistic and visual representations of space are different. This paper pursues the second approach by arguing that the distinction between proximal and distal demonstratives (e.g. this vs. that, here vs. there) does not correspond to an independently established distinction between near and far space in the visual system but is instead based on language-internal factors. Recent neuropsychological and neurophysiological studies suggest that the brain contains separate mechanisms for representing, on the one hand, near or peripersonal space which extends roughly to the perimeter of arm’s reach and, on the other hand, far or extrapersonal space which expands outward from that boundary. In addition, crosslinguistic research suggests that it is very common for languages to have two basic types of demonstrative terms — proximal and distal. This parallelism raises the possibility that the linguistic distinction may derive from the perceptual distinction. However, several arguments support the contrary view that the two distinctions are independent of one another. A substantial proportion of languages in the world have demonstrative systems that divide space into three or more egocentrically-grounded regions, thereby violating the two-way perceptual contrast. Even more importantly, empirical studies of how demonstratives are used in ongoing discourse in different languages suggest that they do not encode quantitative spatial information such as within vs. beyond arm’s reach; instead, they specify abstract semantic notions that, when combined with the unique pragmatic features of communicative contexts, allow speakers to make a virtually unlimited range of spatial distance contrasts. Thus, demonstratives constitute an interesting case of divergence between linguistic and perceptual representations of space. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Far space; Near space; Language-internal factors; Linguistic and perceptual representations; Proximal and distal demonstratives

* Tel.: +1-319-356-3165; fax: +1-319-356-4505.
E-mail address: david-kemmerer@uiowa.edu (D. Kemmerer)

0010-0277/99 - see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0010-0277(99)00040-2
1. Introduction

In recent years a great deal of research has been devoted to the relationship between natural language and spatial perception and cognition (e.g. Miller & Johnson-Laird, 1976; Jackendoff, 1983; Talmy, 1983; Talmy, 1988; Herskovits, 1986; Landau & Jackendoff, 1993; Haviland & Levinson, 1994; Bloom, Peterson, Nadel & Garret 1996; Regier, 1996; Bryant, 1997; Pütz and Dirven, 1997; Senft, 1997; Danziger, 1998; Pederson, Danziger, Wilkins, Levinsons, Kita & Senft 1998). One approach to exploring this relationship has been to concentrate on aspects of the linguistic encoding of space that correspond closely to aspects of the visual system. An excellent example of this approach is Landau and Jackendoff’s (1993) analysis of the distinction between count nouns that encode detailed geometric features of objects (e.g. swan, hand) and locative prepositions that encode schematic topological relationships like ‘‘containment’’ and ‘‘contiguity’’ (e.g. in, on). To explain this linguistic division of labor between ‘‘what’’ and ‘‘where,’’ Landau and Jackendoff suggest that it may derive from an independently established division of labor in the brain between, on the one hand, a ventral pathway that projects from the occipital lobe to the inferior temporal lobe and specializes in representing form, color, and texture features of objects, and on the other hand, a dorsal pathway that projects from the occipital lobe to the parietal lobe and specializes in representing the spatial properties of objects (Ungerleider & Haxby, 1994). In addition to accounting for the different semantic properties of concrete nouns and locative prepositions, the authors point out that the neural distinction may also explain why the former expressions constitute an open class whereas the latter belong to a closed class. Basically, the idea is that the ventral stream may contain a very rich set of primitive geometric units and principles for combining them into larger structures so that the system can generate thousands of distinct object representations (Biederman, 1987); by comparison, the dorsal stream may have a much more restricted set of specifications for the different kinds of spatial relationships that can obtain between objects (Rueckl, Care & Koss., 1988).

In addition to Landau and Jackendoff’s study, much of the work being done in the rapidly growing field of ‘‘cognitive linguistics’’ is explicitly concerned about grounding certain aspects of the linguistic encoding of space in properties of the visual system (e.g. Lakoff, 1987; Langacker, 1987; Langacker, 1991; Regier, 1996). All of this research illustrates how investigating points of convergence between linguistic and perceptual representations of space can lead to many interesting insights. Another line of inquiry, however, which has not received as much attention as the first, is to explore the ways in which spatial expressions diverge from the organizational properties of the visual system and reflect instead language-internal semantic and pragmatic factors (e.g. Bowerman & Bloom, 1996; Levinson, 1997; Kemmerer & Tranel, 1999). These two approaches need not contradict one another because it seems quite plausible that there are both similarities and differences between linguistic and perceptual representations of space. In the ideal scenario, the two approaches would proceed side by side and gradually lead to an integrated, balanced characterization of the interface between linguistic and perceptual repre-
sentations of space that captures not only the areas that do correspond but also the areas that do not.

This paper presents a case study that adopts the second approach. The goal is to show that there are significant differences in the ways in which natural language and the visual system divide space into distinct egocentrically grounded sectors, specifically ‘near’ and ‘far.’ The organization of the paper is as follows. Section 2 is concerned with the perceptual representation of near and far space. I review neuropsychological and neurophysiological research which suggests that the brain contains separate systems for representing space near the body, i.e. roughly within arm’s reach, and far from the body, i.e. roughly beyond arm’s reach. The system for near space appears to be evolutionarily designed primarily for the visuomotor control of arm, hand, and head movements, while the system for far space may be adapted for scanning the visual field, especially the distal field beyond the body, and recognizing ecologically significant objects within it. In Section 3 the focus shifts to the linguistic representation of near and far space. All human languages contain special grammatical terms, namely demonstratives, that serve to distinguish between different sectors of space, using the speaker or addressee as a frame of reference. Cross-linguistic research suggests that a large proportion of languages in the world make a fundamental binary distinction between proximal and distal demonstratives (e.g. *this* vs. *that, here* vs. *there*). This raises the intriguing possibility that a basic design feature of language may be to carve space into two main sectors that correspond closely to the sectors distinguished by the visual system — specifically, near the body (e.g. *this book here*) vs. far from the body (e.g. *that book over there*). I present two arguments, however, for the contrary view that the linguistic division of space into sectors does not parallel the perceptual division. First, a substantial proportion of languages do not have simple two-term demonstrative systems but instead have any of several types of three-term systems; there have even been reports of four-, five-, and seven-term systems. Second, empirical studies of how proximal and distal demonstratives are actually used in ongoing discourse in different languages suggest that the interpretation of these terms is almost never reducible to an objectively restricted distance distinction such as within vs. beyond arm’s reach. Instead, the spatial referential functions of the terms are usually highly relativized to the current speech and situational contexts and depend on extensive shared sociocultural knowledge between the interlocutors. Finally, I conclude in Section 4 by summarizing the main argument of the paper and discussing several implications it has for our conception of the interface between linguistic and perceptual representations of space.

2. The perceptual representation of near and far space

During the past decade, an increasing amount of research in cognitive neuroscience has shown that the way the brain represents space is extremely complex. Although the impression of three-dimensional space that the brain delivers to consciousness is unified and seamless, this does not reflect the structure of the underlying mechanisms
because in fact many different kinds of spatial ‘‘maps’’ for perception and motor control have been found in the brain (e.g. Kosslyn, 1994; Caminiti, 1995; Milner & Goodale, 1995; Colby & Duhamel, 1996; Fogassi et al., 1996; Jeannerod, 1997; Thier & Karnath, 1997; Burgess, Jeffery & O’Keefe 1999). The aim of this section is to review evidence suggesting that the brain contains separate neural systems for representing, on the one hand, near or peripersonal space demarcated roughly by the perimeter of arm’s reach, and on the other hand, far or extrapersonal space extending beyond the sphere of potential motor behavior.

2.1. Neuropsychological evidence

The strongest evidence for distinct representations of near and far space in the human brain comes from studies of subjects with a well-known neuropsychological disorder called neglect. In the majority of subjects the lesion involves the right inferior parietal cortex, especially the supramarginal gyrus, but the disorder can also result from damage in the premotor or supplementary motor cortices in the frontal lobe (Heilman, Watson, Valenstein & Damasio, 1983; Vallar & Perani, 1986; Husain & Kennard, 1996). In the most common form of neglect, the subject ignores an entire side, or hemifield, of egocentric space, usually the left side (e.g. Jeannerod, 1987; Robertson & Marshall, 1993; Halligan & Marshall, 1994; Rafal, 1994; Vallar, 1998). For example, subjects will incorrectly bisect horizontal lines to the right of the midpoint, thus neglecting the left side of the line, and when shown a page covered with isolated lines at different orientations and asked to cross out each one, they will cross out all the lines on the right side of the page but leave the ones on the left untouched.

Recent studies have found that neglect is not a single monolithic disorder but instead can be fractionated into a variety of more specific disorders each of which reflects certain components of the brain’s highly multifaceted architecture for spatial representation (Bisiach, 1997; Vallar, 1998). For purposes of this paper, the most important type of neglect is sometimes referred to as proximal/distal neglect. Some subjects manifest an inability to consciously represent the near sector of space, while other subjects appear to have an impaired representation of the far sector of space. These forms of neglect have been described in several studies (Halligan & Marshall, 1991; Cowey, Small & Ellis, 1994, 1999; see also Shelton, Bowers & Heilman, 1990; Mennemeier, Wertman & Heilman, 1992, and Farné, Zeloni & Ládavas, 1999).

Using exactly the same methods, two different studies found brain-damaged subjects who exhibited opposite types of neglect, one type involving left neglect in near but not far space, and the other involving left neglect in far but not near space. The first study, which was conducted by Halligan and Marshall (1991c), concerned a single subject with a large right temporoparietal lesion. The subject performed very poorly on the standardized Behavioral Inattention Test (Wilson, Cockburn&Halligan 1987) and also manifested classic left unilateral neglect on a line bisection task, making rightward displacements that were on average over an order of magnitude greater than in any member of a group of 20 normal control subjects. The main experiment consisted of two additional line bisection tasks in the
following conditions. First, the subject used an ink pen to bisect horizontal lines of different lengths that were placed on a vertically oriented board at eye level and at a distance of approximately 45 cm, well within arm’s reach. Then he used a laser pointer to perform a similar line bisection task only at a distance of 244 cm, well beyond arm’s reach. The investigators found that the subject exhibited severe left neglect in the first condition (near space) but not in the second condition (far space). Six weeks after the main experiment was completed, the subject was asked to perform the first task again only using a laser pointer held immediately in front of the body. No improvement was found, indicating that the original dissociation was not due to the different response modes. This pattern of results suggests that the subject has a selective impairment of the representation of the near left sector of space.

The second study was conducted by Cowey et al. (1994) and employed the same experimental procedures to test five brain-damaged subjects, all of whom had right-hemisphere lesions and were classified as having left unilateral neglect according to the Behavioral Inattention Test. In the main experiment, the only difference in protocol compared to Halligan and Marshall’s study was that the subjects used a laser pointer in both line bisection tasks. The results showed that none of the subjects exhibited left neglect in the near space condition in which the lines were placed only 45 cm away, but all five of them did exhibit significant left neglect in the far space condition in which the lines were placed at a distance of 244 cm. This suggests that the subjects have a selectively impaired ability to perceive and respond to stimuli in the far left sector of space.

The fact that the single subject in Halligan and Marshall’s study and the five subjects in Cowey et al.’s study demonstrated exactly the opposite performance profiles constitutes strong evidence that the brain contains separate neural systems for representing stimuli in near or peripersonal space on the one hand, and in far or extrapersonal space on the other. However, these studies do not provide information about whether the mental representation of the boundary between near and far regions of space is rigid or gradient. To address this question, Cowey, Small & Ellis (1999) recently carried out another study with thirteen brain-damaged subjects who had right-hemisphere lesions and who met criteria for left unilateral neglect on the basis of the Behavioral Inattention Test.

The subjects performed a series of line bisection tasks with a laser pointer at the following six distances: 25, 50, 100, 200, 300, and 400 cm. Eight of the subjects exhibited left neglect at all of the distances, but the other five showed a clear increase in the severity of left neglect at progressively farther distances. For these five subjects, bisection errors were not significantly different for any single step between two distances; however, errors were significantly worse at the farthest distance (400 cm) than at each of the three nearest distances (25, 50, and 100 cm). These results suggest that the boundary between near and far sectors of space may not be represented as a rigid cut-off point but rather as a continuous change that takes place across the zone corresponding roughly to the extent of arm’s reach. This is consistent with neurophysiological data that will be described below, and also fits with evidence from studies of unilateral left neglect suggesting that the
boundary between left and right hemifields is not rigid but gradual (Kinsbourne, 1993).

Although the studies by Halligan and Marshall, and Cowey et al., provide strong psychological evidence for separate representations of near and far space, the lesion sites for the brain-damaged subjects are too variable to support clear inferences about the anatomical substrates of the different systems. Halligan and Marshall’s subject had a very large right hemisphere lesion that included the inferior parietal cortex as well as much of the lateral and medial temporal cortex and a portion of the inferior frontal lobe, and the subjects in both of Cowey et al.’s studies had right-hemisphere lesions that varied considerably across the parietal, temporal, and frontal lobes with little overlap. Despite these limitations, however, there are a number of other studies that do allow generalizations to be made about the probable anatomical substrates of the different systems for near and far spatial representation. These studies are reviewed in Section 2.3; but before getting to them, it is important to address the question of why the brain has different systems for representing near and far space.

2.2. A functional explanation

Although it is tempting to assume that the sole function of the visual system is to create a detailed replica of the visual world, it may be more appropriate to think of the visual system as having evolved in response to the demands of efficient motor control (Churchland, Ramachandran, Sejnowski 1994; Milner & Goodale, 1995; Colby & Duhamel, 1996; Fogassi, Gallese, Fadiga, Luppino, Matelli & Rizzolatti, 1996; Rizzolatti, Camarda, Fogassi, Gentilucci, Luppino & Matelli, 1994; Rizzolatti, Fadiga, Fogassi & Gallese, 1997). From this point of view, it is possible to understand why the brain appears to contain different systems for representing near and far space. In particular, it is likely that these two systems have been designed by natural selection to solve fundamentally different kinds of computational problems involving motor control.

Near or peripersonal space is the arena in which the visuomotor control of arm and hand movements takes place, and this kind of behavioral control became increasingly important during primate and hominid evolution (Previc, 1990;Jeanneord, 1997; Wilson, 1998). In primates the frequent adoption of a sitting or upright posture enabled the arms and hands to be used less for bodily support and more for manipulatory behaviors, such as retrieving, preparing, and ingesting food, or grooming either the self or conspecifics; and in hominids the development of bipedalism freed the arms and hands even more, allowing them to become progressively more adapted for a variety of skilled actions, including tool manufacture and use. Selection pressures therefore led to the refinement of a set of brain mechanisms specialized for the visual guidance of manual behavior in the peripersonal environment: stereo vision for determining the precise depth of objects in the near visual field; an attention system for guiding the hand to a fixated object when the hand cannot be directly viewed; a smooth pursuit tracking system for visually monitoring the hand as it reaches toward an object and as it brings the object towards the face for inspection or ingestion; and the coordination of multiple spatial representations...
based on effector-specific frames of reference — e.g. eye, head, shoulder, arm, and hand (for detailed discussion see Previc, 1990).

By comparison, far or extrapersonal space is the main arena in which visual search and object scanning and recognition take place, and these kinds of operations also became increasingly important during primate and hominid evolution (Previc, 1990). For example, there was greater reliance on colored fruits as a food source and an enhanced role of facial expression and bodily posture in communication. These selection pressures led to a dramatic improvement in several brain mechanisms designed for representing objects principally, but not exclusively, in the extrapersonal environment. A voluntary saccadic eye movement system evolved to enable visual search and object scanning independent of head movements. This oculomotor system can be thought of as supporting the only kind of immediate action that can be taken toward objects in the distal field, namely foveation or ‘visual grasp’ as well as part-whole analysis or ‘parsing’ by means of sequentially fixating on different components of the object and thereby building up a complete visual representation of its structure (Churchland et al., 1994). In addition to the development of the saccadic system, there were evolutionary advances in optical resolution, color perception, and shape discrimination, especially for faces. It is also worth noting that the value of all of these mechanisms probably increased significantly when hominids made the transition from woodlands to the open savannah where the extrapersonal visual field is larger (again, see Previc, 1990 for detailed discussion).

These evolutionary considerations provide the necessary foundation for understanding why the brain has different systems for representing near and far space. In short, it was highly adaptive in the ancestral primate and hominid environments to be able to construct representations of both spatial fields, but the two types of representation require quite different types of computations; hence distinct systems evolved, each specialized for representing one spatial domain.¹

2.3. Neural substrates

Very little is known about the architecture of the near and far spatial systems in the human brain; however, a number of neurophysiological studies have been done with macaque monkeys, and these studies allow us to glimpse at least the broad outlines of how the two systems are organized in these primates. The following discussion focuses on the studies that provide direct support for the distinction between near and far spatial representations.

First of all, it is possible that separate representations of near and far sectors of space are derived in part from an initially undifferentiated representation of the spatial environment that is anchored in head-centered coordinates (Rizzolatti et

¹ An anonymous referee pointed out that a third kind of perceptually guided motor control involves determining how to move one’s body — in short, how to locomote — so that a distant object which has been detected and positively appraised (by the far system) will fall within the peripersonal sphere and hence be available for manipulation (by the near system). I leave it to future research to explore how the mechanisms required by this kind of planning process relate to the ones that underlie the near and far systems.
This representation may reside in area PO (anterior bank of the parieto-occipital sulcus). Galletti, Battaglini and Fattori (1993) found neurons in this area that respond to visual stimuli in particular spatial locations, and the receptive fields of these neurons do not move with the eyes but rather with the head; hence they are based on a craniotopic rather than a retinotopic frame of reference, leading Galletti et al., to refer to them as "real position" neurons. The authors also suggest that these neurons may provide spatial information to other cortical structures that need the information for representing just the near or just the far visual world.

A large number of brain areas contribute to the representation of near space (see Fig. 1), and more data are available regarding this system than the far system. In computational terms, one important function of the near system is to transform a single representation of peripersonal space in head-centered coordinates into multiple representations of peripersonal space in body-part-specific coordinates; the latter representations can then be used to program movements, especially arm and hand movements directed toward or away from objects. Another function of this system is to visually monitor arm and hand actions relative to the face. As a reflection of these different purposes, the near system can be decomposed into several different functional-anatomical circuits that project from the parietal lobe to the frontal lobe — one for reaching, one for grasping, and one for monitoring limb movements in relation to the face.

The reaching circuit includes at least two parietal areas, 7b and MIP (medial...
intraparietal sulcus), both of which receive input from PO (Colby & Duhamel, 1991). Neurons in 7b fire immediately prior to and during reaching movements (Hyvarinen & Poranen, 1974; Mountcastle, 1975; Mountcastle, 1995; McKay, 1992; Andersen, 1995). In addition, a large proportion of neurons in this area are bimodal since they respond not only to light cutaneous stimulation on the arm, chest, or face, but also to visual stimuli that are within a region of space extending out from the surface of the given body part a certain distance, roughly between 20 and 100 cm (Graziano & Gross, 1995). In accordance with the selectivity of neurons in 7b for near space, lesions in this area cause neglect for near but not far space — for example, the animal no longer attempts to handle or bite food within reaching distance, but continues to react normally to objects placed farther away (Rizzolatti, Matelli & Pavesi, 1983). With regard to area MIP, although it is primarily a somatosensory area, it also contains many distance-sensitive visual neurons that respond more strongly to stimuli within arm’s reach than to stimuli further away (Colby & Duhamel, 1996). Both 7b and MIP project to area F4 in the dorsal premotor cortex of the frontal lobe. Neurons in this area fire during reaching movements (Gentilucci and Rizzolatti, 1990) and have bimodal tactile and visual response properties that are anchored to specific body parts such as the arm, chest, or face, very similar to neurons in 7b (Graziano & Gross, 1995; Fogassi et al., 1996; Graziano, Hu & Gross, 1997a,b; Rizzolatti et al., 1997). Some examples of the receptive fields of these neurons are shown in Fig. 2. Although these neurons are sensitive to visual stimuli within a certain region of space relative to a particular body part, two findings suggest that the boundaries of the spatial region are not rigidly fixed but are instead somewhat fuzzy: first, neurons have a gradient firing response that is strongest to stimuli within the region and steadily declines as stimuli are placed farther away (Graziano, Hu & Gross 1997b); and second, the depth of the region progressively expands as the speed of stimuli towards the body part increases (Fogassi et al.,

Fig. 2. Different types of tactile and visual receptive fields (RFs) of bimodal F4 neurons. Shadowed areas: tactile RFs. Solids around different body parts: visual RFs. Numbers refer to the particular neurons recorded from. (Reproduced with permission from Fogassi et al., 1996, p. 144).
1996). This evidence for a fuzzy boundary between representations of near and far space is consistent with Cowey et al., 1999 studies of human subjects with proximal/distal neglect. Finally, it is important to note that, just as with area 7b, lesions in F4 cause the animal to neglect near but not far space (Rizzolatti et al., 1983). Although it is not clear exactly how the representations of near space in areas 7b and F4 differ, it is likely that the representations in 7b have a predominantly visual function whereas the ones in F4 have a predominantly motoric function (Graziano et al., 1997a; Rizzolatti et al., 1997).

Like the reaching circuit, the grasping circuit is composed of a parietofrontal pathway. Area AIP (anterior intraparietal sulcus) contains neurons that are activated during grasping actions as well as during passive fixation of objects; moreover, these neurons are sensitive to the intrinsic three-dimensional properties of objects such as their size and orientation — properties that are relevant to determining the way in which the hand and fingers should be configured to interact with an object (Sakata, Taira, Mine & Murata, 1992; Sakata, Taira, Murata & Mine, 1995). In addition, pharmacological inactivation of this area causes errors in visually guided grasping movements, such as inaccurate preshaping of the hand during the approach phase (Gallese, Murata, Kaseda, Niki & Sakata, 1994). A similar disorder called optic ataxia is found in many human subjects with lesions in the superior portion of the parietal lobe (Perenin & Vighetto, 1988; Perenin, 1997). AIP projects to area F5 in the ventral premotor cortex of the frontal lobe, which is where the programming of hand movements takes place (Matelli, Camarda, Glickstein & Rizzolatti 1986). The firing properties of neurons in this area are selective for different aspects of grasping behaviors — for example, different kinds of object manipulation and types of grip (Rizzolatti et al., 1986).

The third circuit in the system for representing near space seems to be involved primarily in monitoring arm and hand movements toward and away from the face, especially the mouth; hence this circuit may play an important role in eating behavior. As with the other two circuits, it consists of a parietofrontal pathway. Area VIP (ventral intraparietal sulcus) receives input from PO as well as from several highly motion-sensitive areas (specifically MT and MST), and it contains neurons that are driven largely by stimuli moving at certain speeds and in certain directions within the peripersonal environment (Colby & Duhamel, 1991; Colby & Duhamel, 1996; Colby, Duhamel & Goldberg 1993; Duhamel, Bremmer, Benhamed & Graf 1997). Three different type of neurons are of interest with regard to spatial representation. First, “trajectory” neurons respond selectively to stimuli, especially the arm or hand, that are moving toward or away from the face. For approaching stimuli, these neurons may encode the anticipated point of contact. The second class of neurons are called “ultranear” because they only respond to stimuli that are within approximately 5 cm of the face. Their function may be to register the presence of stimuli that can be reached by the mouth. Neurons belonging to the third class also respond to stimuli that are in close proximity to the face, but not as close as for the ultranear neurons — e.g. 20–100 cm. VIP projects to areas F4 (arm region) and F5 (hand and mouth regions) in the premotor cortex, as well as to the portion of area 8 (frontal eye fields) that is involved in programming eye movements for smooth
pursuit tracking. Thus, this circuit appears to be useful for arm, hand, mouth, and eye coordination in near space.

With regard to the neural organization of the system for representing far space, fewer studies have been conducted and hence less is known. Nevertheless, some of the features of this system are relatively clear. Like the near system, the far system is well-adapted for using visual information to guide motor behavior; however, the far system differs from the near system insofar as it is not designed for controlling arm, hand, and head movements directed toward objects in the peripersonal environment, but is instead designed for controlling saccadic eye movements directed mainly toward objects in the extrapersonal field of view. This system is composed of brain areas in both the parietal and the frontal lobe. Areas 7a and LIP (lateral intraparietal sulcus) contain neurons that respond primarily to visual inputs, including inputs from PO (Colby & Duhamel, 1991). There is evidence that these neurons represent the locations of objects in an oculocentric frame of reference; in other words, a particular location is coded in terms of the distance and direction of the saccadic eye movement that would be necessary to fixate on that location (Colby & Duhamel, 1991, 1996; Colby Duhamel & Goldberg 1996). Neurons fire immediately prior to saccades, and also fire continuously when the animal is holding in working memory a representation of a specific location to which a saccade should be directed after a signal is given. In addition, the response properties of these neurons are heavily modulated by attention, leading Colby and Duhamel (1996), p. 107) to state that “in essence, the activity of an LIP neuron encodes an attended spatial location.” Areas 7a and LIP project to the portion of area 8 (frontal eye fields) that is involved in programming voluntary saccadic eye movements (Bruce & Goldberg, 1985). Remarkably, lesions in this area induce neglect that is much more severe for stimuli in far space, i.e. beyond arm’s reach, than for stimuli in near space, i.e. within arm’s reach — exactly the opposite of lesions in areas 7b and F4 in the near space system (Rizzolatti et al., 1983). These findings therefore support the notion that the parietofrontal circuit extending from areas 7a and LIP to area 8 is designed primarily for visual search and object scanning in the extrapersonal visual field.

Stepping back now from the details of the neurophysiological data, several general conclusions can be drawn about the functional-anatomical organization of the systems for representing near and far space. Both systems are evolutionarily designed for the visual guidance of motor control and are neurally implemented in parietofrontal circuits. But while the near system is geared toward the representation of peripersonal space in limb-centered and head-centered coordinates, the far system is geared more toward the representation of extrapersonal space in gaze-centered coordinates. In other words, the near system represents space for purposes of manipulating or avoiding objects with the body, whereas the far system represents space for purposes of “acquiring” and analyzing objects with the eyes. Because these two kinds of action involve very different kinematic parameters, distinct circuits are needed to carry out the required computations. Although a number of specific cortical areas have been identified in the macaque brain as participating in these circuits, it is likely that the architectures of the circuits are somewhat different in the human brain due to evolutionary modifications. Still, the fundamental division of
labor between near and far systems appears to hold for humans, as shown by the
neuropsychological studies reviewed earlier.

3. The linguistic representation of near and far space

Given the arguments presented above for distinct perceptual representations of
near and far space demarcated roughly by the extent of arm’s reach, the question
arises as to whether that distinction provides the basis for the way natural languages
use demonstrative terms, especially proximal and distal terms (e.g. this vs. that, here
vs. there), to divide space into different egocentrically grounded regions. The aim of
this section is to show that although there is a superficial resemblance between the
perceptual and linguistic distinctions, in fact they do not correspond to one another.
Cross-linguistic studies indicate that the demonstrative systems of many languages
do not make a simple binary proximal/distal distinction. In addition, and more
importantly, there is abundant evidence that the referential scope of proximal and
distal demonstratives is not restricted by the boundaries of immediate motor beha-
vor; rather, these terms encode an abstract language-specific semantic distinction
that can be used to express a potentially unlimited range of spatial distance contrasts
by virtue of being modulated by particular pragmatic contexts.

3.1. The linguistic properties of demonstratives

Demonstratives belong to the class of so-called ‘‘deictic’’ expressions. Generally
speaking, deictic expressions are linguistic elements whose interpretation makes
essential reference to some aspect of the speech situation (Lyons, 1977; Jarvella
& Klein, 1982; Weissenborn & Klein, 1982; Levinson, 1983; Rauh, 1983; Fuchs,
1993; Fillmore, 1997). Different types of deictic expressions vary with regard to
which aspect of the speech situation serves as the anchor, or deictic center, for their
interpretation, such as the identities of the interlocutors (e.g. I vs. you), the time at
which the discourse takes place (e.g. today vs. yesterday), or the spatial locations of
entities or events in the surrounding context (e.g. this vs. that; here vs. there).

The focus in this paper is on spatial demonstratives, which are traditionally
viewed as deictic elements that function as special types of locating expressions.
Basically, a locating expression designates a region of space within which some
entity, the ‘‘figure,’’ can be found, and it does this by making reference to a land-
mark object, the ‘‘ground’’, whose location is presupposed knowledge for both the
speaker and the addressee. For example, in the sentence Your hat is in the closet, the
speaker specifies the location of the hat as being within the region of space delimited
by the boundaries of the closet; moreover, the speaker assumes that the addressee
knows where the closet is. The unique feature of locating expressions that contain
spatial demonstratives is that the ground is not some entity outside the speech context
but rather the speech participants themselves, or more precisely, their location in
space.

Before getting into a detailed discussion of the semantic and pragmatic properties
of spatial demonstratives, it is important for me to describe briefly some of the
syntactic variation that is found with these elements. In what follows, reference will be made to three different syntactic categories of demonstratives: (i) adjectival demonstratives, which modify nouns and sometimes have a determiner-like function (e.g. Look at that horse!); (ii) pronominal demonstratives, which are used as arguments of verbs and other predicates (e.g. Did you see that?); and (iii) adverbial demonstratives, which modify verbs (e.g. Martha drove here). Although adjectival and pronominal demonstratives are phonologically identical in English, they are different in many other languages around the world. Furthermore, there are additional syntactic categories of demonstratives, but since they are not referred to in the following discussion of semantic properties, I will not describe them here.

3.2. Types of demonstrative systems across languages

All languages have spatial demonstratives, but there is variation in the number of distinctions that languages make with respect to the degree of remoteness of entities from the deictic center. All of the major types of demonstrative systems are summarized below; the categories and data are taken from several different crosslinguistic studies of demonstratives (Fillmore, 1982; Anderson & Keenan, 1985; Diessel, 1999, forthcoming). In addition, I report the relative proportion of languages that have each type of system; this information is based on a worldwide sample of 80 languages (Diessel 1999, forthcoming). The purpose of this survey is to make the following points. Many languages encode a distinction between two zones of space, one traditionally referred to as proximal (e.g. this, here) and the other as distal (e.g. that, there). This binary distinction appears to be similar to the perceptual distinction between near or peripersonal space on the one hand, and far or extrapersonal space on the other. However, a large proportion of languages have demonstrative systems that do not make such a two-way distinction but instead divide space into three or more egocentrically grounded sectors. Because these languages violate the near/far perceptual contrast, they constitute an “existence proof” that demonstrative systems need not correspond to it.

In principle, it is possible for a language to have a demonstrative system that does not make any distance distinctions whatsoever; however, no such language has ever been reported. Nevertheless, it is important to note that a very small proportion of languages do have distance-neutral demonstrative terms that are restricted to certain syntactic categories within the overall system, specifically the categories of adjec-

---

2 Some languages have demonstrative systems that are extremely complex due to the fact that they make distinctions not only along the dimension of degree of remoteness from the speaker or addressee, but also along other deictic dimensions such as elevation (up/down), geography (uphill/downhill, upriver/downriver, north coast/south coast), and movement (toward speaker/away from speaker/laterally across visual field of speaker). These additional dimensions will not be discussed here, however, since the focus is on just the dimension of degree of remoteness.

3 The languages in this sample come from many different genetic groups and represent every major linguistic area in the world. The counts for the number of languages with each type of demonstrative system were not done by Diessel but rather by myself; hence I am solely responsible for any counting errors that may exist. The counts are based on the information that Diessel obtained about the languages from technical grammars.
tive and pronoun. Seven of the 80 languages in Diessel’s sample may be like this: French, Czech, German, Alamblak (Sepik–Ramu, New Guinea), Sango (Niger–Kordofanian, Africa), Supyire (Niger–Kordofanian, Africa), and Tok Pisin (creole, New Guinea). All of these languages appear to have distance-neutral adjectival/pronominal demonstratives that are commonly used. In several of the languages, though, the neutral terms are supplemented by other adjectival/pronominal demonstratives that are available to make a proximal/distal contrast. For example, Czech has a term, ten, that is used to designate objects regardless of their distance from the speaker; however, there are two other terms, teno and onen, that indicate a proximal/distal distinction but simply aren’t used very often (Meyerstein, 1972). Alamblak also has a distance-neutral term, but it can be modified by deictic suffixes that encode a proximal/distal distinction and simply aren’t used very often (Meyerstein, 1972). Alamblak also has a distance-neutral term, but it can be modified by deictic suffixes that encode a proximal/distal distinction, as shown in Table 1 (data from Bruce, 1984, p. 81). Anderson and Keenan (1985, p. 280) point out that distance-neutral demonstratives like those in Czech and Alamblak mean simply “present to the speaker” or “present in the extralinguistic context of the utterance” hence they are not really very different from a definite article. This may explain why so few languages contain distance-neutral demonstratives, why such terms are restricted to certain syntactic contexts (adjective and pronoun), and why they are often supplemented by other adjectival/pronominal forms that encode at least a proximal/distal contrast. In all 80 languages in Diessel’s sample, the adverbial component of the demonstrative system always makes some kind of distance distinction, which suggests that distance marking is after all a necessary part of any complete demonstrative system.

A two-term system includes both a proximal and a distal form, with the speaker as the reference point. This kind of system is illustrated by the standard English distinctions between this/that (adjectival/pronominal) and here/there (adverbial). Another example is Vietnamese which, like English, expresses the proximal/distal contrast in word roots: dây/nây (adjectival/pronominal; diacritics indicate tones) and day/ho (adverbial) (data from Thompson 1987, p. 142). Two-term systems are quite frequent in Diessel’s sample, being present in 44 of the 80 languages. It is also intriguing that a sound-symbolic correlation has been found between proximal forms and high front vowels on the one hand, and between distal forms and low back vowels on the other (Woodworth, 1991).

---

4 Here and in what follows, the genetic family and the general geographical location of languages that may not be familiar to the reader are provided in parentheses.
Several different kinds of three-term systems have been described. In all of them, the first term is a proximal demonstrative with the speaker as the reference point; the systems vary in how the other two terms are used. In a person-oriented system, the second term is also a proximal demonstrative but has the addressee as the reference point, and the third term is a distal demonstrative that has both the speaker and the addressee as the reference point. Japanese has a system like this: *kono* means “proximal to speaker,” *sono* means “proximal to addressee,” and *ano* means “distal to both” (these are all adjectival forms; data from Anderson & Keenan, 1985, p. 284). Another language with a person-oriented three-term system is Pangasinan (Austronesian, South Pacific), shown in Table 2 (data from Benton, 1971, p. 88). It is interesting to note that a person-oriented three-term system is like a two-term system insofar as both make a binary distance contrast. The only difference is that in the former type of system the middle term requires that the speaker determine whether an object is relatively close to the addressee’s location, rather than to his or her own location. This type of system is not very well-attested in Diessel’s sample, however, being present in only 10 of the 80 languages.

There are two other types of three-term systems, both of which are distance-oriented and use the speaker as the deictic center. In one type, the first and third terms are usually called proximal and distal demonstratives, and the second is usually called a medial demonstrative because it is prototypically used to refer to entities that are at a distance intermediate between proximal and distal. One well-known language with a system like this is Spanish: *este* means “proximal to speaker,” *ese* means “medial to speaker,” and *aquel* means “distal to speaker” (these are all masculine singular adjectival/pronominal forms; data from Anderson & Keenan, 1985, p. 282). Another such language is Yimas (Sepik–Ramu, New Guinea), which is shown in Table 3 (data from Foley, 1991, p. 112). This is the second most common type of demonstrative system in Diessel’s sample, appearing in 22 of the 80 languages. In the other type of distance-oriented three-term system,

---

**Table 2**
Adjectival/pronominal demonstratives in Pangasinan. The initial /i/ is often dropped, especially if the preceding word ends in a vowel

<table>
<thead>
<tr>
<th></th>
<th>Proximal-speaker</th>
<th>Proximal-addressee</th>
<th>Distal-S/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>(i)-yá</td>
<td>(i)-tán</td>
<td>(i)-mán</td>
</tr>
<tr>
<td>Plural</td>
<td>(i)-rā-ya</td>
<td>(i)-rā-tan</td>
<td>(i)-rā-man</td>
</tr>
</tbody>
</table>

**Table 3**
Adjectival/pronominal demonstratives in Yimas. The proximal and distal roots, *k* and *n*, take number prefixes; the medial root, *m*, takes number suffixes

<table>
<thead>
<tr>
<th></th>
<th>Proximal</th>
<th>Medial</th>
<th>Distal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>p-k</td>
<td>m-n</td>
<td>p-n</td>
</tr>
<tr>
<td>Dual</td>
<td>pla-k</td>
<td>m-pl</td>
<td>pla-n</td>
</tr>
<tr>
<td>Plural</td>
<td>pla-k</td>
<td>m-ra</td>
<td>pia-n</td>
</tr>
</tbody>
</table>
the first term is a proximal demonstrative and the other two terms divide the distal field into separate regions — remote and very remote. The English term *yonder*, which was once used much more frequently than it is today, illustrates the “very remote” category. In most languages with this kind of three-way contrast, however, it is difficult to determine whether the third term simply means “very remote” or if it also, or alternatively, means “invisible.” For example, Fillmore (1982) quotes Dixon (1977, p. 180) as saying that in Yidiny (Pama–Nyungan, Australia), some dialects use the third term to refer to entities that are very far from the speaker, whereas others use it to refer to entities that are outside the speaker’s visual field. Partially because of this indeterminacy, this type of demonstrative system is not well-represented in Diessel’s sample, being present in only 2 of the 80 languages. The ambiguity and rarity of this type of system are not important, though, because as mentioned above, the first type of distance-oriented three-term system is quite common. The significance of this lies in the fact that by clearly distinguishing between three regions of space, the first type of system deviates from the simple two-way perceptual distinction between near and far regions.

Although the number of spatial distinctions made by demonstrative systems appears to be almost always just two or three, there are a few reports of languages that make further distinctions along the dimension of degree of remoteness from the speaker (or addressee). One such language is Tlingit (Na–Dene, North America), which has the following four-way contrast for adjectival/pronominal demonstratives: *yáa* means “this (one) right here” and is thus a proximal term; *hél* means “this (one) nearby” and thus appears to be a medial term; *wéé* means “that (one) over there” and is thus a distal term; and *yóó* means “that (one) far off” and thus appears to be a term for designating very remote entities (data from Story & Naish 1973; Anderson & Keenan, 1985, p. 286). A five-term adjectival/pronominal demonstrative system that includes both distance and speaker/addressee dimensions is found in CiBemba (Bantu, Africa): *ú-nó* means “immediately adjacent to or on the speaker,” *ú-yá* means “nearer the speaker than the addressee,” *ú-yót* means “equally near or relevant to both,” *ú-yló* means “immediately adjacent to or on the addressee,” and *ú-lyá* means “away from both” (diacritics indicate tones; data from Welmers 1973, p. 286; Anderson & Keenan, 1985, p. 288). The richest distance-oriented system that has ever been described is in Malagasy (Western Austronesian, Madagascar), a language that is reputed to have seven different terms for varying degrees of remoteness from the speaker. In order of increasing distance, the terms are as follows: *ety, eto, eo, ety, eny, eroa,* and *ery* (data from Anderson & Keenan, 1985, p. 292). It is not surprising, however, that the semantic status of these terms is controversial (Fillmore, 1982).

In summary, crosslinguistic research indicates that although two-term demonstrative systems appear to be the most common, several other types of systems also exist, and some of them — in particular, distance-oriented three-term systems in which the third term has a medial function — are widely attested. This typological variation clearly shows that the way languages carve space into different egocentrically grounded regions can diverge significantly from the perceptual distinction between near and far sectors.
Now, it might be argued that despite the extent of crosslinguistic variation, the fact that two-term systems are the most frequent in Diessel’s sample suggests that this kind of system can still be considered “basic” and might even be related in some sense to the binary perceptual distinction. An argument along these lines must be evaluated carefully, however. First of all, although it is likely that two-term systems are “basic,” there are many reasons for being cautious about drawing any definitive conclusions from the existing data. Diessel’s sample is large and well-selected in comparison to many other studies in linguistic typology, but it nevertheless has several limitations. It consists of a very small proportion (1.6%) of the roughly 5000 languages in the world, and as with any sample of this size, it over-represents some language groups while under-representing others (Whaley, 1997). In addition, reference grammars, which are the source of Diessel’s data, do not always provide detailed, reliable information about demonstratives, in part because fieldworkers often find it very difficult to determine the nuances of this aspect of the language under investigation.\(^5\) Given these limitations, one can only use Diessel’s sample to make tentative, provisional generalizations about the relative proportions of different types of demonstrative systems in the languages of the world. Another reason for being cautious about regarding two-term systems as “basic” is that such systems are often historically derived from three-term systems by reduction of the middle term (Frei, 1944). Finally, there is evidence that languages spoken in small-scale nonindustrial societies tend to have more numerous and more grammaticalized spatial deictic distinctions than languages spoken in complex industrial societies (Perkins, 1992; see also Denny, 1978; Denny, 1988; Senft, 1997). If this is correct, it raises questions about the proper sociolinguistic context for determining what qualifies as a “basic” demonstrative system.

As for the notion that two-term systems might be related to the perceptual near/far distinction, while there may be some truth to it (see the conclusion), there are several reasons for being skeptical. Most importantly, semantic and pragmatic analyses of how proximal and distal demonstratives are used in conversational discourse indicate that these terms have very broad functions which cannot be reduced to the restricted perceptual distinction between near and far regions of egocentric space. The purpose of the next section is to elaborate this point.

### 3.3. The semantic abstractness and pragmatic flexibility of demonstratives

As noted above, it seems natural to suppose that the crosslinguistically common distinction between proximal and distal demonstratives might correspond to the well-established perceptual distinction between near and far space, with proximal forms referring to entities that are roughly within reach of the speaker and distal forms referring to entities that are further away. Evidence is presented below, however, which shows that this is a highly oversimplified characterization of the forms. I will argue that the inherent meanings of proximal and distal demonstratives are not defined in terms of the range of potential arm, hand, and head movements

\(^5\) Holger Diessel, personal communication.
and hence do not correspond to the perceptual distinction between near and far space. The specifications of “degree of remoteness from the deictic center” that are semantically encoded by these forms are not concrete and quantitatively restricted but are instead abstract and quantitatively variable. In any given utterance, the actual boundaries of the region of space designated by a demonstrative are determined by a combination of the demonstrative’s abstract semantic structure and the unique pragmatic conditions of the speech situation.

The most straightforward way to begin elaborating this argument is by presenting some data from English. Consider the following pair of sentences:

1. This speck is smaller than that speck.
2. This planet is smaller than that planet.

Talmy (1988, pp. 168-9) uses these sentences to illustrate how the proximal/distal contrast specified by the demonstratives this and that is not absolute but relativistic: “The scenes referred to by these two sentences differ greatly, involving tiny objects millimeters apart to huge objects parsecs apart; the scenes’ differences as to the magnitude of size or distance must arise from the lexical elements, they cannot be traced to the deictics…” What this implies is that the notions of “proximal” and “distal” that are semantically encoded by this and that cannot be considered spatial in the strict sense of the term because they do not specify objectively measurable distances; rather, they are technically neutral with regard to actual spatial extent. It is therefore appropriate to treat the notions as abstract semantic features. Together, they constitute a closed language-internal Saussurean system in which each notion is defined not with reference to real concrete space but instead in opposition to the other notion (Saussure, 1916; Ruhl, 1989). Thus the proximal demonstrative this means simply “closer to the deictic center than that,” and conversely the distal demonstrative that means simply “further from the deictic center than this.” When the forms are used in utterances, as in the examples above, their abstract, spatially relativistic semantic values are modulated and made specific by the discourse context and/or the situational context. Because of this pragmatic flexibility, the proximal/distal contrast can be used to designate any kind of spatial remoteness distinction that the speaker wishes to express.

Consider a few more examples, this time involving the adverbial demonstratives here and there. Levinson (1983, p. 80) points out that the statement Place it here “may have quite different implications of precision if said to a crane operator or a fellow surgeon.” In a similar vein, Klein (1982, p. 166) observes that “it is possible to say Here comes my mother when she is at a distance of 100 m, but one can also say There’s my mother when she is at a distance of 10 m.” These examples illustrate quite clearly that the referential function of demonstratives is highly subjective and context-dependent, as opposed to being demarcated by a bodily-based boundary such as within vs. beyond the extent of arm’s reach. As with the pronominal/adjectival forms this and that, the adverbial forms here and there semantically encode an abstract proximal/distal contrast that has the potential to capture an unlimited range of physical distance distinctions by virtue of being combined with the particularities of specific pragmatic contexts.
In a study that focuses on German proximal and distal demonstratives; Fuchs (1993) provides a large amount of data which shows that in using these terms the speaker is not constrained by objective spatial factors but is instead free to express whatever type of distance distinction is relevant to the current social interaction; moreover, the type of distinction that is expressed often depends on extensive shared sociocultural knowledge between the interlocutors in order to be interpreted correctly. For instance, she reports the following exchange:

1. A and B, who live in Göttingen, are on their way to a restaurant outside town — in fact, they have just left the town properly speaking. Over the radio, they hear a song by Purcell:
2. A: das is das Stafford-Ensemble. (that’s the Stafford ensemble.)
3. B: ach — die gerade hier gesungen haben! (oh, the people who just sang here!)

If hier were interpreted in an overly concrete manner, it would refer to the car in which the interlocutors are traveling, but of course that would be absurd. The discourse context indicates that the demonstrative is actually being used to contrast Göttingen, a semiprovincial town which is not often treated to performances of renowned musicians, with more prestigious cities such as European capitols, or perhaps with England, the home country of the ensemble. Note that the demonstrative can be used felicitously in this way even though the interlocutors are not even in Göttingen at the time of speaking. This is possible because the interlocutors share a substantial amount of unconscious background knowledge not only about the topic of conversation but also about the linguistic conventions for using demonstratives. As Fuchs points out, “The function of here is to indicate the [pragmatically] ‘given’ position on any dimension of localization that may be relevant at the moment. The place where we are speaking and, a fortiori, literal physical location is not too frequently relevant to what we are saying’” (p. 18; emphasis in original). Indeed, although one might suppose that proximal vs. distal demonstratives are at least prototypically used to refer to entities and events within vs. beyond the speaker’s immediate peripersonal sphere (Kirsner, 1993), this has never been confirmed by empirical discourse analyses, and studies like Fuchs’s suggest that the assumption may not be true.

The idea that demonstratives are used more often to distinguish between socially determined locations than between purely spatially determined locations is also emphasized by Hanks (1990), who conducted a very detailed investigation of how the demonstratives in Maya (Mexico) are used in discourse. Maya has a fairly complex demonstrative system, and only one part of it is discussed here — specifically, a binary distinction that is very much like a proximal/distal distinction but which Hanks calls inclusive/exclusive instead. The most basic function of the inclusive form way e?, glossed as “here,” is to designate the peripersonal sphere of the speaker; hence the meaning of the term appears, at least superficially, to correspond

---

6 Actually, German has a three-term system that includes two distal forms, da and dort, the precise meanings of which are difficult to specify (Himmelmann, 1997). This distinction is not relevant to the present discussion, however.
to the perceptual sector of near space. However, as with the English and German proximal demonstratives discussed above, the referential range of the term is actually quite broad, which suggests that the notion of ‘inclusive’ that it encodes is abstract and neutral with respect to spatial extent, being modulated and particularized by specific pragmatic contexts. Hanks summarizes the usage conventions of the term as follows (p. 406):

1. the most typical referents individuated with this form include the interactants’ -i’iknal ‘bodily space’ (including the perceptual and kinetic field of corporeal engagements), the space of a single walled room, the first-level soólar space of a single marriage pair with unmarried offspring, the second-level, walled soólar as a whole, the agricultural plots or orchards worked by a single man or ‘team’, the region frequented by the interactants, and the earth inhabited by man, as opposed to the sky inhabited by gods.

The counterpart of way e’? is the exclusive form tol o,’ which is glossed as ‘out there.’ This term is defined in opposition to the inclusive form and can be used to refer to entities that lie beyond an abstract perimeter that freely varies along the dimension of distance from the speaker. Almost any objective situation can be polarized into inclusive and exclusive sectors, regardless of the actual spatial extent of each sector, although most of the situations in which this occurs involve a variety of socioculturally determined ‘spaces’ of different physical sizes such as the ones described in the passage quoted above.

So far the discussion has focused on the two-term demonstrative systems in English, German, and Maya and has shown that the distance contrasts encoded by these systems do not correspond to the perceptual distinction between near and far space but are instead quantitatively variable, depending on the details of particular speech contexts to specify concrete spatial relationships. There is evidence that the same combination of semantic abstractness and pragmatic flexibility is also present in languages with three-term distance-oriented demonstrative systems, such as Spanish. Hottenroth (1982) points out that the Spanish proximal, medial, and distal forms — este, ese, and aquel, respectively (masculine singular pronominal/adjectival forms) — designate increasingly remote concentric circles around the speaker. However, the meanings of the demonstratives ‘can be defined neither positively nor negatively with respect to an objective region of proximity’ (p. 142). Rather, each term encodes an abstract notion that is defined solely in opposition to the other terms such that they constitute a closed semantic system. As Hottenroth puts it, ‘the extension of each region referred to by some demonstrative in a given context is no more than a variable in an abstract meaning formula’ (p. 142). Thus the speaker can willfully expand or contract the concentric circles encoded by the terms to suit the pragmatic demands of the discourse situation. For instance, the proximal terms aquí (adverbial) and este (pronominal/adjectival) can be used with equal felicity to refer to small areas around the speaker or to areas far beyond the objectively near region (p. 139):

1. aquí, en esta habitación (here, in this room)
2. aquí, en esta ciudad (here, in this town)
3. aquí este país (here, in this country)
4. aquí, en este mundo (here, in this world)

Similarly, all three terms — proximal, medial, and distal — can be used by a hospitalized patient pointing to different places on his or her own body to show the doctor where pain is felt. This is possible even though, from an objective point of view, all of the locations are literally on the speaker’s body and therefore within the region of maximal proximity. These examples illustrate that, just as in languages with two-term demonstrative systems, the Spanish three-term system is not based on more or less absolute distance distinctions but is instead highly relativistic and context-sensitive. Text analyses have also revealed that the Latin three-term system, which is the historical source of the Spanish system, is also abstract rather than concrete (Keller, 1946).

4. Conclusion

The purpose of this paper has been to argue that natural languages and the visual system differ significantly in the way they carve up space into separate egocentrically grounded regions. Evidence from both human neuropsychology and primate neurophysiology was presented which suggests that the visual system makes a fundamental distinction between a near or peripersonal region of space that extends roughly to the perimeter of arm’s reach — the realm in which most manual activity takes place — and a far or extrapersonal region of space that expands outward from that boundary — the realm in which visual search and object scanning and recognition usually take place. Then crosslinguistic research was reviewed which indicates that it is very common for languages to have two-term demonstrative systems that also make a binary distinction between proximal and distal regions of space. However, several reasons were given for not treating this linguistic distinction as corresponding to or being based on the perceptual distinction. First, a large number of languages have demonstrative systems that divide the spatial world into three or more separate regions, thereby violating the two-way perceptual contrast. Even more importantly, demonstrative terms simply do not encode quantitative spatial information such as within vs. beyond arm’s reach; instead, they specify abstract semantic notions that, when combined with the unique features of communicative contexts, allow speakers to make a virtually unlimited range of spatial distance contrasts. The way a speaker uses demonstratives actually reveals more about his or her subjective attitude toward the surrounding spatial world than about the objective features of that world. This is because it is the speaker who determines how the oppositional relations that are abstractly encoded by the terms are applied to distinguish between near and far sectors of space. As Hottenroth (1982) points out, objective distances are variables in the meanings of demonstratives; these variables are only given fixed values by specific pragmatic factors in the linguistic and extralinguistic context of the utterance.

In contrast, then, to the analysis of concrete nouns and locative prepositions by Landau and Jackendoff (1993), this analysis of demonstratives suggests that they
constitute an interesting case of divergence between linguistic and perceptual representations of space. A bit of reflection shows, however, that this divergence is not really so surprising. After all, demonstratives appear to straddle the boundaries between visual perception, abstract semantic organization, and context-specific sociolinguistic interaction. It therefore makes sense that they should have connections with all three of these domains without being completely reducible to any single one of them. The typological prevalence of two-term demonstrative systems that encode a binary proximal/distal distinction may be due to some form of influence from the perceptual organization of near and far space (see below); the abstract, quantitatively neutral nature of the notions “proximal” and “distal” is due to their being semantic representations rather than concrete spatial representations; and the remarkable pragmatic flexibility of demonstratives is due to the fact that they are essentially deictic terms that cannot function apart from specific discourse contexts. As Levinson (1983, p. 54) says, ‘‘The properties of deixis should act as a constant reminder to theoretical linguists of the simple but immensely important fact that natural languages are primarily designed, so to speak, for use in face-to-face interaction, and thus there are limits to the extent to which they can be analyzed without taking this into account.’’

The idea that linguistic and perceptual representations of near and far space are different is reinforced by the fact that these two types of representations serve fundamentally different goals. On the one hand, perceptual representations of egocentric space appear to be designed primarily for motor control, as argued in Section 2. In particular, representations of near or peripersonal space are used to program movements of the arms, hands, and head either towards or away from objects, and representations of far or extrapersonal space are used to program saccadic eye movements either for fixating on objects or for visually examining their internal structure. In order to carry out these functions efficiently, the perceptual representations must capture fairly precise metric distance information, and they must also have their frame of reference anchored in the appropriate body part. On the other hand, the representations of egocentric space that are encoded by demonstratives appear to be designed for a very different purpose, namely communicating a virtually unlimited range of possible distance contrasts that are determined by the unique pragmatic features of particular speech situations. Demonstratives are closed-class items, and like other closed-class items (e.g. prepositions, case-markers, determiners, auxiliaries, etc.) they are used very frequently and contribute to the assembly of a skeletal semantic framework which is ‘‘fleshed out,’’ so to speak, by the richer, more idiosyncratic conceptual properties of open-class items such as nouns, verbs, and adjectives (Talmy, 1988; Slobin, 1997). In order to serve this function properly, the semantic content of demonstratives must be general enough to be applicable across an extremely wide spectrum of linguistic contexts. Thus, unlike the metrically precise spatial representations that are used in perception, the linguistic notions of ‘‘proximal’’ and ‘‘distal’’ that are encoded by demonstratives are necessarily very abstract. Of course, it is possible for speakers to communicate quantitatively specific distance information if they wish, but this requires the employment of open-class measure expressions such as 32.4 cm. It
should be clear, then, that functional considerations support the view that linguistic and perceptual representations of near and far space are essentially different. If the meanings of demonstratives were based directly on egocentrically-grounded perceptual representations, there would have to be so many terms, each with a fairly precise distance specification, that they would be too unwieldy to use; indeed, this would defeat the whole purpose of the closed-class lexicon, which is to provide a small inventory of semantically schematic terms that can be processed rapidly and modulated to fit a variety of contexts. Conversely, if perceptual representations of near and far space had the same kind of abstract, quantitatively neutral content as demonstratives, they would not be able to guide accurate body movements, and this would violate the adaptive principles of the sensorimotor system.

Despite all of these considerations, however, it is still quite remarkable that the available evidence suggests that two-term demonstrative systems are the most common in the world. As noted in Section 3.2, the crosslinguistic prevalence of this type of system remains to be confirmed by further research, but if it is confirmed, as seems likely, the question naturally arises as to why two-term systems are so frequent. After all, it is logically possible for five- or seven-term systems, as in CiBemba and Malagasy, to be the most widespread. It seems reasonable to speculate that there may actually be some kind of deep connection between the linguistic proximal/distal distinction and the perceptual near/far distinction.7 Comparative paleoneurological data indicates that around 3–4 million years ago a major reorganizational change from a pongid to a hominid brain took place which included expansion of the region at the junction of the parietal, occipital, and temporal lobes (POT) (Holloway, 1995). This region is famous for being dubbed “the association area of association areas” by Geschwind (1965). Perhaps one of its functions is to restructure spatial representations in such a way that their content becomes more abstract and hence more available for cognitive purposes, as opposed to being limited to sensorimotor control. The abstract representations could be manipulated for various cognitive tasks by means of executive networks in the dorsolateral and ventrolateral prefrontal cortices, which are known to have robust reciprocal connections with the POT region (Pandya & Yeterian, 1996). An important application of the abstract representations may be for language-specific semantic structure. With regard to demonstratives, it is conceivable that, as a legacy of its evolutionary origins, the POT region contains a predisposition for a binary proximal/distal distinction. If so, this could explain why the majority of languages make such a distinction in their demonstrative systems. It is worth noting that other domains of language-specific semantic structure which the POT region might subserve include locative prepositions (Damasio, Grabowski, Hichwa & Damasio 1998; Kemmerer & Tranel, 1999) and many properties of verb meaning, such as aspectual, thematic-role, and force-dynamic features (Wilkins & Wakefield, 1995).

More generally, it is conceivable that the POT region contributes significantly to the uniquely human tendency to use spatial representation as a framework or scaffolding for structuring and reasoning about other conceptual domains, such as time.

7 The following proposals were inspired in part by comments from an anonymous referee.
causation, possession, emotion, kinship, status hierarchies, and so forth (e.g. Pinker 1989; Pinker, 1997; Lakoff, 1990; Jackendoff & Aaron, 1991; Wierzbicka, 1994; Johnson-Laird, 1996).

Several testable predictions follow from the hypothesis that the abstract meanings of proximal and distal demonstratives are implemented in a different brain region than the metric spatial representations used by the near and far perceptual systems. First, a functional neuroimaging study should find that the major posterior “hot spot” for demonstratives is the POT region; moreover, this activation should be strongest in the left hemisphere, not just because the left hemisphere is dominant for language but also because it is dominant for categorical as opposed to metric representations (Kosslyn, 1994). On the other hand, different areas in the parietal, occipital, and temporal lobes should be activated during sensorimotor tasks that require the near and far perceptual systems, and these activations should be strongest in the right hemisphere since it is dominant for the kinds of metric representations used in sensorimotor control. Second, right-hemisphere brain-damaged subjects who neglect either near or far space should still be able to use demonstratives appropriately within the truncated spatial field that they consciously represent. For example, within their conscious field they should be able to voluntarily expand or contract the imaginary boundary between the referential zones of this and that, or here and there, just as easily as a normal person. This expectation comes from the assumption that the semantic structures encoded by proximal and distal demonstratives are mediated primarily by the left hemisphere POT region, which is still intact in these subjects.

In conclusion, this study has shown that the ways in which experience is structured for purposes of linguistic communication do not always reflect the ways in which it is structured for purposes of visual perception and motor control (see also Kemmerer & Tranel, 1999). In this regard, the study differs from other work that has focused instead on discovering points of convergence between linguistic and perceptual representations of space (e.g. Lakoff, 1987; Langacker, 1987; Langacker, 1991; Landau & Jackendoff, 1993; Regier, 1996). A major implication of the present paper is that research on both fronts is necessary if we are to achieve a complete understanding of the interface between language and perception.

Acknowledgements

I owe an enormous debt of thanks to Holger Diessel, who was willing to not only share his extensive data on demonstratives with me, but also engage in countless discussions on the topic and share with me his many insights. It is by no means an overstatement to say that without him I would never have been able to write this paper. Still, all errors that remain are of course my own responsibility. I would also like to thank Daniel Tranel for bringing the Cowey, Small and Ellis (1999) paper to my attention, and two anonymous referees for their comments on an earlier version of the manuscript.

Supported by Program Project Grant NINDS NS19632
References


