Discourse Deixis and Discourse Processing

Bonnie L. Webber
University of Pennsylvania, bonnie@inf.ed.ac.uk

Follow this and additional works at: http://repository.upenn.edu/cis_reports

Recommended Citation
http://repository.upenn.edu/cis_reports/610


This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cis_reports/610
For more information, please contact repository@pobox.upenn.edu.
Discourse Deixis and Discourse Processing

Abstract
Computational approaches to discourse understanding have a two-part goal: (1) identifying those aspects of discourse understanding that require process-based accounts, and (2) characterizing the processes and data structures they involve. To date, in the area of reference, process-based accounts have been developed for reference via anaphoric pronouns and via definite descriptors. In this paper, I propose and argue for a process-based account of deictic reference in text. This account adds precision to common notions of discourse entity, discourse segment and focus and to relationships between the three.

Comments
Acknowledgements: This research was partially supported by U.S. Army grants DAA29-84-9-0027, DAA29-84-K-0061, DARPA grant NOOO14-85-K-0018, NSF grants MCS-8219196-CER, IRI84-10413-AO2
Abstract

Computational approaches to discourse understanding have a two-part goal: (1) identifying those aspects of discourse understanding that require process-based accounts, and (2) characterizing the processes and data structures they involve. To date, in the area of reference, process-based accounts have been developed for reference via anaphoric pronouns and via definite descriptors. In this paper, I propose and argue for a process-based account of deictic reference in text. This account adds precision to common notions of discourse entity, discourse segment and focus and to relationships between the three.

*This work was partially supported by ARO grant DAA29-884-9-0027, NSF grant MCS-8219116-CER and DARPA grant N00014-85K-0018 to the University of Pennsylvania. It was done while the author was on sabbatical leave at Medical Computer Science, Stanford University in Spring 1988. Thanks go to Jerry Hobbs, Mark Steedman, James Allen, Becky Passonneau, Ethel Schuster, the Situation Semantics Group at CSLI, Stanford (Craig Roberts, Annie Zaenen, Mark Gawron, John Nerbonne & Steve Levinson), and the Sloan Lunch Groupies at Penn for their helpful comments. A preliminary version of this paper was presented at the 26th Annual Meeting of the Association for Computational Linguistics, Buffalo NY, June 1988.
1 Introduction

This paper represents another step in the enterprise of characterizing linguistic phenomena and their implications for theories of discourse. The particular phenomenon to be discussed is textual use of this and that. Such usage is very common, although apparently not encouraged: Strunk & White in their Elements of Style (3rd Edition) warn writers:

This. The pronoun this, referring to the complete sense of a preceding sentence or clause, cannot always carry the load and so may produce an imprecise statement.

However, despite this warning, there is ample data for examining textual use of this and that and its implications for theories of discourse.

The theories of discourse that seem to me most cognitively plausible are ones that take its most common features as following from two separate factors:

1. speakers' intentions that take more than one sentence to describe;
2. independent attentional mechanisms that enable information to be conveyed linguistically with minimal lexical effort.

Speakers can, in part, manipulate listeners' attention through their text, thereby changing the context in which subsequent text will be understood. (See Figure 1.) In this view, an account of this and that addresses the pair of questions:

- How does text change context in such a way that context provides an interpretation for this and that?¹
- How are instances of this and that in text processed such that their intended interpretations are found?

In this paper, I try to provide an initial answer to both these questions, and in doing so, place some additional requirements on theories of discourse.

The paper is organized as follows: In the next section (Section 2), I review the constructs that current discourse theories provide for describing context and

¹Steve Isard [Isard75] is the first person I know of to have asked this question (in general, not just for deictics), thereby going beyond the question standardly addressed by formal semanticists like Montague of how context affects interpretation.
changes in context. In Section 3, I try to give a more precise characterization of the phenomenon that Strunk & White loosely describe as “referring to the complete sense of a preceding sentence or clause”. In Section 4, I then propose an account of how this and that achieve their intended interpretations, and finally, in Section 5, I describe both the advantages of this approach and some complicating factors.

2 Context and Context Change

Theories of discourse understanding give us three different sorts of constructs for describing context and its changes: discourse entities, discourse segments and attention. Each has a rather different history and purpose, which I will describe briefly in turn.

2.1 Discourse Entities

Discourse entities seem to have been first introduced by Lauri Kartunnen in 1976 [Kart76] under the name discourse referent, to provide a uniform way of
explaining what it is that noun phrases and pronouns in a discourse refer to. That is, noun phrases should not be taken to refer to things in the world, but rather to (mental) entities in a listener's evolving model of the discourse, often called simply a discourse model [Webb82, Garn87]. Discourse entities may correspond to something in the outside world, but they do not have to. Minimally, theories hold that discourse entities are evoked into a listener's discourse model in response to the listener's interpreting noun phrases in a text, but theories differ somewhat in what they take to be the precise relationship between text and discourse entities. In theories that posit discourse entities, the term refer is often used to mean "refer in the discourse model" and not "refer in the outside world". Referent then means the entity in the model picked out by a linguistic expression. This is the usage I have adopted in this paper.

The basic features of a discourse entity are that (a) it is a constant within the current discourse model and that (b) one can attribute to it, inter alia, properties and relationships with other entities. (It is for this reason that Bill Woods once called them "conceptual coathooks"). In some theories, different parts of the discourse model (often called spaces) correspond to portions of the text or discourse with different modalities, including hypotheticals, counterfactuals, beliefs of different agents, etc. Depending on which space in the model is currently active (under construction), the same noun phrase (e.g., "the key") or pronoun (e.g., "it") may refer to very different discourse entities.

2.2 Discourse Segments

The second construct available for describing context and changes in context is the discourse segment. While discourse segmentation is generally taken to be a chunking of a linguistic text into sequences of related clauses or sentences, as James Allen has noted:

... there is little consensus on what the segments of a particular discourse should be or how segmentation could be accomplished. One reason for this lack of consensus is that there is no precise definition of what a segment is beyond the intuition that certain sentences naturally group together [Allen87, p. 398-9]

Beyond this common intuition, intuitions vary: among computational linguists, Grosz & Sidner [GS85] have taken a discourse segment to be a chunk of text that expresses a common purpose (what they have called a discourse segment purpose) with respect to the speaker's plans; Hobbs et al. [HMSE88] have taken a discourse segment to be a chunk of text that has a common meaning; while Nakhimovsky [Nakh88], considering only narrative, has taken a discourse
segment to be a chunk of text that describes a single event from a single perspective. Theories also differ as to their minimal discourse segment. Hobbs takes it to be a sentence, and Polanyi [Pola86], a clause. Grosz & Sidner seem to take a sentence as the minimal segment needed to express a single purpose, but do not assume that every sentence constitutes a distinct discourse segment.

In general, the notion of discourse segment has been defined recursively, such that a discourse segment is either

1. a minimal discourse segment
2. a sequence of discourse segments.

The resulting segmentation of a text will therefore be describable as a set of embedding structures or trees. A simple discourse structure is illustrated as an embedding structure in Figure 2a, and as a tree in Figure 2b.

As Passonneau has pointed out (personal correspondence), this ignores the possibility of interpreting a stretch of text as belonging to two adjacent segments in a sequence, serving essentially as a transition between them. Including this possibility complicates what it would mean to have a sequence of discourse segments, but would not alter the recursive nature of the definition itself.
If one wants to take discourse segmentation as a factor in context change, then it cannot be done post facto, after a text has been read and understood: it must be doable on-line. While characterizing the features of this process is an area of active research (cf. [GS85], [HL87], [Cohen87], [Reich]), for the purposes of this paper, I will assume that it can be described in 'tree'-terms as follows:

- Multi-clause segments correspond to non-terminal nodes of the tree, and single-clause segments to terminal nodes.\(^3\)
- To start, an initial root node is established. (This will not necessarily be the final root node after segmentation is complete.)
- Clauses are processed in linear order.
- A new single-clause segment (terminal node) is added to the growing tree in one of two ways: by attachment or by adjunction:
  - A node is attached to an existing non-terminal node by making it the new rightmost daughter of that node. (See Figure 3.)
  - A node \(C_j\) is adjoined to an existing non-terminal node \(S_i\) by creating a new non-terminal node \(S_k\) and attaching \(S_i\) as its left daughter and \(C_j\) as its right daughter. If \(S_i\) was the root node, \(S_k\) becomes the new one (See Figure 4a.) If \(S_i\) was an internal node, then \(S_k\) replaces \(S_i\) in the tree. (See Figure 4b.)
  - A node \(C_j\) is adjoined to an existing terminal node \(C_i\) by creating a new non-terminal node \(S_k\), inserting it in the tree in place of \(C_i\), attaching \(C_i\) as its left daughter and \(C_j\) as its right daughter. (See Figure 4c.)

Where and how a new clause is attached depends on what clauses it "naturally groups together with", in terms of its meaning, purpose, viewpoint, etc., in ways

\(^3\)This simplifies the tree by eliminating single branching structures of single-clause segment nodes going to clause nodes. Unfortunately, this presentational simplification is at the expense of somewhat complicating the algorithm.
(a) adjunction of $C_3$ to root node $S_1$

(b) adjunction of $C_3$ to internal node $S_2$

(c) adjunction of $C_3$ to terminal node $C_2$

Figure 4: Addition of nodes by Adjunction
that are currently the focus of active research, as noted above. 4

In any case, my assumption that on-line discourse segmentation resembles this type of tree-construction process has the following consequences:

- It is only segments on the right frontier of the tree that are involved in the attachment or adjunction of a new clause/segment. (For example, in Figure 2, at the point of processing clause $C_{j+1}$, the right frontier comprises segments \{$S_{k2j}$, $S_{k21}$, $S_{k2}$, $S_k$\}.) These are the segments whose meaning, purpose, etc. the listener can be taken as actively attending to. 5

- Adjoining a new clause/segment may involve tree restructuring as the existing daughters of a node are grouped together as a segment unto themselves. Thus the discourse segmentation may change over time.

Notice that it is still a separate question as to how segmentation relates to context, in the sense of Figure 1. For example, if segmentation simply produces larger chunks of meaning, it may not affect context beyond simply adding to it. On the other hand, if it imposes a structure on context it may affect the use and interpretation of other linguistic devices. For example, Grosz [Grosz81], Reichman [Reich] and Grosz & Sidner [GS85] have argued that discourse segmentation indirectly constrains the interpretation of definite noun phrases. Other researchers ([Nakh88,Webb87,Webb88]) have argued that through a Temporal Focus associated with a segment, segmentation constrains the interpretation of tense. Now in this paper, I will argue that the mental correlates of discourse segments have another effect on context - themselves providing referents for the deictic pronouns this and that.

2.3 Attention

The third sort of construct provided by theories of discourse for describing changes in context is the listener's attention. This appears in theories in various

4By assuming that a new clause/segment is immediately attached into a growing structure, I am ignoring the interesting possibility that "growing" constituents might in part be separated from deciding where to attach them in the tree, as in Marcus's deterministic parser [Marcus79].

5I believe this corresponds to the intuition behind Grosz & Sidner's stack representation of Attentional State, cf. Section 2.3. (Any directed path through a tree from root to tip can be represented as a stack.) In any case, the rest of a discourse segmentation tree may well blur out of a listener's awareness, with the result that the tree as an instantiated data structure may actually only be an artifact of post hoc analysis. In fact, it is not even clear how much of a right frontier a listener can be taken as actively attending to at any one time, such that there is some linguistic property that all segments on this frontier share. My data here on deixis can be interpreted as arguing for three or four segments, but further study on this and other aspects of attention in discourse are clearly needed.
guises:

- Foreground/Background [HoppT80, Givon87]
- States of Activation [Chafe87]
- Focusing [Grosz81]
- Attentional State [GS85]
- Centering [GJW83]

All of these notions have been explicitly proposed, at least in part, to account for patterns of concept verbalization – for example, when the pronunciation of concept descriptions can be attenuated, when concepts can be specified using explicit pronouns or zero-anaphors, when an unmodified definite noun phrase can be used to refer to a concept, when particular intonation structures and/or marked syntactic constructs are appropriate, etc. Three of these notions I will now describe in more detail, as they will be relevant later in the paper.

Chafe proposes [Chafe87] that a concept (objects, events, and properties) may have one of three activation states:

**active** – A concept that is currently “in a person’s focus of consciousness” because of its recent explicit verbalization. In such a state, a concept can be verbalized in a very reduced way (e.g., with an unstressed pronoun, a zero anaphor, etc.)

**semi-active** – A concept in a person’s “peripheral consciousness” (i.e., of which s/he has “background awareness”) either because of belonging to “a set of expectations associated with a schema” evoked by a text or because of having been active but not being “refreshed”. In this state, a concept can be verbalized as an unmodified definite noun phrase.

**inactive** – A concept in a person’s long-term memory that is neither active nor semi-active. Concepts in this state will be verbalized with indefinite noun phrases or modified definite noun phrases.

Thus in Chafe’s model, a listener’s attention to concepts reflects both their recency of mention and a block structure associated with whatever schema is currently active.

In Grosz & Sidner’s model [GS85], a listener’s attention correlates with the perceived structure of the discourse. In particular, there is a focus space associated with each discourse segment, as well as a *discourse segment purpose*
or DSP, as noted earlier. Corresponding to a tree hierarchy of DSPs (which they call *Intentional Structure*), there is a *stack* of focus spaces, representing the listener's *Attentional State*. The focus spaces on the stack correspond to the discourse segments whose DSPs are on the *right frontier of Intentional Structure*. Given this, a concept's activation correlates with its focus space's position in the stack: ones in the focus space at the top of the stack are most "activated". Hence, an unmodified noun phrase is more likely to pick up its referent from a segment closer to the top of the stack than one further down. Picking up a referent from further down in the stack may, in fact, indicate that the segments higher up the stack can now be taken as "closed", with attention shifting back to a more inclusive segment. (Other researchers relating some notion of *activation* with segmentation are [Reich] and [Fox87].)

Grosz, Joshi & Weinstein [GJW83] attempt to account for more local properties of a listener's *attention*. In their *centering* theory, at any point in a discourse following its initial clause, there is a distinguished entity called the *backwards-looking center* (Cₘ) and an ordered list of entities, including Cₘ, called *forward-looking centers* (Cₖ). Cₖ includes all entities that have been mentioned either explicitly or implicitly in the previous clause, and their ordering reflects, in large part, their syntactic position in that clause. Because of this recency of mention, a listener is assumed to be actively aware of all of them. The Cₘ is distinguished from the other members of Cₖ by being the one the speaker is assumed to be talking about (focused on), while the others are what the speaker may go on to talk about. Because of this, while it will be possible to refer pronominally to any member of Cₖ by virtue of recency, a subsequent clause will sound most natural (i.e., will be most easily processed) if

1. the current Cₘ, if referred to, is referred to pronominally;
2. any other member of the current Cₖ is referenced pronominally only if the current Cₘ is not mentioned in the clause or if it is referenced pronominally as well.

As I will discuss further in Section 4, what is for me significant about this *centering* theory is that (1) there is one entity, Cₘ, that is a natural referent of a pronoun in the next clause and (2) there is a small distinguished set of entities Cₖ that may be referenced pronominally as well.

To interpret this research on *attention* in terms of *text* and *context*, what it says, in part, is that context should not be considered a homogeneous concept soup, but rather, a dynamically changing structure of concepts, in which some are more accessible than others at a given time and may permit more inferential operations than others in ways that reflect not just the organization of long-term memory but local and global properties of the particular *text*.
3 The Phenomenon of *this* and *that*

Having considered what theories of discourse provide for describing changes in context, I want to move on to characterize more precisely what it means, in Strunk & White's terms, "to refer to the complete sense of a preceding sentence or clause", as in the following two examples.

**Example 1**

It's always been presumed that when the glaciers receded, the area got very hot. The Folsum men couldn't adapt, and they died out. *That's* what is supposed to have happened. It's the textbook dogma. But it's wrong. They were human and smart. They adapted their weapons and culture, and they survived.

**Example 2**

The tools come from the development of new types of computing devices. Just as we thought of intelligence in terms of servomechanism in the 1950s, and in terms of sequential computers in the sixties and seventies, we are now beginning to think in terms of parallel computers, in which tens of thousands of processors work together. *This* is not a deep, philosophical shift, but it is of great practical importance, since it is now possible to study large emergent systems experimentally. [Hill88, p.176]

In neither of these examples, do the pronouns *this* and *that* refer to discourse entities introduced into the listener's discourse model by noun phrases. The stuff of their interpretation – which comes from material introduced clausally – is the phenomenon I shall now try to characterize more precisely.

The first piece of evidence I want to introduce consists of distributional statistics. I took six texts written by six different authors and tabulated instances of subsequent reference that used either of the pronouns *it*, *this* or *that*. Of 96 instances of pronominal reference to clausal material found in these texts, only 15 (~16%) used the pronoun *it* while the other 81 (~84%) used either *this* or *that* (19 instances of *that* and 62 instances of *this*). On the other hand, looking at all instances of pronominal reference to discourse entities evoked by

noun phrases, of 81 such references, 79 (~98%) used it while only 2 (~2%) used this or that. (Passonneau [Schiff84a, Schiff84b] reports a similar distribution for it and that in spoken discourse. She does not report results for this.) Given such results, it makes sense to call the phenomenon discourse deixis.

There is an important reason for confining my investigation to written (primarily objective) expositions rather than including spoken texts as well. That is that spoken text makes additional use of this and that in ways that objective written text rarely does. For example, in first-person accounts (which are very common in speech, less so in written objective texts), narrators use this and that in ways that reflect some subjective notion of nearness to or distance from themselves [AnKe85]. In addition, spoken texts employ stress — in particular, contrastive stress — which is well-known for altering what it is that a referring phrase can be used to refer to. Since the purpose of this paper is to argue what is necessary for an account of this and that, not for what is sufficient, it seems to me permissible to confine my investigation to written text.

The second piece of evidence concerns what it is that the pronouns this and that are referring to. Consider the following example:

Example 3

There’s two houses you might be interested in:

House A is in Palo Alto. It’s got 3 bedrooms and 2 baths, and was built in 1950. It’s on a quarter acre, with a lovely garden, and the owner is asking $425K. But that’s all I know about it.

House B is in Portola Vally. It’s got 3 bedrooms, 4 baths and a kidney-shaped pool, and was also built in 1950. It’s on 4 acres of steep wooded slope, with a view of the mountains. The owner is asking $600K. I heard all this from a real-estate friend of mine.

Is that enough information for you to decide which to look at?

What I want to show is that in this passage, that in the second paragraph does not refer to House A (although all instances of it do): rather it refers to the packet of information about House A presented there. Similarly (all) this in the third paragraph does not refer to House B (although again, all instances of it do): rather it refers to the packet of information about House B presented there. That in the fourth paragraph refers to the combined packet of information about the two houses taken together.

That in each case it is a packet of information that this and that are accessing and not the houses, can be seen by interleaving the two descriptions, a technique often used in discourse when comparing two items:
Example 4

There’s two houses you might be interested in:
House A is in Palo Alto, House B in Portola Vally. Both were built in 1950, and both have 3 bedrooms. House A has 2 baths, and B, 4. House B also has a kidney-shaped pool. House A is on a quarter acre, with a lovely garden, while House B is on 4 acres of steep wooded slope, with a view of the mountains. The owner of House A is asking $425K. The owner of House B is asking $600K.

# That’s all I know about House A. #

This I heard from a real-estate friend of mine.

Is that enough information for you to decide which to look at?

Here houses A and B have been described together, and the failure of that and this to refer successfully in the second paragraph to the same referents as in Example 3 indicates that it is not the houses being referred to or what one has learned about them from the text. Rather, this and that must be interpreted as specifying what was said about each house, viewed as a thing unto itself — a packet’s worth of information — that is, the mental correlate of a discourse segment, which I will call a sequent. In Example 4, the packaging is different than in Example 3, even though the content is the same. In Example 4, there is only one packet of information about both houses together, hence the failure of the individual pointers. The only deictic that works is the final that, which successfully refers to the information conveyed about both houses together.

As noted in Section 2, discourse segment is usually taken to be a recursive structure. Thus my third piece of evidence that this and that specify sequents comes from examples in which successive instances of this and/or that appear to derive their interpretation from segments at different levels of embedding, as in the following example.

Example 5

...it should be possible to identify certain functions as being unnecessary for thought by studying patients whose cognitive abilities are unaffected by locally confined damage to the brain. {1} For example, binocular stereo fusion is known to take place in a specific area of the cortex near the back of the head. {2} Patients with damage to this area of the cortex have visual handicaps but {3} [they] show no obvious impairment in their ability to think. {3} This suggests that stereo fusion is not necessary for thought. {1} This is a simple

---

7I am using the term sequent in very broad analogy to its use in logic to denote a structured collection of propositions in which one subset “proves” the other. I would not be averse to alternative suggestions.
example, and the conclusion is not surprising. . . . [Hill88, p. 185]

Here brackets have been added to indicate discourse segments, with subscripts indicating the depth of embedding. The most likely interpretation of the first this is the observation that visual cortex-damaged patients have visual handicaps but no impairment to their thinking abilities (i.e., the interpretation of segment 2), while the most likely interpretation of the second this is the whole “brain damage” example (i.e., the interpretation of segment 1).³

Here brackets have been added to indicate discourse segments, with subscripts indicating the depth of embedding. The most likely interpretation of the first this is the observation that visual cortex-damaged patients have visual handicaps but no impairment to their thinking abilities (i.e., the interpretation of segment 2), while the most likely interpretation of the second this is the whole “brain damage” example (i.e., the interpretation of segment 1).³

The final piece of evidence I want to present concerns the range of features that must be associated with the mental representation of a discourse segment for subsequent reference by this or that. These include: its speech act (B1 in Example 6); its form of expression (B2); the event it specifies, if it specifies an event (B3); the action it specifies, if the event it specifies involves an action (B4); as well as its complete sense (B5).

Example 6

A: Hey, management has promoted Fred to second vice president.
B1: That's a lie.
B2: That's a funny way to describe the situation.
B3: When did that happen?
B4: That's a weird thing for them to do.
B5: Oh, that's why his wife is so happy.

4 Resolving Discourse Deictics

Having characterized the phenomenon I am calling discourse deixis and what context minimally has to provide for its interpretation (i.e., what I am calling sequents—mental correlates of discourse segments), I want to move on to discuss the interpretive process itself. First I will discuss constraints on the process, and then how it differs from the process of interpreting anaphoric pronouns. Finally I will discuss where in the process differences between this and that come into play.

³Segment 3 alone might be thought to provide the correct interpretation if the first this were followed by something like “is obvious when they are asked to solve word problems presented orally.” (cf. Section 4) It has also occurred to me that Example 5 can be thought of as a topic-comment construction in which the topic takes more than one clause to describe.
4.1 Interpretive Constraints and their Application

In this section, I will argue that the process that provides interpretations for this and that is subject to two separate sets of constraints – one attentional, the other semantic – namely,

1. the interpretation of this and that must come from a sequent corresponding to a discourse segment on the right frontier (cf. Section 2);
2. their interpretation must be compatible with that of their matrix clause.

My evidence that the only sequents that can provide interpretations for discourse deictics are ones associated with discourse segments on the right frontier consists of (a) it being true of the 81 sequent-referring instances of this and that found in the six texts mentioned above and (b) the oddity of examples like the following variation of Example 3. (The clauses are numbered for later discussion.)

Example 3'

(1) There’s two houses you might be interested in:
   (2) House A is in Palo Alto. (3) It’s got three bedrooms and two baths, and was built in 1950. (4) It’s on a quarter acre, with a lovely garden, and (5) the owner is asking $425K.
   (6) House B is in Portola Vally. (7) It’s got three bedrooms, four baths and a kidney-shaped pool, and (8) was also built in 1950. (9) It’s on 4 acres of steep wooded slope, with a view of the mountains. (10) The owner is asking $600K. (11) I heard all this from a real-estate friend of mine. (12) #But that’s all I know about House A.
   (13) Is that enough information for you to decide which to look at?

What is at issue is the interpretation of that in clause 12. The rest of clause 12 constrains the interpretation of that to be information about House A. However its position in the text, assuming that it is unstressed, is only compatible with its being interpreted in one of very few ways, including:

- something related to clause 11, as in “But that’s all she said.” (where that is interpretable as referring to the same thing as “all this about House B that I heard from a real-estate friend of mine”);
- something related to the interpretation of clauses 2-11 (the information regarding both houses, similar to the perceived interpretation of that in
Info on two houses of interest

Figure 5: Discourse Segmentation at the point of processing "But that's all ..."

clause 13.

Schematically, one might represent the discourse segmentation at the point in the processing that roughly as in Figure 5. The oddity of Example 3' comes from the conflicting demands of text position and clause predication in the process of resolving that.

Again let me emphasize that I am only considering written text and unstressed instances of this and that. It is well-known that stressing a pronoun can shift its preferred referent. In the case of clause 12, stressing that, reinforced by information conveyed by the rest of the sentence, allows it to be interpreted as the block of information about House A, even though that sequent is no longer being attended to.

Notice that even if it is true that unstressed this and that must be identified with a discourse segment on the right frontier, there is still an ambiguity as to which segment. To see this, consider the first part of Example 5 as a “discourse completion task”.

Example 5

...it should be possible to identify certain functions as being unnecessary for thought by studying patients whose cognitive abilities
are unaffected by locally confined damage to the brain. For example, binocular stereo fusion is known to take place in a specific area of the cortex near the back of the head. Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think. This ....

At this point in the discourse, there are many possible ways of completing the last sentence, among them –

a. *This* is obvious when they are asked to solve word problems presented orally.

b. *This* suggests that stereo fusion is not necessary for thought.

c. *This* is only a simple example, and the conclusion is not surprising.

In (a), it makes most sense to interpret *this* as specifying the proposition that patients with damage to the particular area of the cortex near the back of the head show no obvious impairment in their ability to think. In (b) *this* is identifiable with the more inclusive claim that patients with damage to the particular area of the cortex near the back of the head have visual handicaps but show no obvious impairment in their ability to think. Finally, in (c) it makes most sense to interpret *this* more inclusively as the entire example about binocular stereo vision. How the listener chooses to interpret the deictic depends on what is compatible with the meaning of the rest of the sentence. As with other types of ambiguity, there may be a default preference for one particular form of construal over the others (cf. [CS85, Steed88]) but it is easily overridden by context.

This ambiguity as to what sequent the deictic pronouns *this* and *that* are identified with seems to me very similar to the ambiguity associated with what some linguists (cf. [Lyons79]) see as the logically prior use of deixis for pointing within a shared physical context, as in Figure 6.

In Figure 6, *this* could be interpreted as either the business, the pictures, or the physical gallery. Both Quine [Quine71] and Miller [Miller82] have observed in this regard that all pointing is ambiguous: the intended demonstratum of a pointing gesture can be any of the infinite number of points "intersected" by the gesture or any of the structures encompassing those points. (Or, one might add, any interpretation of those structures.) The ambiguity here as to how large a segment on the right frontier is providing an interpretation for *this* or *that* is very similar.

9Presumably Al Junior will have enough context to resolve this more precisely, or he will be smart enough to ask.
As I noted at the beginning of this section, a second constraint on the interpretation of a discourse deictic is that it be compatible with the interpretation of its matrix clause. The interpretation of some sequent as a whole may not be so compatible. The process that provides interpretations for discourse deictics must be able to coerce [HMSE88,MS88] a sequent to one of its properties, as in Example 6 (repeated here).

Example 6

A: Hey, management has promoted Fred to second vice president.
B1: That's a lie.
B2: That's a funny way to describe the situation.
B3: Did that happen before or after the scandal broke?
B4: That's a weird thing for them to do.
B5: Oh, that's why his wife is so happy.

Moreover, since this and that can also be interpreted as some discourse entity that is explicitly part of a sequent, as in Examples 7 and 8 below, the process that interprets discourse deictics must also be able to do such “whole-part” coercions as well.
Example 7

A: Hey, management has promoted Fred to second vice president.
B1: Oh, that's who's replacing Harry.
B2: That's a very important post.
B3: Oh!, that's who we have to blame then.

Example 8

A: Bob got a 90 on the exam. Betsy got a 95, and Gerry got a 93.
B: Well, that's who I would have expected to do well.

4.2 Processing Differences between Anaphors and Deictics

If one looks at the linguistics literature, it is clear that many linguists believe that deixis is different from anaphora. Ehlich [Ehli82] makes the clearest statement of what he takes the difference to be. Deictics he sees as:

- a linguistic instrument for achieving focusing of the hearer's attention towards a specific item which is part of the respective deictic space

whereas anaphors are

- a linguistic instrument for having the hearer continue (sustain) a previously established focus towards a specific item on which he has oriented his attention earlier.

As a linguist though, Ehlich does not give a precise definition of "deictic space", "previously established focus" or what it means to "orient one's attention towards an item". However, the vocabulary of processing discussed here can be used to give substance to Ehlich's distinction.

Following [GJW83], I assume that in processing the bulk of anaphoric pronouns in text\(^\text{10}\), the language processor makes use of a distinguished entity \(C_b\) and a small set of only slightly less distinguished entities \(C_J\). It offers the \(C_b\) as the most natural referent for a pronoun and after that,\(^\text{11}\) the members of \(C_J\). If these local, highly salient assignments fail, conscious reasoning begins or

\(^{10}\text{not the ones linguists write papers about}\)

\(^{11}\text{either because }C_b\text{ is incompatible with the role assigned to it or because there is more than one pronoun in the clause}\)
confusion sets in, as in interpreting “they” in the following example:

**Example 9**

“One difficulty for many young people in New York City is that while part-time jobs tend to be concentrated in Manhattan, they tend to be concentrated in other boroughs, and that makes it difficult to work after school.” [New York Times, August 14, 1988]

The point is, in normal circumstances, the processing of anaphoric pronouns appears to involve a very local, highly constrained pair of data structures, $C_b$ and $C_f$. It is not unreasonable to assume that a number/gender compatible pronoun (or zero-anaphor, in languages like Japanese) is immediately interpreted as specifying the $C_b$. Any later-discovered incompatibility in that assignment switches the processor to members of $C_f$ and then to conscious reasoning or confusion.

This is in contrast with the processing of discourse deictics. First, as shown above, the space of likely referents for deictics is different and much larger than for anaphoric pronouns, being any sequent corresponding to a segment on the right frontier, plus any property of such a sequent or any entity within it.\(^{12}\)

Moreover, most instances of *this* and *that* in text occur as the first noun phrase in a clause.\(^{13}\) At this point, the only knowable constraint on the interpretation of one of these deictics is that it derives from a segment on the right frontier. If one assumes that the language processor does not wait for all constraints to come in before interpreting a discourse deictic, it must be able to develop and make use of a *partial interpretation* based on whatever constraints it has. I see this as analogous to spatio-temporal uses of *this* and *that*, where a person may first recognize a general pointing gesture, and then tries to figure out what is being pointed to specifically, based on what the speaker says about it and perhaps general heuristics about what might be worth pointing to.

This need for partial interpretations that can be constrained as more of the clause is processed further distinguishes the interpretation of deictics differs from that of anaphora. In Section 5 however, I will present evidence for a particular link between discourse deictics and anaphoric pronouns, which will give substance to Ehlich’s assertion that deictics “achieve focussing of the hearer’s attention”.

\(^{12}\)This provides a definition for Ehlich’s “deictic space”.

\(^{13}\)Of the 81 clausally-referring instances of *this* and *that* pronouns I tabulated, 60 (~74%) were in subject position in standard SVO clauses, 19 were post-verbal noun phrases, and 2 were preposed adverbials (e.g. “after that”). Hence ~77% were first noun phrases. This is not the case with anaphoric pronouns.
Notice that this step of further constraining a pointing gesture also allows for a uniform treatment of *this* and *do this* (*that* and *do that*). That is, given a preposed *this* or *that*, a listener may not be able to determine right off (or for some ways into the sentence) whether it serves as the object of *do* or it plays some other role in the clause: s/he cannot tell until s/he reaches the gap (or “trace”) co-indexed with it, as in Example 10.

**Example 10**

1. Gladys told Sam last night that Fred was a complete jerk.
2a. Anyway, *that*s what Fred believes that Gladys said T1.
2b. Anyway, *that*s what Fred believes that Gladys did T1.14

There are two ways out of this for listeners:

1. they can commit themselves to one interpretation (as action, situation, or entity) and retract it later if it turns out to be incompatible with the syntax, or
2. assuming actions are properties of their respective events, listeners can take the interpretation of a discourse segment on the right frontier and coerce that sequent into a component action if that is what is required by syntax.

The latter seems to me more likely.

### 4.3 Differences between *this* and *that*

Nothing in the interpretive process so far described makes a distinction between *this* and *that*. In the many cases where the two appear interchangeable (cf. Example 11), there appears to be no difference in what they can access.

**Example 11**

a. Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think. *This* suggests that

---

14In an e-mail message I received, I found a preposed *that* serving as both the object of *do* and the object of a regular verb: Several universities have made computer science a separate school. But *that* is not necessarily what we want or even what we could do.

This example can be taken as a “speech” error or as additional evidence that at some level, both forms are considered the same.
stereo fusion is not necessary for thought.

b. Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think. That suggests that stereo fusion is not necessary for thought.

In the cases where the two are not interchangeable (cf. Example 12), the reason for the oddness does not seem to follow from any difference in accessibility (i.e., an ability to identify the intended referent), but rather from some kind of clash between the “psychological distance” implied by the deictic (near/far) and the respondent’s felt “psychological distance” from its referent. (In Example 12, it seems reasonable for B to specify A’s statement as that – distant from B – but odd for B to specify it as this – close to B.)

Example 12

A: Hey, management has promoted Fred to second vice president.
B1: That’s a funny way to describe the situation.
B2: #This is a funny way to describe the situation.

Compare Example 12 to the following example where there is no “attitudinal clash”. Here, it seems reasonable for B to consider the event of Fred’s promotion either as something “psychologically” close to him or as something far away.

Example 13

A: Hey, management has promoted Fred to second vice president.
B1: Wow! That’s wonderful for us.
B2: Wow! This is wonderful for us.

Thus I would claim that this near/far deictic dimension (with spatial near/far extended to “attitudinal” near/far) only comes into play either in filtering possible referents or even later, in attributing speaker attitude towards the one chosen. (Robin Lakoff [Lakoff] discusses psychological distance in her not on this and that, but does not tie it in with processing.)

Here is some additional suggestive anecdotal evidence regarding “attitudinal aspects” that other researchers might like to consider:

• In technical reports and conference papers I have looked at, this predominates strongly over that. Here people are presenting their own research results and arguing for their own theories.

• In the few newspaper editorials and op-ed columns I have looked at, that predominates strongly over this. These editorials and columns have all
been critical of some government policy.

- In one of the technical reports, this appears 13 times in positive clauses and once in a negative clause ("If it doesn't know this, ... "). That, on the other hand, is used only twice, both times in negative clauses ("since that has not yet been done, ... ") and "Worse than that, ... ").

I will have nothing more to say about these observations in the current paper, as they are peripheral to my main interest in attention and processing.

Before I close this section though, let me show how the current account also explains the observation that listeners, when quizzed on what a particular this or that refers to, may vary widely in their response.

As noted, a sequent can correspond to the interpretation of more than one natural-language clause. However, those clauses themselves may not be retained in the listener's memory. In such a case, describing the referent of a particular discourse deictic would be a constructive act and hence subject to variability, including how much the listener chooses to describe explicitly.\(^\text{15}\) This variability goes beyond any ambiguity in how inclusive a segment provides the referent.

To summarize, in Section 2, I argued for the existence of sequents – mental correlates of discourse segments – as providing referents for deictic pronouns in text. In this section, I have argued that determining the referent of a discourse deictic makes use of two different sets of constraints – one attentional, the other semantic. While a listener is aware of the attentional constraints right away, more of the clause may have to be processed before semantic constraints become clear as well. The consequences for automatic text understanding systems are that:

1. such systems must build up an individual representation of each chunk of text taken to be a segment (i.e., a sequent is not just context for pronoun and noun phrase interpretation);

2. they must be able to construct and reason with partial interpretations and perform the type of reasoning that has been called coercion.

\(^{15}\)Schuster [Schus88] makes a similar point with respect to text describing a sequence of actions related by Goldman's [Gold75] generates relation: if a that follows, when asked about its referent, the listener may either describe it as the entire relation or just the final action.
5 Discourse Entities and Activation

So far my discussion of discourse deixis has primarily been in terms of only one of the three discourse notions mentioned in Section 2 - discourse segments. Now I want to present some data relevant to the other two - discourse entities and activation - and argue that it provides additional evidence for:

- another relationship between discourse segments and discourse entities;
- a rough partition of a listener’s attention into
  - those concepts s/he can be taken as actively attending to, that can be specified with an anaphoric pronoun. (These I will call centered, following [GJW83].)
  - those that s/he is also attending to, but only as background to the centered entities, that can be specified with a deictic pronoun or noun phrase. (These I will call background.)

There is some overlap between these two, resulting in speakers’ abilities to refer to some entities with either deictic pronouns or anaphoric pronouns.

- a sharp contrast between the above-mentioned background concepts, which can be specified deictically, and inferrable concepts, which can only be taken as associated with something that the listener is attending to and can only be specified with a definite noun phrase.

Evidence for the first two points comes from a common pattern of usage in which specification using this or that is followed by co-specification using it. This is illustrated in the following examples:

Example 1

It’s always been presumed that when the glaciers receded, the area got very hot. The Folsum men couldn’t adapt, and they died out. That’s what is supposed to have happened. It’s the textbook dogma. But it’s wrong. They were human and smart. They adapted their weapons and culture, and they survived.

Example 2

The tools come from the development of new types of computing devices. Just as we thought of intelligence in terms of servomechanism in the 1950s, and in terms of sequential computers in the sixties
and seventies, we are now beginning to think in terms of parallel computers, in which tens of thousands of processors work together. This is not a deep, philosophical shift, but it is of great practical importance, since it is now possible to study large emergent systems experimentally. [Hil88, p.176]

Example 15

I don’t think this can be taken seriously either. It would mean in effect that we had learned nothing at all from the evaluation, and anyway we can’t afford the resources it would entail.

Example 16

The Texas attorney general said that the McDonald’s announcement represented “a calculated effort to make the public think that they were doing this out of the goodness of their heart when, in fact, they were doing it because of pressure from our office”. [Philadelphia Inquirer, 13 June 1986]

One can account for this pattern in the following way:

(1) The sequents associated with discourse segments on the right frontier, while active (explicitly verbalized) play the part of background to the discourse entities that the listener is most attending to. It is only this small set of centered entities that are accessible anaphorically.

(2) However, just as this and that can be used in pointing to parts of the background in spatio-temporal deixis, they can be used to access the sequents associated with segments on the right frontier in their role as background to these centered entities.

(3) Once background material has been referenced with a deictic pronoun, it becomes a centered discourse entity, simply by virtue of its mention as a noun phrase in the clause.

(4) Once centered, the normal situation for centered entities holds, and the resulting entity is accessible to reference by an anaphoric pronoun.

Note that I do not mean to imply that one cannot refer deictically to the same thing more than once – that by being specified deictically and hence focussed, something is removed in some sense from the background. Multiple co-referring deictics are certainly possible, for example
Example 17

They wouldn't hear to my giving up my career in New York. That was where I belonged. That was where I had to be to do my work. [Peter Taylor, *A Summons to Memphis*, p.68]

Example 18

By this time of course I accepted Holly’s doctrine that our old people must be not merely forgiven all their injustices and unconscious cruelties in their roles as parents but that any selfishness on their parts had actually been required of them if they were to remain whole human beings and not become merely guardian robots of the young. This was something to be remembered, not forgotten. This was something to be accepted and even welcomed, not forgotten or forgiven. [Peter Taylor, *A Summons to Memphis*, p.217]

But this pattern is the exception rather than the rule – a literary device marking the clauses as strongly parallel – comments on the same thing. (In cases like Example 17, the two clauses could have been presented in either order, which does not appear to be the case in the more common pattern of reference described above.) All it implies is that even something in the active background is specified deictically, it is not removed from there. On the other hand, it does become a centered entity and, as such, the easiest, least marked way of specifying it is with an anaphoric pronoun.

My third point – that deictics provide evidence for distinguishing background concepts from inferrable ones – comes from looking at deictic noun phrases. I have shown that a deictic pronoun can only be interpreted as referring to a sequent associated with a discourse segment on the right frontier, or a property of that sequent, or a discourse entity explicitly part of it – that is, a concept taken to be attended to but nevertheless background. The same holds for deictic noun phrases. This distinguishes them from definite noun phrases, which can also be interpreted as referring to things that are merely inferrable from the current situation – what Chafe has called semi-active concepts – for example

Example 19

a. John and Mary decided to go on a picnic.

b. While they remembered most things, they forgot to put the picnic supplies in the cooler.

c. So when they got to the park, the beer was warm.
Here the beer is easily interpretable as referring to an entity corresponding to the beer that is part of John and Mary's picnic supplies, where the picnic supplies are part of the current situation. By contrast, a similar example with a deictic noun phrase sounds definitely odd:

Example 20

a. John and Mary decided to go on a picnic.
b. While they remembered most things, they forgot to put the picnic supplies in the cooler.
c. #So when they got to the park, that beer was warm.\(^{16}\)

Another example illustrates this contrast in a different way: given a context that admits the use of either a definite noun phrase or a deictic noun phrase, the two noun phrases will pick out different referents, even if their descriptive content is the same - for example,

Example 21

a. Some files are superfiles.
b. To screw up some one's directory, look at the files.
c. If one of them is a superfile, delete it.

Example 22

a. Some files are superfiles.
b. To screw up some one's directory, look at those files.
c. They will tell you which of his files is absolutely vital to him.

In Example 21, the definite noun phrase the files is interpreted as referring to an entity corresponding to the files in that person's directory – that is, to a inferrable concept associated with something currently being attended to. In Example 22 on the other hand, the deictic noun phrase those files is interpreted as referring to an entity corresponding to the files that are superfiles, i.e., to a discourse entity explicitly in the interpretation of a sequent associated with a segment on the right frontier – part of the background.

Thus there is a difference between concepts being attended to, even as background, and those that can only be inferred.\(^{17}\)

\(^{16}\)A similar point is made in [Lakoff], although she does not provide any explanation.

\(^{17}\)With respect to Chafe's proposal [Chafe87] discussed in Section 2, this argues against simply lumping together concepts that have been mentioned explicitly but have not been refreshed and ones that are only inferrable. While Chafe is correct in taking both to be accessible to unmodified definite noun phrases, deictic noun phrases are limited to the former.
6 Conclusion

In this paper, I have proposed and argued for a process-based account of deictic pronouns in text. Its main contributions lie in having (1) enhanced the stature of discourse segments and their interpretations (sequents); (2) further specified properties of attention in discourse; (3) uncovered additional relationships between the notions of discourse entity, discourse segment, and attention; and (4) provided some meat to the intuition that deictics are different from anaphors.

7 References


Grosz81 Grosz, B. The Representation and Use of Focus in a System for Under­


Isard75 Isard, S. Changing the Context. In E. Keenan (ed.), Formal Seman­


Lakoff Lakoff, R. Remarks on This and That. CLS 10, 1974, pp.345-356.


MS88 Moens, M. and Steedman, M. Temporal Ontology and Temporal Refer­


Schiff84a Schiffman, R. Categories of Discourse Deixis. 29th Meeting of the International Linguistics Associations, New York University, 1984.

Schiff84b Schiffman, R. The Two Nominal Anaphors *it* and *that*. Papers from the Chicago Linguistics Society, 20th Regional Meeting, July 1984, pp.344-357.


