

Auditory primes and visual targets (translation in parentheses)

Auditory Primes			Visual Target	
Auditory Primes			Control	
Test				
schaap	(sheep)	blik	(tin)	HOND (dog)
bakker	(baker)	vrede	(peace)	LERAAAR (teacher)
melk	(milk)	storm	(storm)	SINAS (lemonade)
geweer	(rifle)	havik	(hawk)	KANON (canon)
spin	(spider)	krijt	(chalk)	TOR (beetle)
fluit	(flute)	knop	(knob)	GITAAR (guitar)
pistool	(gun)	klodder	(blotch)	ZWAARD (sword)
wesp	(wasp)	pracht	(splendour)	KEVER (beetle)
kerrie	(curry)	stuiver	(dime)	KANEEL (cinnamon)
zwaluw	(swallow)	praline	(chocolate)	MUS (sparrow)
papier	(paper)	vlegel	(lout)	GUM (gum)
tennis	(tennis)	drama	(drama)	ZWEMMEN (swimming)
vloot	(violin)	kristal	(crystal)	PIANO (piano)
broer	(brother)	rund	(cattle)	NEEF (cousin)
paard	(horse)	troep	(mess)	KAT (cat)
kapitein	(captain)	salade	(salad)	ADMIRAAL (admiral)
schoen	(shoe)	trots	(pride)	HEMD (shirt)
krant	(newspaper)	viam	(flame)	BOEK (book)
tijger	(tiger)	kruimel	(crumb)	OLIFANT (elephant)
dokter	(doctor)	komeet	(comet)	SLAGER (butcher)
banaan	(banana)	pilaar	(pillar)	PEER (pear)
hommel	(bumblebee)	twijfel	(doubt)	VLINDER (butterfly)
ijzer	(iron)	patroon	(pattern)	BRONS (bronze)
sherry	(sherry)	mantel	(coat)	BIER (beer)
koffie	(coffee)	magneet	(magnet)	WATER (water)
spijker	(nail)	reuma	(rheumatism)	ZAAG (saw)
tante	(aunt)	spiegel	(mirror)	OPA (grandad)
generaal	(general)	provincie	(province)	SOLDAAT (soldier)
moeder	(mother)	plakkaat	(poster)	NICHT (niece)
karper	(carp)	tiran	(tyrant)	HARING (herring)
vlieg	(fly)	span	(pair)	MIER (ant)
zilver	(silver)	peuter	(toddler)	KOPER (copper)
merel	(blackbird)	tomaat	(tomato)	KRAAI (crow)
potlood	(pencil)	orgaan	(organ)	BALPOINT (ballpoint)
stoel	(chair)	land	(country)	KAST (cupboard)
broek	(trousers)	ruit	(windowpane)	SOKKEN (socks)

Centering Attention in Discourse

Susan E. Brennan

Department of Psychology, State University of New York, Stony Brook, New York, USA

Centering is a set of processes by which people can coordinate their attention during conversation. Speakers signal where their attention is via particular lexical, syntactic and prosodic choices. These choices make discourse entities salient to addressees and may also signal whether speakers expect addressees' attention to be already centered on the same entities. Two questions arise: What linguistic devices do speakers use to make an entity salient in a discourse? And how do speakers re-refer to discourse entities that move in and out of the center of attention? I manipulated speakers' center of attention using a videotaped basketball game. Speakers tended to refer to prominent entities as sentence subjects. When they referred to entities as sentence *objects*, they were more likely to re-refer next by repeating the full noun phrase verbatim rather than pronominalising. This happened even though both a recency-of-mention strategy and world knowledge would have provided enough information to uniquely identify the referent of a pronoun. When speakers pronominalised entities that had *not* been centered syntactically, they tended to make them salient prosodically by lengthening the duration of those pronouns.

INTRODUCTION

During communication, people build models that help them keep track of what a discourse is about (Garnham & Oakhill, 1992; Sag & Hankamer, 1984; Webber, 1979). These models are constrained in both space and time by the limits on working memory (Johnson-Laird, 1983). That is, only a few

Requests for reprints should be addressed to Susan E. Brennan, Department of Psychology, State University of New York, Stony Brook, NY 11794-2500, USA.

I wish to thank Dwight Bolinger, Herbert Clark, Anne Fernald, Morton Ann Gernsbacher, Richard Gerrig, Steven Greene, Ellen Isaacs, Livia Polanyi, Carl Pollard, William Posner, Craig Roberts, Michael Schober, Peter Sells, Ivan Sag, Michael Tanenhaus, Marilyn Walker, Thomas Wasow and the anonymous reviewers for their valuable comments on this paper and the research it describes. This research was supported by a National Science Foundation Graduate Fellowship and by Grant No. BNS87-09812 from the National Science Foundation.

and drawing inferences about them takes cognitive effort. Potentially ambiguous referring expressions such as pronouns and definite references are resolved using such models (Johnson-Laird, 1983; Sag & Hankamer, 1984; Sidner, 1981). Consider a situation where one speaker, Alonza, is watching a basketball game and describing it to her friend Joy, who is sitting where she can't see the game:

Alonza: number forty-five o' the Spartans has the ball.
number ten is bringing it down, the Spartans.
and passes it off to ...
number forty-two ...
number forty-two passes it to forty-five ...
and forty-two goes up and it was fouled.
by number thirty-five of the Wolverines.
ok ...
Joy: is he takin a sh- foul shot?
Alonza: yup.
number forty's on the line.
Joy: forty?
Alonza: number forty of the Spartans.
oh, excuse me, forty-two.
ha hah ha ha.
and the first shot's good.
and he goes up for his shot ...
and it's good.

Although there are several potential referents for a third-person singular masculine pronoun ("forty-five", "ten", "forty-two" and "thirty-five"), Joy is able to use "he" to refer to player forty-two. When Alonza mis-speaks, calling that player "forty" instead of "forty-two", Joy gets her to repair the error. Then Alonza goes on to refer to player forty-two as "he". Joy and Alonza's ability to collaborate so closely in establishing the referents of these noun phrases makes it seem as if they have a shared mental workspace containing entities available for inspection and manipulation by them both.

What influences speakers' choices in conversation? Consider conversation as an opportunistic activity. People often begin to speak while they are still planning what they are going to say (Levelt, 1989). They have many options: they can mention an entity as a subject or an object of a sentence; they can use a full noun phrase or a pronoun; they can choose a word that biases the agency of an action towards the subject or towards the object of the sentence; they can feature a phrase by fronting it; they can subordinate a clause; they can stress a word or not stress it. Such linguistic

said, "every difference makes a difference".¹ Speakers' choices reflect their limited attention and the status of information in the discourse.

Information is typically marked in an utterance to distinguish given information from new (Chafe, 1976; Clark & Haviland, 1977; Halliday & Hasan, 1976; Haviland & Clark, 1974). The given information in an utterance helps an addressee know where to associate the new information with respect to a discourse model under construction. In English, given information is typically de-stressed and placed near the beginning of a sentence, while new information is typically stressed and placed near the end (Bock, 1977). There are several kinds of given and new information. Given information can be *known* to the speaker and addressee ("hearer-old" according to Prince, 1992), or *predictable* in the syntax of an utterance so that it is deletable (such as in conjunction reduction), or *salient*, where "the speaker assumes that the hearer has or could appropriately have some particular thing/entity/... in his/her consciousness at the time of hearing the utterance" (Prince, 1981). New information can be *brand-new* to the addressee, requiring him to create a novel mental entity for it ("hearer-new" according to Prince, 1992), or *inferrable* from known information (Prince, 1981; 1992), or finally, both "hearer-old" and "discourse-new" (Prince, 1992). This last category includes information that is unused in what Prince (1992) has called the "prior discourse stretch" but that is already known to the addressee when the speaker first introduces it. Within this category, information that has not been previously evoked in the discourse at all can be distinguished from information that *has* been evoked earlier but that is not now *salient* to the addressee (Prince, 1981).

Clark and Marshall (1981) have proposed that successful definite reference depends on the mutual knowledge of speakers and addressees. However, because conversation unfolds over time, only a small portion of what is mutually known can be salient at a particular moment and directly relevant to the interpretation of a pronoun. Successful pronominal reference requires that the focuses of attention of speaker and addressee be coordinated (Sidner, 1983), to maximise the likelihood that a referring expression specifies the same entity for both of them. Chafe's information-status account presumes that a particular speaker uses a pronoun because the pronoun's referent is already active in the addressee's model (Chafe, 1987). It is also possible that coordination between speakers and addressees comes about in a "generic" way, simply because the comprehension and production systems may have evolved so that what is easy for one is easy for the other (Dell & Brown, 1991; for discussion of cognitive constraints on production, see Bock, 1982). I will not distinguish these possibilities here,

¹D. Bolinger (pers. comm.); for discussion of this point, see Bolinger (1977).

speaking and hearing are closely adapted.

Centering is a theory of how a speaker signals her center of attention by her linguistic choices and how such choices help an addressee track changes in the focus of attention and interpret expressions in a discourse context. The centering theory originated in the field of artificial intelligence (Grosz, 1977; Grosz, Joshi, & Weinstein, 1983; 1986; in press; Grosz & Sidner, 1986; Joshi & Weinstein, 1981; Sidner, 1979; 1981; 1983). Centering integrates linguistic notions of topic and givenness, psychological concepts of salience and information status, and computational accounts of how discourse structure affects the interpretation of referring expressions. It fits within a mental models framework and provides specific mappings from linguistic form onto a discourse model. After surveying some of the details of centering theory, I will describe evidence for centering from text comprehension studies. Then I will present a production study that manipulates speakers' attentional state in a conversational setting to test some hypotheses derived from centering about syntactic and prosodic choices.

The Centering Theory

In a discourse, local structure can be distinguished from global structure (Grosz & Sidner, 1986; Kintsch & van Dijk, 1978). Grosz and Sidner model the global attentional structure of discourse using a stack that operates with the usual "last on, first off" principle; after the end of an interruption or a side sequence, speakers and listeners find themselves back at the point where they left off (Grosz & Sidner, 1986). What gets pushed onto and popped off the stack are global discourse segments or "focus spaces", corresponding to stretches of discourse segmented according to intentions—that is, each segment can be characterised by a single main purpose of the participants (Grosz & Sidner, 1986). A focus space is a data structure containing mentally represented entities that were mentioned explicitly or that are accessible implicitly (via inference) during a discourse segment. Focus spaces are assumed to be locally segmented by utterances that are sentences (e.g. Brennan, Friedman, & Pollard, 1987; Dahl, 1986; Grosz et al., 1983; 1986; in press; Kameyama, 1986) or sentential clauses (see Suri & McCoy, 1993, on the segmentation of complex sentences). Psychologically, the current discourse context corresponds to the contents of working memory available for interpreting or generating an utterance.

Centering is defined as the local focusing of attention within a global discourse segment or focus space (Grosz et al., 1983; 1986; in press; Joshi & Weinstein, 1981). The term "center" was proposed by Joshi and Weinstein (1981), presumably to escape the many associations—some of them quite incompatible—that are evoked in linguistics by the term "focus". The *center*

is the single most important discourse entity at a particular moment; in terms of information status, it is both discourse-old (Prince, 1992) and currently salient. It has more in common with Reinhart's notion of sentence topic than with her notion of focus (see Reinhart, 1982), although it is not quite the same as a topic; Reinhart's "topic" is a linguistic expression, whereas a "center" is an entity in a model.

Several computational accounts have modelled how surface syntactic cues affect the availability of discourse entities in the service of reference resolution (Grosz et al., 1983; 1986; in press; Kameyama, 1986; Reichman, 1985; Sidner, 1981); however, here I will concentrate on Grosz, Joshi and Weinstein's centering constraints and rules, as formalised in a centering algorithm by Brennan et al. (1987). At any moment, there is assumed to be one entity that the discourse is most "about"; this is the discourse center or *backward-looking center* (also known as the backward center).² The backward center is what is presumed to be the most salient entity at that moment (corresponding to the local topic) by both the speaker and the addressee. Each utterance is also associated with an ordered set of discourse entities, or *forward-looking centers* (also known as forward centers) that consist of all the discourse entities realised (either explicitly or implicitly) in that utterance. The algorithm ranks forward centers according to their syntactic prominence as expressed in the current utterance, a ranking assumed to approximate their likelihood of becoming the next backward center. The most highly ranked forward center is called the *preferred center*.³ The algorithm selects as backward center the most highly ranked entity on the previous utterance's forward center list that is also realised in the current utterance. Table 1 illustrates backward, forward and preferred centers.

Syntactic prominence is determined according to the hierarchy of grammatical relations; that is, sentence subjects are ranked more highly than objects, and subcategorised functions of the main verb are ranked more highly than adjuncts. This ranking represents the relative availability of entities as potential referents for upcoming referring expressions. It is consistent with evidence that topics of sentences are more likely to be co-specified by surface subjects than by surface objects (Reinhart, 1982; see also Zubin, 1979), that conceptually accessible items are more likely to appear as subjects than as objects (Bock & Warren, 1985), and that an object may not serve as the topic unless it is prosodically marked by de-stressing (Bolinger, 1979). Centering predicts that speakers will tend to mention highly salient entities as subjects rather than as objects.

²The backward center approximates what Chafe (1987) has called a "starting point".

³The preferred center corresponds to what Levelt (1989, p. 118) calls "the speaker's focal center". Unfortunately, the phrase "focal center" adds to the topic/focus confusion; Grosz et al. (1983) originally used "center" to mean *not* the stressed or new information that is often labelled "focus", but the salient and given information of a local discourse "topic".

previous utterance, and whether or not the backward center of the current utterance coincides with its preferred center (Brennan et al., 1987). The algorithm uses these ranked transitions to interpret adjacent utterances being about the same entities if at all possible and to resolve the meanings of pronouns. The co-specifier of a potentially ambiguous pronoun is computed from the previous utterance's forward centers in a way that maintains the most coherent transition—that is, the smallest attentional shift—between the two utterances.

Evidence from Text Comprehension Studies

Most studies of how people comprehend referring expressions in discourse make three assumptions: first, that understanding a definite reference involves either searching backward through the text of the discourse or else assessing an entity in a semantic model represented in working memory; second, that the text of a discourse is segmented (whether into turns of utterances, clauses, sentences, or some other units); and third, that discourse segment boundaries restrict the availability of potential referents during processing. The segmentation assumptions are supported by evidence that the most recent sentence is generally available for verbatim recall, whereas the sentence one back is not (Jarvella, 1971). Pronouns appear to depend on their interpretation partly on recency of mention (Clark & Sengul, 1979) —“referential distance” (Gernsbacher, 1989). But recency does not fully account for the availability of referents. In a test of a “discontinuity model” (Clark & Sengul, 1979), pronouns and definite noun phrases were interpreted fastest when the antecedent was in the immediately previous sentence, while those expressions having antecedents two or more sentences or clauses back were interpreted equally slowly, regardless of the number of intervening sentences. In another study, when a discourse entity was introduced in a sentence followed by an intervening sentence in which it was not mentioned, a pronoun co-specifying the same entity in a third sentence was read more slowly than without the intervening sentence, unless the entity had been introduced in a question; a question appears to keep entities in the local discourse context available until the answer arrives (Malt, 1988).

A “discourse center hypothesis”, that pronouns identify the entity that a discourse is most about at a particular point and that noun phrases shift reader’s focus of attention, was tested by Hudson, Tanenhaus and Dixon (1986) using two-sentence discourses. Each of these discourses described a situation involving two men or two women, and in each, the discourse center (corresponding to the backward-looking center) was introduced as the subject of the first sentence. In the second sentence, the discourse center was realised as either the subject or the object, and the subject of the sentence was either a pronoun or a full noun phrase. A full noun phrase (NP)

TABLE 1
Examples of Backward, Forward and Preferred Centers from a Sample Discourse, Segmented^a into Utterances

Utterance	Backward Center	Forward Centers	Preferred Center
and now Wolverines have the ball	[Wolverines] ^{bc}	{[Wolverines] [ball]} ^d	[Wolverines]
they're going down	[Wolverines]	{[Wolverines]}	[Wolverines]
number thirty passes	[]	{[thirty] [ball] [forty-one]}	[thirty]
it off to forty-one	[forty-one]	{[forty-one] [shot]}	[forty-one]
the shot			
and he misses	[forty-one]	{[forty-one]}	[forty-one]

^aSegmentation is crucial to determining centering data structures; however, local segmentation into “utterances” is underspecified in Grosz, Joshi and Weinstein’s original proposal (for discussion, see Suri & McCoy, 1993; 1994). Here, segmentation is into clauses.

^bThe assumption is that the initial utterance in this example continues to have the same backward center as the previous utterance.

^cAn item in square brackets represents a center that arises from a particular referring expression in the utterance (in this case, “Wolverines”).

^dItems in curly brackets represent a list of centers.

Additional aspects of surface syntax are relevant to the ranking of forward centers. For instance, a hierarchy of markedness for noun phrases and pronouns was proposed by Givon (1983) and tested by Fletcher (1984) as a way speakers can signal the degree of topic continuity from previous sentences in English text. This hierarchy ranges from zero anaphora and unstressed pronouns (signalling highest and second-highest degrees of topic continuity), to cleft/focus constructions (signalling lowest degree of topic continuity); similar givenness or accessibility scales relating information status to linguistic form have been proposed by Ariel (1990) and Gundel, Hedberg and Zacharski (1993). Pronouns signal that reference is being maintained within an episode (Tomlin, 1987). Consistent with these proposals, the centering framework predicts that speakers will pronominalise the backward center.

To model how the center of attention changes over time, the centering algorithm classifies attentional transitions between utterances. Each utterance, unless it is discourse-initial or changes the topic, is linked to the utterance before it by its backward center. A discourse is assumed to be more “coherent” when it continues to be about the same topic, and so, from the comprehender’s point of view, attention is shifted only when necessary (for instance, when the previous backward center is not realised at all in the current utterance). There are four possible ranked transitions between successive utterances, based on two independent factors: whether or not the backward center of the current utterance is the same as it was for the

subject position of the second sentence was read faster when it co-specified the discourse center than when it did not. Pronouns co-specifying the center were read even faster than full NPs co-specifying the center. But pronouns co-specifying a non-centered entity were read more slowly than full NPs co-specifying a non-centered entity. And when non-centered entities were co-specified by pronouns (as opposed to full NPs), sentences were more often rejected as not making any sense.

A set of studies specifically testing the centering theory measured self-paced reading times for short discourses (Gordon, Grosz, & Gilliom, 1993) and found a repeated name penalty (slower reading times) when the backward center was repeated as a full NP instead of being pronominalised; this repeated name penalty occurred for entities just mentioned in subject position but not for those in object position.

These findings support the idea that mentioning a brand new entity in subject position causes the reader to mentally instantiate that entity in the discourse model, and this takes even longer than understanding the “inappropriate” co-specification of a full NP with an existing center. Converging evidence comes from another study showing longer reading latencies for pronouns than for full NPs when the referents were not currently activated, and longer latencies for full NPs than for pronouns when the referents were activated (Guindon, 1985). In an experiment by Hudson-D’Zmura and Tanenhaus, three-sentence discourses were judged to be coherent less often and more slowly when an ambiguous pronoun co-specified a semantically plausible referent that had not been syntactically centered. Even when world knowledge makes it most plausible for a pronoun to co-specify an entity that is *not* the backward center, there appears to be an immediate or automatic preference to interpret the pronoun as co-specifying the backward center (Hudson-D’Zmura, 1988; Hudson-D’Zmura & Tanenhaus, in press).

An alternative way to rank the entities in the forward center list would be by the order of mention in the sentence, as this appears to contribute to both salience (Gernsbacher & Hargreaves, 1988) and topicality (Horn, 1986). For simple sentences with canonical English word order, ranking by order of mention would give the same result as ranking by grammatical relation. Studies by Gordon et al. (1993) teased apart order of mention and grammatical relation using sentences in which the subject was not always the first NP; they concluded that the most coherent discourses were those realising the backward center as a subject, but that *both* subjecthood and initial surface position contribute to making an entity salient in the forward center list.

To sum up, pronouns are neither harder nor easier to process than full NPs—it depends on the attentional context. Often, an addressee can rapidly fix the referent of a pronoun without having to wait for inferential processes

to finish; in fact, the addressee may have an immediate expectation for a pronoun’s co-specifier (Greene, McKoon, & Ratcliff, 1992).

To Pronominalise or not to Pronominalise

A speaker may actually construct and track a model of the addressee’s attentional state, or else her choice of a pronoun may simply reflect her own attentional state (for a discussion, see Dell & Brown, 1991). What guides a speaker’s choice to pronominalise? Here I will focus on production in a conversational setting to test three predictions derived from centering:

Prediction 1. When a speaker mentions an entity as a surface subject, as opposed to as a surface object, she implicitly indicates to her listener, “my attention is here”. That is, an item that seems likely to figure prominently in the near future is more likely to be mentioned with a full NP subject than a full NP object. Furthermore, entities mentioned as subjects are more likely to be continued as the backward center than are those mentioned as objects.

Prediction 2. An entity should not be pronominalised unless it is already salient; that is, pronouns mark what the speaker thinks is the center of attention for both herself and her listener (the backward center). When a speaker does choose to pronominalise, she confirms “we are both attending to this, it’s salient”.

Prediction 3. When a speaker re-refers to an entity that she just mentioned as a full NP object, she should not use a pronoun, since the entity it co-specifies is not yet salient. She should mention the entity in subject position to mark “I’m shifting attention onto this”.

Let us consider an alternative to centering. Several computational approaches to discourse have modelled the resolution of pronouns and definite noun phrases using semantic information and world knowledge, without taking advantage of the kinds of syntactic constraints that centering uses (e.g. Hobbs, Stickel, Appelt, & Martin, 1993; Sowa, 1984). While knowledge-based theories often succeed in resolving referring expressions in this manner (Walker, 1989), they do not model human discourse processing. An entirely knowledge-based algorithm would not reproduce an addressee’s immediate tendency to interpret a pronoun as co-specifying the backward center, even when this results in an implausible interpretation. In addition, approaches that use knowledge alone do not predict when a speaker is likely to produce a pronoun instead of a noun phrase.

A traditional assumption about pronouns is that they can replace any linguistic constituents in which there is no added information relative to the intent of the speaker and that they exist as place-holders in the surface

structure (Olson, 1970). Another traditional assumption is that pronouns save effort:

There are many kinds of anaphors in English. Their fundamental function is presumably to enable speakers to avoid wasting breath and listeners to avoid wasting energy on unnecessary verbiage. If a reference or sense has already been established, it can be communicated either by a pro-form or by a manifest "gap" in an utterance. (Johnson-Laird, 1983, p. 393)

One of the implications of these traditional assumptions that is relevant to a knowledge-based view is that a speaker should be able to pronominalise when an antecedent has been realised in the discourse already and whenever there is enough information so that her listener can easily and unambiguously interpret what she means by the pronoun. On such a view, a pronominalised entity need not be particularly salient in the discourse model; it need only be identifiable. What I will call a *knowledge-based strategy* is the idea that a pronoun co-specifies the most recent, semantically plausible noun phrase that agrees in gender and number. On a knowledge-based strategy, speakers would produce pronouns rather than full noun phrases whenever referring to the most recently mentioned, semantically plausible discourse entity. The predictions that follow contrast with those of centering; for instance, a speaker should pronominalise the very next reference to something she has just mentioned as a full NP object, as long as there is sufficient world knowledge so that the pronoun is not ambiguous.

To test these predictions, I designed an experiment to see if speakers use grammatical relations and pronouns as a function of the prominence of physical entities in a physical environment, as reflected in the salience of entities in a discourse model.

METHOD

Pairs of people were audiotape-recorded during a conversational task. Their attention was manipulated with a videotape of a basketball game. One person viewed the videotape and described it to the other, who was seated behind a barrier.

Subjects

Twenty-eight Stanford undergraduates (23 males, 5 females) who identified themselves as basketball fans volunteered to serve as subjects in exchange for research credit in an introductory psychology class. All of them were native speakers of English, and none had prior experience announcing

sporting events. The subjects participated in pairs; one pair happened to be friends and the other pairs appeared to be unacquainted.

Materials

The stimulus was a videotape of a men's college basketball game between perennial rivals, the Michigan State Spartans and the University of Michigan Wolverines. The videotape was recorded from a live television broadcast. I selected the domain of basketball for several reasons. First, it was dynamic, fast moving, and likely to influence the order and pace of the verbal descriptions. Second, it exerted a reasonably high degree of control over the viewer's potential focus of attention (as defined simply by what was visible and who had the ball). This made it possible to direct what people were talking about, to compare across speakers, and later, to reconstruct what they meant. And third, the domain provided speakers with many opportunities to refer to entities that had the same gender and number.

Since it was desirable for people to refer to the individual basketball players as often as possible, the particular game was chosen because the numbers on individual players' uniforms were frequently visible. Before watching the videotape, each member of the pair was given a program listing the players on each team by name and number. The videotape began with a 5 min pre-game warm-up show, included in order to familiarise people with the players and to make the game more engaging. The game portion of the videotape included approximately 15 min of broadcast time (or about 7 min of actual playing time) taken from near the end of the first half of the game. All commercials were edited out and the sound was erased from the entire game portion of the tape. As with most broadcasts of sporting events, occasionally the score or statistics about a player were superimposed on the screen.

A concise script of the action of the basketball game was created in order to code the set of events that manipulated the speakers' attention. This script was a list of all the players who handled the ball and included some additional information of sporting interest such as shots and fouls. The form of this script was propositional. Each "reference event" began when a player got possession of the ball and ended when he no longer had possession. Three consecutive events from the script appear in Table 2. These events were classified as either high prominence or low prominence, based simply on the action of the game. In the high-prominence events, a player seemed likely to continue being the center of attention for a while—for instance, when he took the ball all the way across court, tried to shoot, was involved in a foul, or fell down. In the low-prominence events, possession of the ball was changing rapidly, with great uncertainty as to what would happen next. I selected a total of 86 events, using three criteria: (1) that these events should begin

about 1 min into the videotape to give the subjects time to get used to the task; (2) that the players' numbers should be visible on the videotape; and (3) that there should be an equal number of both kinds of events. There were 43 events in the low-prominence category and 43 in the high-prominence category.

Procedure

The members of each pair were told when they arrived for the experiment that one of them would be doing a play-by-play description of a basketball game, and that the goal was to see how well the other could visualise the action of the game. After the pair watched the pre-game show together, the videotape was stopped and a coin flipped to determine who would be the "announcer" and who would be the "audience" partner. Then, the announcer remained in front of the television set and the audience partner went to sit behind a barrier where neither the game nor the announcer was visible. The audience partner was given an additional sheet of paper that contained three questions and was instructed to write down answers whenever a light went on. These questions were: "What player has the ball now?", "What team is ahead?" and "What is the score?" The light signalling the questions went on at arbitrary points during the game and could not be seen by the announcer, although the announcer knew what the audience's task and questions were. The announcer was instructed to describe the action of the basketball game so that the audience would find it interesting, be able to visualise it, and be able to answer the questions. Announcers were told that they need not try to sound like professional announcers.

The audience's task was included to provide the pair with a purpose for communicating. The questions—in particular, "What player has the ball now?"—were used to motivate the announcer to describe the game in sufficient detail and to refer to individual players rather than to teams. This did not happen in pilot trials that lacked a motivating task for the audience partner. Then, announcer partners (like professional play-by-play announcers) tended to refer more to the teams than to individuals in possession of the ball, as in "Blue's comin' down the court now ...". Because the purpose of the questions was mainly to motivate the task, and because

TABLE 2
Three Consecutive Events Excerpted from the Propositional Script for the Basketball Game

<i>Event</i>	<i>Prominence</i>
Blue-30 (Thompson) runs-to-basket, passes	High prominence
Blue-11 (Joubert) gets-ball, passes	Low prominence
Blue-41 (Rice) shoots, misses	High prominence

announcers' descriptions, no data were collected from them. The audience partner was allowed to talk freely to the announcer as long as the light was not on.

Analysis of Transcripts

The tape-recorded speech was transcribed, and the transcripts were checked twice against the original audiotapes. The videotape manipulated the announcers' attention so that their verbal descriptions would include a series of expected reference events. A reference event was considered to be realised when a portion of the transcribed discourse described all or part of the event as listed on the propositional script. A reference event consisted of the continuous discourse that began when a player took possession of the ball and ended when someone else got the ball or when the ball went out of play. Each transcript was matched with the propositional script and segmented into realised reference events. This was done by me and also by a research assistant blind to the hypotheses of the experiment. The two sets of coded transcripts were then compared. There were segmentation differences in fewer than 3% of the referring events; we discussed these differences until we reached agreement. Table 3 contains examples of how four speakers described the events summarised by the propositional script in Table 2.

Next, we coded each reference event that was realised in the discourses

TABLE 3
Four Speakers Describe the Events Listed in the Script in Table 2

<i>Speaker</i>	<i>Description of Reference Event</i>
Speaker 4	"Wolverines movin' it slowly up the back court, over to number ... back to forty-one, he shoots from three points ... no."
Speaker 6	"Michigan's bringin' the ball down a little bit slowly. Thompson, the top court, out to Joubert, shot at the top of the key by Rice, is no good ..."
Speaker 9	"Thompson bringin' the ball up the court now. Joubert must be out, I don't see 'im in there. Yes, he was substituted in for. I think. No, there's Joubert. Joubert with the ball, brings it over to Rice, Rice, for about a fifteen footer, he misses, in and out."
Speaker 14	"Number thirty passes it off to ... forty-one. Forty-one goes up for the shot, and he misses."

relations of these references. The referring expressions coded fell into these categories: personal pronouns, relative pronouns, NP subjects, NP objects, NPs appearing after a copula (e.g. “it’s Joubert shooting”), disembodied NPs (noun phrases with missing verbs, e.g. “now Joubert, coming down the court”), missing or dropped subjects, and non-restrictive modifiers (e.g. “Todd Wolfe, number 42”, was coded as a NP followed by a non-restrictive modifier). Full NP subjects, the relative pronoun “who” acting as a subject and full NPs appearing after copulas, were counted separately but combined in the full NP subject category for the analyses (unless otherwise stated). NP objects were objects of prepositions or of verbs. Each realised reference event was coded for its complete sequence of referring expressions. A typical sequence of referring expressions to one player might consist of [NP subject, pronoun].⁴ Each realised reference event was categorised according to the prominence (high or low) assigned by the propositional script, and the referring expressions within each event were counted. Non-restrictive modifiers and restarts due to speech errors were excluded from these counts.

RESULTS

The recorded conversations indeed reflected the action of the basketball game; out of a possible total of 1204 events from the propositional script, the 14 announcers spontaneously described 884 of them, or 73%. In 5 of the 14 pairs, the audience partner helped motivate the task at the outset by explicitly asking the announcer for more information. If the classification of reference events into high and low prominence were valid, low-prominence events should be more likely to be omitted from the discourses than high-prominence ones. And they were: on average, a speaker left out 15.6 of the 43 low-prominence events and 7.2 of the 43 high-prominence events [$t(13) = 6.05$, $P < 0.001$]. Of the 884 events realised in the discourses, 198 (22%) contained one or more masculine singular pronouns, most of which co-specified a previous full noun phrase in the same realised event.

Prominence and First Mention in a Reference Event

When a speaker describes the action of a basketball game, she has many options. She can say “X has the ball” or “X gets the ball” or “X brings it in”, mentioning player X as a surface subject. She can also say “Y passes to X”, “Y dribbles over to X”, or simply “over to X” or “back to X”, introducing player X as a surface object. In general, speakers introduced entities during reference events as subjects more often than as objects. The question here is

⁴The grammatical roles of pronouns are not reported here, as no predictions were made.

prominence events, thus indicating where they expected the center of attention to be next. Prediction 1 is that prominent entities (players who have possession of the ball during a high-prominence event) should be more likely to be introduced into a segment as full NP *subjects* than low-prominence entities, and this was found to be the case [53.8% *vs* 35.1%; $t(13) = 3.9$, $P = 0.002$; see Table 4]. Low-prominence entities were more likely to be introduced as full NP *objects* than high-prominence entities [52.3% *vs* 29.1%; $t(13) = 5.90$, $P < 0.001$]. So prominent entities were expressed as surface subjects in a way that reflected their salience in the discourse model.

Prominence and Pronominalisation

Consider prediction 2: that speakers should use pronouns to express the most salient entity in the discourse (e.g. the backward center), as opposed to pronominalising whenever they could do so on a knowledge-based strategy. The logic of testing prediction 2 was as follows: The first expression in an realised event was almost always a full NP. Therefore, the referring expressions of interest would be those that followed the first one in a reference event, co-specifying the same entity. Because of the structure of these reference events, any referring expression following the first full NP and co-specifying the same entity could be a pronoun without causing any ambiguity problems for the addressee. So if speakers really used pronouns whenever they could do so without being ambiguous, all referring expressions *downstream* from the initial one should be pronouns. We computed the percentage of downstream referring expressions that were pronouns (co-specifying the same player as the first full NP), where there were no intervening references to other players within the reference event. Only 55% of the downstream referring expressions were pronouns, while 45% were full NPs [no significant difference: $t(13) = 0.75$, $P = 0.46$]. The fact that pronouns were not used more often than full NPs to co-specify the given or discourse-old entities suggests that the choice of a pronoun determined by something other than recency.

I then compared the proportion of downstream referring expressions that were pronouns in high-prominence events to the proportion in low-prominence events. If speakers used pronouns simply to refer to the most recently mentioned, semantically plausible entity, then pronouns should have constituted equal proportions of the references downstream from the first one in an event, regardless of whether the event was of high or low prominence. But this was not the case. Pronouns made up a greater percentage of the downstream references in high-prominence events than low-prominence events [60.5% *vs* 27.7%; $t(13) = 2.84$, $P < 0.05$]. Since speakers attended more to high- than to low-prominence events, this res-

the center of attention.

Repetition of full NPs

What happened to discourse entities that were first mentioned during a reference event as surface objects? In those events when an entity was first mentioned in object position but then went on to become salient in the discourse model, how did the speaker move it into the center of attention? Prediction 3 says that after speakers mention entities as full NP objects, they should re-mention the entities again as full NPs, most likely in subject position, before pronominalising (that is, if they continue to talk about them). A knowledge-based strategy predicts the opposite: speakers should be able to avoid "unnecessary" repetition and pronominalise with the very next reference after mentioning an entity in object position.

The basketball game discourses were particularly appropriate for testing this syntactic prediction for two reasons. First, there was abundant world knowledge that could be used to interpret pronouns—after all, a player can't shoot or pass the ball unless he gets it first—and so this semantic redundancy might cause speakers *not* to have to center entities syntactically before pronominalising. It certainly seemed as if the announcers relied a great deal on their partners' knowledge about basketball games. For instance:

1. "back to **forty-one**, **he** shoots from three points ... no!"

This speaker seemed to assume that the addressee would understand that player forty-one was shooting all the way from the three-point line, a position very far from the basket (in these examples, expressions co-specifying the same entity are printed in boldface). The second reason that the basketball discourses provide a good test of prediction 3 is that the task itself favours a recency strategy. That is, the action of the game was

TABLE 4
How Speakers Introduced Entities into Reference Events: Overall Frequency (and Mean % per Speaker)

Event type	Low Prominence	High Prominence	Total
Object NPs	215 (52.3%)	148 (29.0%)	363 (39.5%)
Subject NPs	127 (35.1%)	266 (53.8%)	393 (45.7%)
Post-copular NPs	7 (1.7%)	26 (4.5%)	33 (3.6%)
Disembodied NPs	33 (9.6%)	45 (9.3%)	78 (9.0%)
Pronouns or zeros	1 (1.4%)	16 (3.4%)	17 (2.3%)
<i>Total</i>	383 (100%) ^a	501 (100%)	884 (100%)

^aBecause of rounding, some columns add up to 100.1%.

recency bias, and speakers had to keep up the pace of their descriptions. Since a pronoun is shorter than a proper name, number or description, speakers might be expected to avoid repetition and to pronominalise as soon as possible, as in example (1).

If, on the other hand, speakers fashion their referring expressions in ways consistent with the centering theory, there are several ways they can make a discourse entity salient enough to move into the discourse center. First, speakers could repeat the full NP in subject position before pronominalising:

2. "and now Wolverines have the ball, they're going down.

Number thirty passes it off to ... **forty-one**.

Forty-one goes up for the shot, and **he** misses."

Or they could re-mention the discourse entity of interest as a relative pronoun that serves as the subject of a clause:

3. "Eleven goes out front, and ...

he passes to **forty-one**,

who takes a shot from the left-hand side of the free throw, and **he** makes it.

Two points for the white team."

And, finally, they could repeat the full NP verbatim, in the absence of any verb:

4. "Rice with the ball in the middle, over to **Joubert**, **Joubert** with a nice move, basket's off, shot's off."

The second NP in constructions like "Joubert with a nice move" in example (4) was analysed as a post-copular NP subject, only with the copula deleted, following Ferguson (1983). Examples (2), (3) and (4) are consistent with the third centering prediction, and example (1) is counter to it. In example (1), the entity is introduced into the event as a sentence object immediately before it is mentioned as a personal pronoun, whereas in examples (2) and (3) there is intervening mention as a full NP subject or relative pronoun subject before mention as a personal pronoun. Of the 884 realised events, there were 50 containing one or more personal pronouns, in which the player was introduced at the beginning of the event with a full NP object. In 23 of these the entity was not syntactically centered by being re-mentioned as a subject before being pronominalised, and in 27 it was; these numbers are not significantly different. But the critical comparison is this one: When a player is introduced during an event as an object and then mentioned again, is that second mention more likely to be a full NP subject or a pronoun?

Verbatim repetition of a player's name or number, as in examples (2) and

(4), was surprisingly common. In fact, whenever a speaker introduced a player into an event using a full NP in object position, they were more likely to re-refer to that player by repeating the full NP verbatim than by pronominalising [$F(1,13) = 7.13, P < 0.02$]. Of the 367 events in which speakers introduced a player as a full NP in object position, 85 re-references were verbatim repetitions of the full NP and 23 re-references were pronouns. In these events, there was little or no intervening speech between the repeated references, and in all cases the second full NP was a subject, as in example (2), or a post-copular NP, as in example (4). This count was a conservative estimate in that it excluded non-restrictive modifiers where the speaker might use a second full NP to add identifying information,⁵ and I counted only cases where the second full NP was a verbatim repetition of the first.

Just what were these speakers doing when they repeated a full NP verbatim? According to Bolinger (1979), the decision to repeat a noun rather than use a pronoun is a motivated one; such repetition depends on how necessary the speaker finds it that the referent be *re-identified*: "the majority of reidentifications probably occur after a break of some kind" (Bolinger, 1979, p. 308). Tomlin (1987, p. 458) concurs: "Whenever attention focus is disrupted, the speaker reinstates reference with a full noun, no matter how few clauses intervene between subsequent references". For our purposes, such attentional "breaks" would correspond to segmenting the discourse on the global level, consistent with the fact that during most reference events, the very first mention of an entity was as a full NP.

But the need to re-identify an entity does not account for the repetitions that occurred *within* the reference events. There were no breaks or interruptions in these 85 verbatim repetitions. According to the centering theory, the motivation for repeating a full NP can be quite different from re-identification. These repetitions "move" the full NP from object position into subject position, mark its referent as prominent, and make its corresponding discourse entity the backward center of the utterance, salient enough for pronominalisation.

Table 5 summarises the different kinds of double nominalisations that appeared in the basketball discourses. I counted all the reference events that contained two full NPs as their first and second referring expressions. The most frequent category, verbatim repetition while moving the NP from object to subject position, appeared to serve the function of bringing an entity into the discourse center. With the first three kinds of repetitions listed

⁵Players were referred to by number in 39 of these repetitions, and by proper names in 46; sometimes a proper name would be added as a non-restrictive modifier after a number, or vice versa. Each speaker tended to refer to all players by name or by number, but not both.

Re-Nominalisations	Low Prominence	High Prominence	Total
Verbatim repetition from object to subject ^a	43	42	85
Verbatim repetition from subject to object	0	6	6
Verbatim repetition from subject to subject	5	13	18
Repetition after interruption or side sequence	1	30	31
Repetition due to repair or clarification	0	19	19
Re-nominalisation to add information	4	13	17
Total	53	123	176

^aThe subject category includes post-copular subjects and disembodied subjects (with copula deleted).

in Table 5, the speaker chose to repeat a full NP verbatim instead of pronominalising. The second most frequent category appeared to serve the function of re-identifying an entity after an interruption or side sequence, as Bolinger expected. Repetition was also used as a self-repair or to clarify the identity of ambiguous players. The re-nominalisations reported in Table 5 do not seem to be an artifact of the task's forcing speakers to be excessively specific; after all, the basketball discourses did include many third-person singular masculine pronouns. The verbatim repetitions that served to promote a NP object to a NP subject contained no intervening verb phrases or interruptions. The only intervening words in these repetitions were connectives such as "and" or locative adjuncts such as "on the right-hand side of the court" as in:

5. "Thirteen has the ball up front again,
passes it to **forty-two** on the right-hand side of the court,
forty-two puts the ball on the floor,
he passes it to twenty-three."

Prosody

Prosody can be an important cue to the attentional structure of discourse (Grosz & Sidner, 1986; Hirschberg & Pierrehumbert, 1986). Clearly, stress can affect the interpretation of a pronoun. Consider the contrast between these hypothetical examples (adapted from ones attributed to George Lakoff):

c. Herb called Gordon a Republican, and then *he* insulted him.

Stressing the pronouns in example (7) has the effect of breaking the expected co-specification of “he” with “Herb” and “him” with “Gordon”, for the interpretation that it’s an insult to be called a Republican.

Stress is used to mark the information status of information in a discourse. Brand-new (hearer-new) information tends to be introduced using high pitch, an “attention marker”, while already evoked information tends to be mentioned using low pitch (Brown, 1983). Prosodic salience is expressed not only by pitch, but also by amplitude, duration and spectral content, and these variables are not independent of one another (Fry, 1958; Gay, 1978). In particular, the cue of vowel quality can mark some pronouns as reduced, making them unaccentable. Such encliticised pronouns lack any rhythmic disjuncture with an adjacent word or “host” that is ordinarily the following verb or preposition (Selkirk, 1984). The reduction of vowels dispells ambiguity, so that other, full-vowel elements can be accented with a greater variety of intonational shapes. It is unusual for a pronoun not to have a reduced vowel; in fact, “no cue other than the full vowel is needed to mark the accent on the pronoun” (Bolinger, 1986). A longer duration alone may be enough to mark the lack of salience of a discourse-new referent (Chafe, 1976); after all, the second use of a word in a discourse tends to be shortened relative to its first use (Fowler & Housum, 1987). In terms of Prince’s (1981) given/new taxonomy, while information in the discourse center is *salient* given information, this information may be related to *predictable* given information as well. Predictable information is recoverable from the preceding discourse and frequently deletable (Prince, 1981). So when pronouns are not reduced,⁶ it may be that they are marked as being not predictable from the preceding syntax; that is, although they are given in the sense that their referents are known to the hearer, they are also new, in Prince’s (1981) senses of being unpredictable and not currently salient.

Recall the small number of instances where a player was mentioned by being pronominalised immediately after being introduced into an event as the object of a sentence. What might these apparent violations of syntactic centering indicate? Were speakers relying on the characteristics of the situation and the inferencing abilities of their addressees to be understood, or were they simply being careless? Or could they have been using other, non-syntactic, centering mechanisms? Another candidate for a centering mechanism could be prosody.

⁶I use “not reduced” rather than accented here because even when a pronoun is given some prosodic prominence, it still does not usually receive the primary pitch accent of the phrase. For more on the contribution of pitch accents to prominence, see Pierrehumbert and Hirschberg (1990).

Because the center of attention shifted rather quickly among a limited set of players in the basketball game, personal pronouns frequently had several evoked (Prince, 1981) but displaced (Brown, 1983) potential co-specifiers. If entities could be centered using some combination of prosody and syntax, then discourse-new information should be more phonologically prominent than salient given (or discourse-old) information. This suggests that referring to a non-centered entity (mentioned in the previous local discourse segment with a full NP object) using a prosodically prominent pronoun could be consistent with the centering theory after all; that is, although pronominalising itself does not move a discourse entity into the center of attention, stressing a pronoun may. In the centering framework, such prosodic marking might enable the listener to understand not only that the speaker’s attention is on the entity that the pronoun co-specifies, but also that she is trying to make it salient, since the listener is probably not attending to it already. Lengthening the pronoun could be used to make the discourse entity it co-specifies immediately salient.

To test for prosodic differences, I reanalysed the set of realised events that were introduced by full NP objects and that also contained at least one pronoun. Recall that 27 of these (set I) contained an intervening NP subject or post-copular NP, and thus were consistent with the syntactic predictions of centering, while 23 of these (set II) pronominalised directly after a NP object, and thus were inconsistent with the syntactic predictions of centering. Of the realised events in sets I and II, 14 were eliminated from further analysis either because the pronoun was not in the nominative case and thus its length would not be comparable to those of other instances of “he”, or else because there was intervening syntactic material consisting of verb phrases with missing subjects, sentential complements, or other potentially distinct local discourse segments between the first mention of the player in object position and its second mention. This was done to make the two sets of events as free from additional “interruptions” as possible. This left 36 reference events, of which 19 used syntactic centering (set I) and 17 did not (set II). I digitised these 36 events and measured the length of the first “he” in each one. For this, I used an IBM personal computer and a speech analysis software package, MSLPitch. The software presented the digitised speech as a spectral envelope aligned with F0 contours on a graphic display. To measure the length of a pronoun, I set two cursors (constituting a window) on the display while listening to the segment of speech within the window. I segmented each pronoun as accurately as possible, while listening to the speech and visually inspecting the waveform. While marking pronoun lengths, I was blind to whether events belonged to set I or set II.

Pronouns co-specifying entities that had not been syntactically centered were longer in duration than those that had been. There was a reliable difference of 28 msec between the lengths of pronouns in sets I and II [when

each "he" was weighted the same regardless of speaker: $F(1,34) = 22.05$, $P < 0.001$. Pronouns in set II, whose co-specifiers had not already been syntactically centered, averaged 120.4 msec in length ($SD = 22.3$), and pronouns in set I, whose co-specifiers had already been syntactically centered, averaged 92.6 msec ($SD = 12.3$). When average length differences were compared within each of the seven speakers who happened to contribute events to both sets, the difference was still reliable [$F(1,6) = 13.11$, $P < 0.02$]. This amount of temporal lengthening could result from pronouncing a full vowel and may be sufficient to mark a pronoun as no longer predictable given information, but as *new* in the sense of being *unrecoverable from the preceding syntax* (in the terms of Prince, 1981).

Finally, the basketball discourses contain many examples of speakers dropping subjects altogether (in the next example, missing subjects are denoted by "@"):

8. "For Michigan State, **guy's** dribblin' a whole lot,
@ runnin' around,
he's in trouble, heavy defense,
@ can't handle the pressure,
@ passes it back out to number forty-two ..."

In such examples, some but not all of the missing subjects could be attributed to the conjunction reduction possible with predictable given information (Prince, 1981). Note also that dropping sentence-initial NPs and copulas ("prosiopesis") has been identified as a feature of sports talk by Ferguson (1983). Whether the attentional mechanisms that enable a speaker to use a pronoun are related to those that enable dropping one is beyond the scope of this study.

DISCUSSION

These data strongly support the centering account of pronominalisation and fail to support a knowledge-based account. Speakers' choices of grammatical relations and referring expressions were predicted by the centering hypothesis. These choices were consistent with a ranking scheme that considers surface subjects more salient than surface objects. The data may also be consistent with a first-mention scheme; these two possibilities are not separated here (but see Gordon et al., 1993). In the small number of cases where speakers did not syntactically center an entity before pronominalising it, they tended to prosodically lengthen the pronoun. This may have contributed further to the salience of its referent, perhaps by signalling that it was not recoverable by simply looking to the preferred center from the previous clause.

The evidence presented here suggests some ways people can coordinate

their attention during conversation. Speakers can indicate what is salient to them by mentioning an entity as a surface subject as opposed to a surface object. They can confirm that something is given and already salient by referring to it with a pronoun. They can repeat full noun phrases for reasons other than re-identifying referents—that is, in order to move entities into the center of attention. The evidence shows that speakers do not use pronouns to refer to the most recently mentioned, semantically plausible entity, but in a way that systematically marks what appears to be salient, both to them and their addressees. Such choices could help addressees confirm and predict speakers' movement of attention during the conversation, enabling both partners to coordinate their attention over their individual mental models.

Perhaps the most compelling evidence for centering comes from the repetition result. When speakers referred to an entity as a full NP object, they were much more likely to re-refer to it immediately by repeating it verbatim as a full NP than by pronominalising it. This was the case even though using a pronoun would usually not have been ambiguous.

Do descriptions of basketball games provide a strong test of the centering hypotheses? Consider the task: it was designed to be engaging to speakers and addressees, to get speakers to generate chains of references to the same entity, and to maximise the likelihood of potentially ambiguous third-person singular pronouns. The task was also biased against centering prediction 3 and towards the alternative strategy in two ways. First, speakers were supposed to (and did) keep up with the pace of the game. Since re-referring with a pronoun always took up less speaking time than re-referring with a repeated full NP, speakers could have tried to use a pronoun whenever they could "get away" with doing so, such as whenever the pronoun would co-specify the most recent full NP. Second, speakers could have relied on addressees to use inference to resolve pronouns; this would have freed speakers from having to center entities syntactically (by repeating a full NP object in subject position) before pronominalising. So the most efficient way to describe the game using a knowledge-based strategy would have been with a series of agent-action-recipient sentences; the recipient in one sentence would become the agent of the next, where it would be specified with a pronoun. This did not happen; although task and event structures favoured the knowledge-based strategy, the data supported centering.

Could the present results have been biased by the features of "sports announcer talk"? Sports announcers typically use unusual syntactic elements such as preposed predicates, "heavy modifier" constructions, and dropped sentence-initial NPs and copulas (Ferguson, 1983). Of heavy modifiers and preposed predicates, Ferguson claims: "Most listeners to sportscasts are probably unable to use such devices without considerable practice, even though they may be thoroughly familiar with the devices in written English" (p. 163). This was clearly not the case with the subjects in

this experiment, all novice announcers. Apparently these speakers were able to use some of the same syntactic devices that professional announcers use to describe repetitive, highly constrained actions, while at the same time adapting to the idiosyncratic requirements of the task and referring repeatedly to players as individuals (as opposed to referring more often to teams or formations of players, as professional play-by-play announcers tend to do and as my announcer subjects did in pilot trials before I added the audience partners' task).

The features of sports announcer talk produced by the speakers in this study appeared to evolve naturally from the requirements of the discourse. Canonical English word order requires that subjects of sentences be articulated first; however, when a player cannot be named immediately (for instance, if his number isn't immediately visible), the most opportunistic strategy is to describe the player or his action up front and identify him later (Ferguson, 1983; Green, 1980). This would lead naturally to the greater use of proposed predicates and heavy modifiers. I coded non-restrictive modifiers (a type of heavy modifier) but did not include them in any counts, so this kind of construction did not figure in any of the present results. It would, of course, be desirable to replicate these findings using corpora of spontaneous speech from other domains.

Many other questions relevant to centering attention in discourse and resolving referring expressions remain unanswered. For instance, inferences based on world knowledge sometimes cause an addressee to revise a pronoun's original interpretation (Greene et al., 1992; Hudson-D'Zmura & Tanenhaus, in press). Inference-making is guided by many kinds of information; for instance, the awareness of who the protagonist is in a story (Francik, 1985; Morrow, 1985a; Stark, 1986), causal bias towards grammatical subjects or objects implicit in certain verbs or parallel constructions (Au, 1986; Frazier et al., 1984; Grober, Beardsley, & Caramazza, 1978), temporal bias and episode shifts (Anderson, Garrod, & Sanford, 1983), and foregrounding and backgrounding due to verb aspect (Morrow, 1985b). Although the centering framework acknowledges that both syntactic and semantic constraints guide pronoun resolution, it has not explicitly addressed the question of whether the contribution of world knowledge should be integrated with syntactic constraints in the same algorithm for pronoun resolution. While knowledge affects the ultimate interpretation of a pronoun, its contribution probably comes after the contribution made by grammatical relation (Hudson-D'Zmura, 1988; Hudson-D'Zmura & Tanenhaus, in press), and often these factors converge on the same referent.

The data reported here demonstrate that speakers use linguistic devices to encode the salience of an entity in the discourse model. But how does the salience of an entity in the model *decay*—abruptly or gradually? The

centering algorithm does not model the decay of salience explicitly; basically, when the backward center of one clause is not realised in the next clause, it gets displaced by another backward center. But salience may decay based on additional factors. Consider these hypothetical examples:⁷

9. Smith to Jones. Jones shoots.
10. Smith to Jones. He shoots.

vs

11. Smith hands the ball over to Jones. Jones shoots.
12. Smith hands the ball over to Jones. He shoots.

While centering predicts that repeating the full NP in (9) and (11) should be better than pronominalising in (10) and (12), intuition seems clearer with respect to the first pair than the second [the native speakers I have asked prefer (9) to (10), while some of them prefer (12) to (11)]. One possible explanation is based on the relative distance between the pronoun and its two potential antecedents, "Smith" vs "Jones". Perhaps the entities realised by "Smith" and "Jones" are in competition and their salience decays with speaking time or textual distance. In (9) and (10), there would be more ambiguity, and so re-referring with a full NP sounds better than with a pronoun. With the greater number of intervening syllables in (11) and (12) Smith's salience has decayed somewhat relative to that of the NP object (Jones), which may then displace it as the most salient item (or preferred center), even though it is only an object. This would explain why a pronoun may be more acceptable in (12) than in (10). Current centering algorithms would need to be modified if the data show a need to model the relative decay of salience of discourse entities over time, due to intervening syntactic material.

An alternative explanation of these examples would be that the additional distance after "Smith" in (11) and (12) allows additional time for inferential processing, so the pronoun isn't ambiguous any more. The addressee would have more time to make the inference that the player who has the ball must be the one who shoots. This explanation is compatible with the centering framework and with Hudson D'Zmura and Tanenhaus's finding that people immediately interpret ambiguous pronouns as co-specifying the center, even when this co-specification is semantically implausible. It seems likely that the best way to integrate the contributions of different sources of constraint is to consider pronoun resolution as a dynamic process where different kinds of information come into play at different times. Centering, as it has been described here, models the immediate bias due to syntactic constraints.

⁷I thank Kathryn Bock and an anonymous reviewer for bringing these examples to my attention.

To further explore examples like these in the current corpus, I re-examined the 50 events beginning with a full NP in object position and containing at least one pronoun, and coded those where the last mention of the player in the previous event (corresponding to "Smith") had been in subject position. This yielded 30 reference events: 20 centered the entity with a repeated full NP, like examples (9) and (11), and 10 did not, like (10) and (12). The "Jones-Jones" examples had a mean distance of 8.7 syllables between the first subject and object (that is, between "Smith" and "Jones" in the first sentence) and the "Jones-he" examples had a corresponding mean distance of 11.7 syllables between the first subject and object; however, this difference was not reliable ($P > 0.25$). While the direction of the means supports the intuition that a pronoun is more acceptable in (12) than in (10), these mean distances are surprisingly high [much longer than examples (11) and (12)]. Systematic tests about addressees' preferences for examples like (9) and (12) are called for. Regardless of such intuitions about examples like these, the fact remains that speakers in the current study were much more likely to produce examples consistent with syntactic centering like (9) and (11) than like (10) and (12).

Another unresolved issue concerning the centering framework is that the mappings among Grosz and Sidner's three levels of discourse structure (linguistic, intentional and attentional) have not been fully specified. The current study takes a step towards understanding the mapping between one kind of linguistic expression—the noun phrase—and the attentional structure of discourse. But speakers can use many other kinds of linguistic devices such as cue words, changes in rate of speech, changes in pitch range, or a variety of rhetorical devices to signal structural elements such as side sequences, interruptions or other discourse segments (Fox, 1987; Grosz, 1977; Grosz & Sidner, 1986; Hirschberg & Pierrehumbert, 1986; Linde, 1979; Polanyi & Scha, 1986; Reichman, 1985). All these devices have not been clearly integrated into one computational framework.

Another unanswered question concerns discourse segmentation. The boundaries of linguistically determined local discourse segments are assumed to restrict the availability of discourse entities. So the kind of local discourse segments that are assumed (such as utterances, sentences, clauses, complements, etc.) will have consequences for local focusing algorithms. Centering algorithms have been vague about the segmentation problems arising with complex sentences (Suri & McCoy, 1993; 1994). In addition, if we are to model the language processing that occurs during conversation, segmentation has to be done on-line, and not by looking ahead in the text of the discourse in order to segment what has gone previously.

A theoretical question has to do with the "size" of the backward center and the level of description of the entities that can occupy it. In the current experiment, I investigated a set of entities with relatively equal status and

granularity in the discourses, coding only NPs and masculine singular pronouns co-specifying basketball players. The transcripts also included many other pronouns; for instance, "it" co-specified a variety of entities—a shot, the ball, the basket, the game, the score, and entire situations. Sometimes the backward center was undoubtedly the basketball, not the player; many more utterances contained both "it" and "he" than contained two masculine singular pronouns. Centering, of course, does not preclude an utterance having multiple pronouns; the original formulation by Grosz, Joshi and Weinstein says only that if any entity in an utterance is pronominalised, the center must be (although this rule may not be strong enough, given the "explicit mention penalty" found by Gordon et al., 1993). Does an utterance with multiple pronouns specify a model where several entities have varying salience (see Reichman's, 1985, proposal) or else a model with a single proposition encompassing both pronominalised entities (e.g. specifying a player's relationship with the ball)? The current study was not designed to systematically elicit "it", so no predictions were made and "it" was not analysed. A complicating issue is that just as a speaker may have several purposes, a discourse may be about several different topics at once. It may be appropriate to model *levels* of topical or given information on a scale similar to the "focal scale" that has been proposed for new or foregrounded information by Wilson and Sperber (1979). An untested aspect of centering theory is the cognitive extent of the center itself.

Centering, as a process, enables a speaker to guide an addressee's attention. A speaker takes a perspective on the situation and describes it accordingly, attending to different elements at different moments, depending on her interest and purpose. Her speech plays over her mental model of the situation like a flashlight, illuminating pieces of it for her listener, who is creating his own similar mental model. Centering, as a framework, integrates notions of topic, given and new, salience, shared knowledge, accessibility of discourse entities in working memory, and the resolution of referring expressions. But centering, as a discourse-processing algorithm, remains underspecified. Centering algorithms have been motivated in part by the need to build natural language understanding programs that can interpret pronouns (e.g. Brennan et al., 1987; Kameyama, 1986; Sidner, 1981). These algorithms have limitations. For instance, with respect to the one discussed here, there could be other ranking schemes for the forward center list, corresponding to other parameters that influence an addressee's attention. In addition, only those entities realised as noun phrases are considered as candidates for the backward center, and other potentially salient propositions such as verb phrases are neglected. The algorithm takes a very narrow view of discourse segmentation and locality as the backward center of an utterance must be realised in the immediately

preceding utterance. Existing implementations do not model forward centers realised implicitly via inference.

Finally, in communication between autonomous agents, there will always be occasions where both local focusing and inferencing processes fail. A pronoun may not be resolved by the end of a sentence (Garnham, Oakhill, & Cruttenden, 1992; Greene et al., 1992), and the referent a speaker proposes by a pronoun may be genuinely ambiguous to an addressee. When that happens, the speaker and addressee have other options, such as marking the problem and taking additional turns to repair or clarify it. Pronominalisation and pronoun resolution are not modelled in their entirety as syntactic algorithms.

Manuscript received August 1993
Revised manuscript received July 1994

REFERENCES

- Anderson, A., Garrod, S.C., & Sanford, A.J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology*, 35A, 427-440.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
- Au, T.K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language*, 25, 104-122.
- Bock, J.K. (1977). The effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior*, 16, 723-734.
- Bock, J.K. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.
- Bock, J.K., & Warren, R.K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47-67.
- Bolinger, D. (1977). *Meaning and form*. London: Longman.
- Bolinger, D. (1979). Pronouns in discourse. In T. Givón (Ed.), *Syntax and semantics 12: Discourse and syntax*. New York: Academic Press.
- Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press.
- Brennan, S.E., Friedman, M.W., & Pollard, C.J. (1987). A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, pp. 155-162. Stanford, CA.
- Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler & D. Ladd (Eds.), *Prosody: Models and measurements*. New York: Springer-Verlag.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In C. Li (Ed.), *Subject and topic*, pp. 25-55. New York: Academic Press.
- Chafe, W. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins.
- Clark, H.H., & Haviland, S.E. (1977). Comprehension and the given/new contract. In R. Freedle (Ed.), *Discourse production and comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In

- Clark, H.H., & Sengul, C.J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition*, 7, 35-41.
- Dahl, D.A. (1986). Focusing and reference resolution in PUNDIT. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, pp. 1083-1093. Philadelphia, PA: AAAI-86.
- Dell, G.S., & Brown, P. (1991). Mechanisms for listener adaptation in language production. In D.J. Napoli & J.A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman*, pp. 105-129. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Ferguson, C.A. (1983). Sports announcer talk: Syntactic aspects of register variation. *Language in Society*, 12, 153-172.
- Fletcher, C.R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, 23, 487-493.
- Fowler, C.A., & Housum, J. (1987). Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489-504.
- Fox, B. (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.
- Frank, E.P. (1985). *Referential choice and focus of attention in narratives*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Frazier, L., Taft, L., Roeper, T., Clifton, C., & Erlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory and Cognition*, 12, 421-430.
- Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126-152.
- Garnham, A., & Oakhill, J. (1992). Disclosure processing and text representation from a "mental models" perspective. *Language and Cognitive Processes*, 7, 193-204.
- Garnham, A., Oakhill, J., & Cruttenden, H. (1992). The role of implicit causality and gender cue in the interpretation of pronouns. *Language and Cognitive Processes*, 7, 231-255.
- Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, 21, 347-353.
- Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition*, 32, 99-156.
- Gernsbacher, M.A., & Hargreaves, D.J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27, 699-717.
- Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.
- Gordon, P.C., Grosz, B.J., & Gilliom, L.A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311-347.
- Green, G.M. (1980). Some wherefores of English inversions. *Language*, 56, 582-601.
- Greene, S.B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 266-283.
- Grober, E.H., Beardsley, W., & Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition*, 6, 117-133.
- Grosz, B.J. (1977). The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA.
- Grosz, B.J., & Sidner, C.L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12, 175-204.
- Grosz, B.J., Joshi, A.K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA.
- Grosz, B.J., Joshi, A.K., & Weinstein, S. (1986). *Towards a computational theory of discourse interpretation*. Unpublished manuscript, CSLI, Stanford University, Stanford, CA.

- Grosz, B.J., Joshi, A.K., & Weinstein, S. (in press). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*.
- Guindon, R. (1985). Anaphora resolution: Short term memory and focusing. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pp. 218-227, Chicago, IL.
- Gundel, J.K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.
- Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hirschberg, J., & Pierrehumbert, J. (1986). The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pp. 136-144, New York.
- Hobbs, J.R., Stickel, M.E., Appelt, D.E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, 63, 69-142.
- Horn, L. (1986). Presupposition, theme, and variations. In *Papers from the Twenty-Second Regional Meeting of the Chicago Linguistic Society*, pp. 168-192, Chicago, IL.
- Hudson-D'Zmura, S. (1988). *The structure of discourse and anaphor resolution: The discourse center and the roles of nouns and pronouns*. Unpublished PhD dissertation, University of Rochester, Rochester, NY.
- Hudson-D'Zmura, S., & Tanenhaus, M.K. (in press). Assigning antecedents to ambiguous pronouns: The role of the center of attention as the default assignment. In M. Walker, E. Prince, & A. Joshi (Eds), *Centering in discourse*. Oxford: Oxford University Press.
- Hudson, S.B., Tanenhaus, M.K., & Dell, G.S. (1986). The effect of the discourse center on the local coherence of a discourse. In *Proceedings of the 8th Annual Conference of the Cognitive Science Society*, pp. 96-101. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Jarvella, R.J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10, 409-416.
- Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.
- Joshi, A.K., & Weinstein, S. (1981). Control of inference: Role of some aspects of discourse structure—Centering. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 385-387, Vancouver, BC.
- Kameyama, M. (1986). A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pp. 200-206, New York.
- Kintsch, W., & Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Linde, C. (1979). Focus of attention and the choice of pronouns in discourse. In T. Givón (Ed.), *Syntax and semantics 12: Discourse and syntax*. New York: Academic Press.
- Malt, B.C. (1985). The role of discourse structure in understanding anaphora. *Journal of Memory and Language*, 24, 271-289.
- Marslen-Wilson, W., Levy, E., & Tyler, L.K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R.J. Jarvella & W. Klein (Eds), *Speech, place, and action*. Chichester: John Wiley.
- Morrow, D.G. (1985a). Prominent characters and events organize narrative understanding. *Journal of Memory and Language*, 24, 304-319.
- Morrow, D.G. (1985b). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language*, 24, 390-404.
- Olson, D.R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review*, 77, 257-273.
- interpretation of discourse. In P.R. Cohen, J. Morgan, & M.E. Pollack (Eds), *Intention in communication*. Cambridge, MA: MIT Press.
- Polanyi, L., & Scha, R.J.H. (1986). Syntactic and semantic aspects of discourse structure. In L. Polanyi (Ed.), *The structure of discourse*. Norwood, NJ: Ablex.
- Prince, E.F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radiating pragmatics*. New York: Academic Press.
- Prince, E.F. (1992). The ZPG letter: Subjects, definiteness, and information-status. In S. Thompson & W. Mann (Eds), *Discourse description: Diverse analysis of a fund raising text*. Philadelphia, PA: John Benjamins.
- Reichman, R. (1985). *Getting computers to talk like you and me*. Cambridge, MA: MIT Press.
- Reinhart, T. (1982). Pragmatics and linguistics: Analysis of sentence topics. *Philosophica*, 2, 53-94.
- Sag, I., & Hankamer, J. (1984). Toward a theory of anaphoric processing. *Linguistics and Philosophy*, 7, 325-345.
- Selkirk, E.O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Sidner, C.L. (1979). *A computational model of co-reference comprehension in English*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Sidner, C.L. (1981). Focusing for interpretation of pronouns. *American Journal of Computational Linguistics*, 7, 217-231.
- Sidner, C.L. (1983). Focusing in the comprehension of definite anaphora. In M. Brady, R.C. Berwick (Eds), *Computational models of discourse*. Cambridge, MA: MIT Press.
- Sowa, J.F. (1984). *Conceptual structures: Information processing in mind and machine*. Reading, MA: Addison-Wesley.
- Stark, H.A. (1986). *Keeping track of characters in narrative*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Suri, L.Z., & McCoy, K.F. (1993). Focusing and pronoun resolution in particular kinds of complex sentences. In *Proceedings of the Workshop on Centering Theory in Natural Occurring Discourse*, Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, PA.
- Suri, L.Z., & McCoy, K.F. (1994). RAFT/RAPR and centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics*, 20, 301-317.
- Tomlin, R.S. (1987). Linguistic reflections of cognitive events. In R.S. Tomlin (Ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins.
- Walker, M.A. (1989). Evaluating discourse processing algorithms. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 251-261, Vancouver, BC.
- Webber, B.L. (1979). *A formal approach to discourse anaphora*. New York: Garland Press.
- Wilson, D., & Sperber, D. (1979). Ordered entailments: An alternative to presupposition theories. In C.K. Oh & D.A. Dineen (Eds), *Syntax and semantics 11: Presupposition*. New York: Academic Press.
- Zubin, D. (1979). Discourse function of morphology: The focus system in German. In T. Givón (Ed.), *Syntax and semantics 12: Discourse and syntax*. New York: Academic Press.