Neil Kumar

Human Language Technology

Professor Church

12/15/17

The Purpose of Prose: An Analysis of "Attention, Intentions, and the Structure of Discourse"

Introduction and History

Introduction

When studying human language technology, we looked at the field from a more quantitative point of view where we saw various algorithms, graph structures, and theories describe many phenomena of written and spoken text. However, we looked at the field from a somewhat insular perspective, where we primarily focused on research that aimed to predict, be it through regressions, n-gram models, and predictive models, something about text, typically parts of speech or sentence structure. From the great debate of Minsky versus Chomsky on syntactic structures to Shannon's analysis of entropy in the English language, most of the work we looked at tended to ignore the actual meaning of the words themselves in favor of syntactic and part-of-speech analyses. That is not to say that these studies are not important, since we saw that this research was very fundamental in establishing computational linguistics to where it is today, but a lot of this research failed to touch upon an often-overlooked data point, the meaning of the words themselves. Therefore, the goal of this paper is to analyze and emphasize the importance of a particular paper, "Attention, Intentions, and the Structure of Discourse" by Barbara Grosz and Candace Sidner. Grosz and Sidner, in their 1986 paper, introduced a novel process to efficiently and procedurally understand textual discourse, or conversation, by breaking it into its fundamental parts, namely its linguistic, intentional, and attentional structures. By building upon

research primarily from the 1970's and 1980's, Grosz and Sidner provided a framework from which other computational linguists could study language from a more humanistic point of view, which proved to be very useful.

History

Prior to Grosz and Sidner, various researchers in this more humanistic approach to textual analysis provided the groundwork from which the two could build their language model. Often cited as the father of this field, H.P. Grice published works that were fundamental in establishing language models that could encapsulate the meaning behind textual discourse. In his 1969 paper "Utterer's Meaning and Intentions", Grice breaks text, specifically textual conversation, into "utterances", where "utterers", or speakers, provide utterances that have different types, statements, questions, and commands for example. When looking at conversations, different speakers switch between different utterances and different utterance types, where each singular utterance is "incomplete" and only captures its "true" meaning in the context of the conversation. A phrase's "timeless meaning" refers to an utterance's most probable meaning in an incomplete context, while its "applied timeless meaning" refers to the utterance's meaning within the context of the whole conversation and the speaker themselves. Finally, Grice calls the "occasion meaning" of an utterance the meaning of an utterance within the context of the conversation, speaker, and any outside knowledge. For example, if utterer U says "I shall be helping the grass to grow", in an incomplete context this utterance could resemble the idiom "pushing up the daisies" if it exists in the grammar, but in most conversations its occasion meaning should be that the person wants to grow the plants (Grice 147-148). This way of "tagging" text culminates in Grice's own algorithms, where the parameters $A, f, r,$ and c refer to the audience, features of the utterance, responses, and modes of correlation, respectively (Grice 163). Grice attempts to

predict the probability of statements such as "*U* uttered *x* intending *A* to think *x* possesses *f*" given immense tagged data that populates the parameters (Grice 164). Therefore, in this paper Grice provided a framework with which to understand the intricate contextual meaning of phrases, or their occasion meanings.

In 1975, his paper "Logic and Conversation" built upon his previous work by refining what he meant by "conversation" and dictating elements he deemed necessary for "successful" conversation. Grice took a more "natural" approach to language, in his words, compared to the more "formal" symbolic analysis we saw in class which looked at language more in terms of parts-of-speech rather than meaning (Grice 42). When looking at data to analyze, he said that successful textual conversation abided by the "Cooperative Principle", where successive exchanges build upon each other and contribute to a common purpose (Grice 44). Additionally, he believed that all utterances spoken should abide by four general maxims, "Quantity", "Quality", "Relation", and "Manner" (Grice 45). Statements should have enough informational "quantity", contributing enough information as required and no more, display "quality", only saying what you know is true based on evidence, be relevant, relating to the current topic, and be presented in an unambiguous manner, avoiding obscurity. Although this paper was far more qualitative than his previous, linguists heavily adopted this perception of conversation because it presented a way to easily understand discourse in its inherently humanistic form by pinpointing its most relevant features. In conjunction with his more algorithmic view of meaning, Grice provided a template with which to understand conversation procedurally, solidifying his contribution to the field.

Building off of Grice, researchers in the early 1980's still took this humanistic approach to discourse but attempted to quantify many of the more qualitative facets of conversation. In "A Theory and Grammar of Spontaneous Discourse" from 1981, Rachel Reichman created the "Context Space Theory" to model conversations, where conversations are a sequence of "conversational moves", or distinct utterances that serve an overarching communicative goal (Reichman 13). Like Grice's Cooperative Principle, Reichman's Context Space Theory looks at discourse in terms of connected phrases, but Reichman looks at these phrases much more systematically and quantitatively. She suggests that each phrase contributes to the conversation's "Systemic Grammar", where a phrase is given a role in the conversation's meaning and determined by the words that make it up (Reichman 20). For example, if the current phrase is "Let's go.", based on preconditions, or training data, the systemic grammar would tag this as a command. Phrases are then connected by "clue words" like "except", "but", and "anyway" to build "Augmented Transition Networks (ATNs)", which are undirected graphs where the states are conversational roles and the arcs are the grammatical transitions connecting them (Reichman 24).

Figure 1

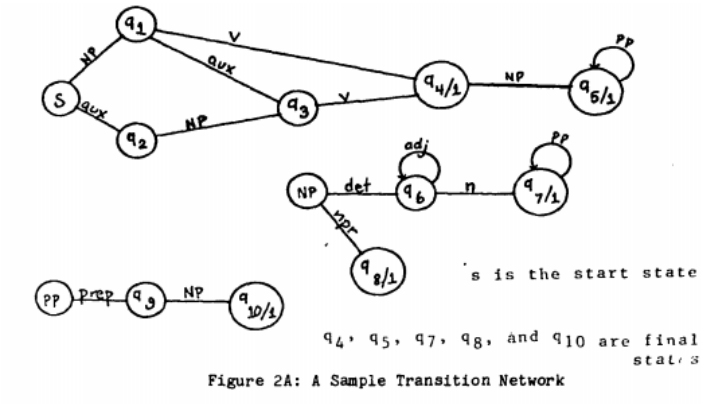


Figure 2A: A Sample Transition Network

Incorporating Chomsky's theories on syntactic structures, Reichman was able to build a rule-based theory of discourse that calculated the likelihood of both generated and sampled conversations while incorporating more detailed views of textual meaning. 1983's paper "Relational Propositions in Discourse" by William Mann and Sandra Thompson adapted Reichman's model on discourse but focused solely on the implicit relationships between phrases, or their "relational propositions" (Mann 2). Reichman already understood the importance of "cue words" like "but" that changed the relationship between sequential phrases, but she did not account for implicit meanings that described deeper relationships. For example, if "I'm hungry." was followed by "Let's go to a restaurant", Mann and Thompson's model would suggest that a "so" could be placed in between and correctly identify the causal relationship. Able to identify other relationships like reasoning, justification, and evidence between phrases, Mann and Thompson's work served as a large basis for Grosz and Sidner to identify intention in discourse. Finally, an important precursor to Grosz and Sidner's paper was Philip Cohen and Hector Levesque's "Speech Acts and Rationality" which looked at conversations as a combination of "actions" which are taken with respect to the different speaker's beliefs, goals, and more. Given conversational data, phrases are tagged as primitives like "BELIEF" and "GOAL" and cue words like "later" are tagged as "AFTER", producing logical statements like "*IF* BELIEF *TRUE THEN* ACT *THEN* GOAL" (Cohen 50-51). Compared to other research, Cohen and Levesque's emphasis on conversational logic aided Grosz and Sidner heavily with regards to their analyses of intentional structure and attentional states.

Ultimately, the goal of this paper is to emphasize the importance of Grosz and Sidner's study with regards to computational linguistics in this more humanistic form. In detailing the major

theories proposed by these two, most of which is rooted in prior research like the examples above, I hope to shed light on how this paper influenced other researchers, papers, and related fields. Furthermore, I hope to state how this paper can influence future research and be incorporated in adjunct areas like speech cognition.

"Attention, Intentions, and the Structure of Discourse": An Overview

To fully understand this paper's importance, you should know its inner workings. In my analysis of it, I will be looking at its three most core elements, the components that discourse can be broken down into, the issues that arise when processing discourse, and the applications of this model on real data.

Discourse Structural Analysis

Grosz and Sidner developed a unique model of discourse that aims to answer two questions, namely (1) Given a conversation, how many interleaved "discourses" exist? and (2) What makes discourse "coherent"? Before I clarify the above questions, a clear understanding of how they would solve this problem is needed. They formalized that discourse is made of three components, "the structure of the actual sequences of utterances in the discourse", or the linguistic structure, a "structure of intentions", helping to define discourse structure and coherence, and an "attentional state", necessary when processing sequential discourse (Grosz 175). The above definitions will become much clearer as I delve into specifics ahead.

The two define "discourse" to be "a piece of language behavior that typically involves multiple utterances and multiple participants", essentially independent statements by a number of

participants (Grosz 176). When differentiating speakers, the two call the first speaker the "initiating conversational participant (ICP)" and the second speaker the "other conversational participant (OCP)" (Grosz 176). Given text that provides sufficient phrases for utterances and speakers for participants, the different components of discourse described above can be extracted.

1) Linguistic Structure

The first component of discourse is its "linguistic structure", which refers to the structure of the individual utterances, aggregating into "discourse segments" (Grosz 177). Borrowing some of Mann's research on segmented dialogues, the two say that this linguistic structure consists of the discourse segments and an embedding relationship between them. Additionally, there is a significant relationship between individual utterances and the discourse segments they create. Because utterances (individual phrases) can be used to convey information about discourse structure (changes in the conversation), then the context of the overall conversation (discourse structure) influences the interpretation of the individual express as well (utterances). Given these definitions, certain words can be pinpointed which change the discourse structure, known as "clue words" (Grosz 178). Phrases like "but then" not only provide clues about the discourse structure, but they also can indicate changes in the intentional or attentional states of the discourse.

2) Intentional Structure

Second, Grosz and Sidner discuss the "intentional structure" as each discourse segment's contribution to an overall purpose. This overall purpose is called the "discourse purpose (DP)",

or “the intention that underlies engaging in the particular discourse” (Grosz 178). Since discourse is broken into discourse segments, each segment naturally has its own intention, called the “discourse segment purpose (DSP)” which specifies how this specific segment contributes to the overall purpose. Two structural relationships provide great insight into intention, including dominance and satisfaction-precedence. Grosz and Sidner say an intention DSP1, “dominates” another, DSP2, if “an action that satisfies one intention (DSP2) is enacted to provide the satisfaction of another (DSP1)” (Grosz 179). Therefore, the two show that the order in which DSPs are satisfied is important because it contributes to a dominance hierarchy, eventually influencing the overall purpose. Similarly, they say “DSP1 satisfaction-precedes DSP2 if DSP1 must be satisfied before DSP2”, also providing clues towards the DP (Grosz 179). As overarching maxims, Grosz and Sidner aim to “determine” the meaning of these DSPs, namely knowing what is intended by whom, and the “recognition” of them, or the process that leads a participant to identify what the intention is.

3) Attentional State

Finally, the third pertinent component to discourse is its “attentional state”, which is “a property of the discourse itself…recording the objects, properties, and relations that are salient at each point in the discourse” (Grosz 179). More clearly, each speaker has their own understanding of the conversation where different attentional states are called “focus spaces” and contribute to an overall “focus structure” that connects states through transitions. Since DSPs provide background to the purpose of the conversation, focus spaces inherently rely on the intentional structure of the discourse as well.

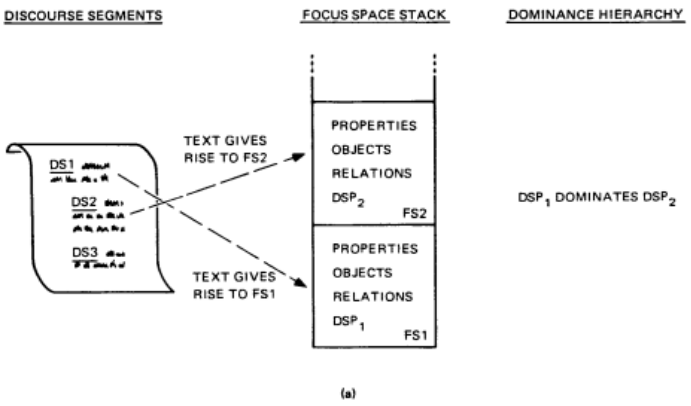
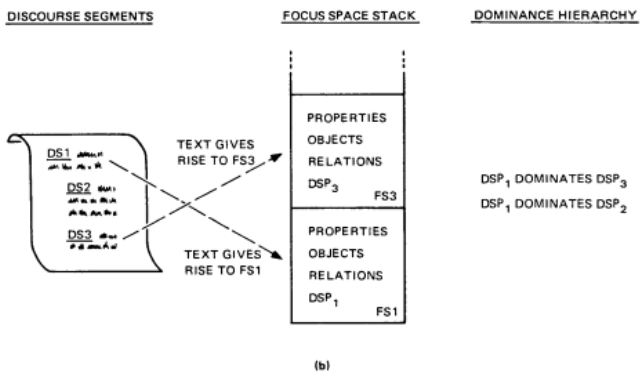


Figure 2



To provide better understanding of this focusing structure, Figure 2 above demonstrates how discourse segments (titled "DS1" etc. on the left) are mapped to focus spaces, where "higher" focus spaces are informed by lower ones on the stack. Part (a) indicates that because DSP1 dominates DSP2, then FS2 (Focus Space Stack 2) is informed by FS1, so any information like a "red ball" in DS1 will be interpreted as the same "red ball" in DS2 if it is mentioned again. Unlike (a), (b) shows that when a third DSP is being interpreted (DS3) FS3 pushes FS2 out of the top based on the dominance hierarchy given, mimicking the DSP relations already in place. Therefore, Grosz and Sidner emphasize two important properties of focusing structures: first, the focusing structure relies heavily upon the intentional structure, where cue words like "first" in the discourse determine relative positions in the focus space. Second, the focusing structure changes as the conversation continues. In comparison to the intentional structure, the focusing

structure can only look at a certain subset of the discourse at a time (the stack), while the discourse-level intentions build and inform each other over time. Additionally, when the discourse is being processed only the attentional state can interpret utterances directly; the intentional structure must wait to finish processing. Ultimately, the attentional state acts as a "repository for the contextual information" needed to process utterances at each point in the conversation (Grosz 182).

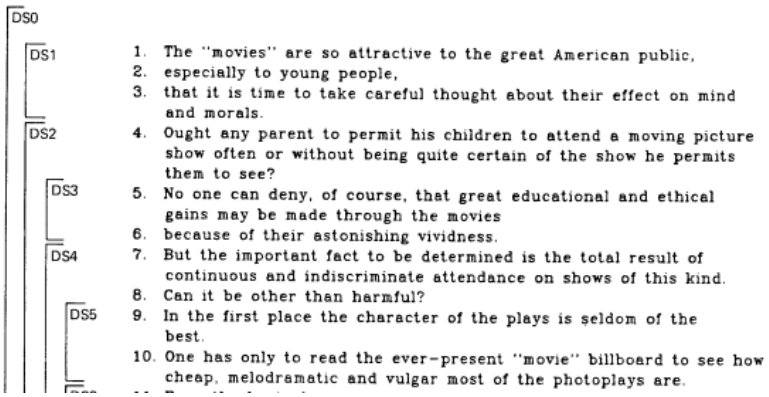
Discourse Processing Examples and Issues

To better convey the meaning of Grosz and Sidner's model, I will be analyzing two examples put forth by them.

1) Rhetoric Example

Example 1, given in Figure 3, comes from Cohen and analyzes DSPs marked by Cohen herself.

Figure 3



I0: (Intend ICP (Believe OCP PO))

where PO = the proposition that parents and teachers should guard the young from overindulgence in the movies.

I1: (Intend ICP (Believe OCP P1))

where P1 = the proposition that it is time to consider the effect of movies on mind and morals.

I2: (Intend ICP (Believe OCP P2))

where P2 = the proposition that young people cannot drink in through their eyes a continuous spectacle of intense and strained activity without harmful effects.

I3: (Intend ICP (Believe OCP P3))

where P3 = the proposition that it is undeniable that great educational and ethical gains may be made through the movies.

I4: (Intend ICP (Believe OCP P4))

where P4 = the proposition that although there are gains, the total result of continuous and indiscriminate attendance at movies is harmful.

I5: (Intend ICP (Believe OCP P5))

where P5 = the proposition that the content of movies (i.e., the character of the plays) is not the best.

Figure 4

I0 DOM I1
I0 DOM I2
I2 DOM I3
I2 DOM I4
I4 DOM I5

Figure 5

Although Grosz and Sidner use the same tagged DS's as Cohen, Figures 4 and 5 indicate their unique analysis methods, where Figure 4 gives the primary DSP of each component and Figure 5 is their dominance relationships. P1…P5 in Figure 4 are extracted to be five intentions that the participant (OCP) are supposed to believe, but the DSPs are not always explicitly expressed through the utterances that make them up. For example, DS1's purpose is "explicit" because the speaker "intends" for the OCP to believe a certain proposition, while DS2 is "derivative" because it relies upon DS5 and more utterances (Grosz 185). Using these distinct propositions, the interactions between expressions, and cue phrases that indicate discourse boundaries (eg. "But" starting DS4), Grosz and Sidner are able to mostly understand discourse by noticing attentional changes, segmenting ideally different phrases, and tying the pieces together to create an overall purpose.

2) Task-Oriented Example

Example 2, shown in Figure 6, is a more task-oriented conversation pulled from Grosz's earlier research.

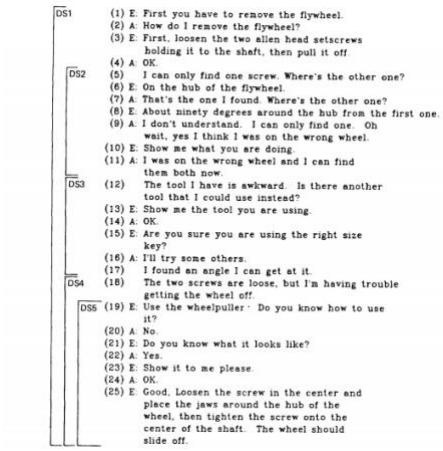


Figure 6

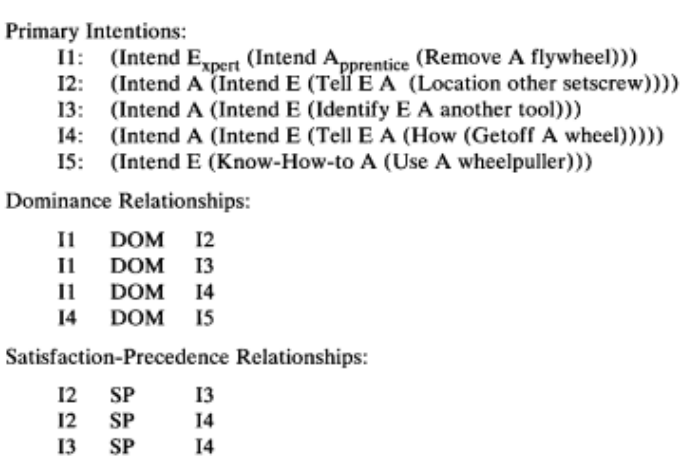


Figure 7

Unlike the previous example, this conversation revolves around the speaker (ICP) telling the listener (OCP) to perform some actions. Since this conversation involves different speakers initiating different DS's, there must be some sort of "cooperative behavior" between the two, sharing knowledge of the task when necessary (Grosz 185). Since the OCP has to carry out tasks from the ICP, the OCP will have to perform subtasks that work towards larger tasks, creating the dominance relationships illustrated in Figure 7. For example, DS1, DS2, and DS3 are all related, where DS1 indicates that to remove a flywheel, DS2, finding the screws in the wheel, must first happen then the OCP must find the proper tool to remove them (DS3). These satisfaction-precedence and dominance relationships create trees of intentions, where the most probable tree becomes the intentional structure and informs the attentional state.

3) Issues with Processing Discourse

Grosz and Sidner identify three critical issues when attempting to process discourse, including how "the ICP indicates and the OCP recognizes" a new discourse segment, how the "OCP recognizes the DS purpose", and "how the focus space stack operates" (Grosz 188). First, although cue phrases play an integral part in recognizing new DS's, they often fail in producing

clear domination or satisfaction-precedents and can often confuse the attentional state. For example, if the ICP says "I want you to arrange a trip for me to Palo Alto. It only be 1 week." followed by "But anyway, my day was great." then the attentional state might not know the previous context for this new statement and be confused about which DS it fits into. Second, for a conversation to be "coherent and comprehensible, the OCP must be able to recognize the DP/DSPs", but this becomes hard as the OCP must able to see which DSPs dominate others and understand their correct relationships (Grosz 188). Because OCPs cannot fully recognize the meaning of the conversation until the end, they have to have some understanding of the DSP's during the discourse and hopefully gain information from the ICP through indicators, intentions, and propositions. However, this is often hard for the OCP to achieve, so fully comprehending conversations has its issues. Finally, the focus space stack in often unwieldy because two DS's might be close to each other in the attentional structure and will often use locality in determining dominance hierarchy, sometimes creating incorrect relationships and messing up the attentional state. Although these issues have held Grosz and Sidner back from analyzing volumes of varying data, they have been able to identify many useful applications of this model.

Applications of this Computationally Logistic Model

1) Interruptions

Grosz and Sidner identify their model's ability to detect interruptions in a conversation as an important application of their theory. They call interruptions "strong" when a current DSP is relationally independent of all prior DSPs, creating "true interruptions" and digressions (Grosz 192). On the other hand, an interruption is "weak" when a current DSP is only independent of the DSP before it, creating "flashbacks" (Grosz 192).

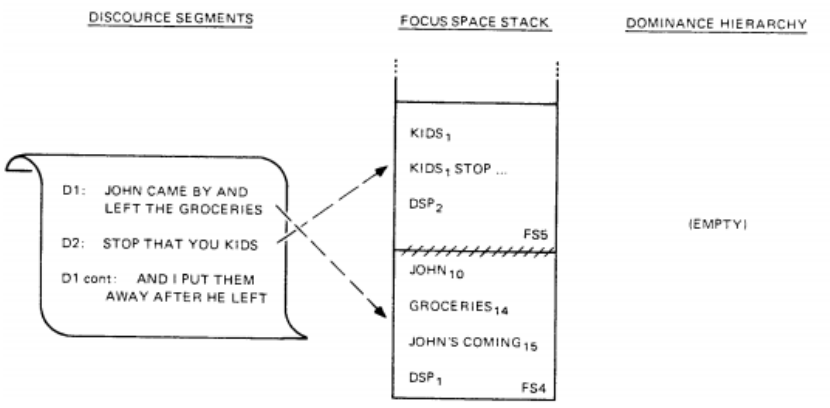


Figure 8

As shown in Figure 8, because D2 occurs inside D1 one would think that the two DS's are related; however, because D2 is a "strong" interruption the intentional states for DSP1 and DSP2 are distinct so FS1 and FS2 are independent, indicated by the slashed boundary. Another strong interruption is a digression, which is similar in that the current DSP is independent of all others but provides the intentional basis of a new DS. For example, if a conversation is occurring and a speaker interjects with "Speaking of Bill, that reminds me…", then this phrase acts as a digression to a new focus space but provides stack and intention shifts by mentioning Bill. Finally, interruptions can be "weak" as in the example of flashbacks, which are independent of the previous DSP yet is dominated by some previous DSP. For example, if you are talking about Bill then a speaker interjects with "Oh, I love Google." then you reply "How did Bill get that job at Google?", then the OCP's interjection was a flashback because it satisfied a previous DSP (regarding Google) then returned to the original DSP (Bill). The ability to identify different kinds of interruptions proves very useful to understanding a conversation's meaning because they provide great insight into discourse structure and affect processing.

2) Cue Words

Grosz and Sidner also cite their ability to find "cue words", or phrases that are subtler in indicating changes in meaning, as an important application of their work. When cue words are introduced, the focus stack model changes as a "new dominance" relationship emerges, where the focus space on the top of the stack changes to that of the current DS (Grosz 196). However, because this cue word ties the new DS to a previous one, we gain understanding of the dominance hierarchy within the conversation. Additionally, cue words can be used to see if DSPs are connected and can also be used to see if the attentional state has changed.

Attentional Change	
(push)	now, next, that reminds me, and, but
(pop to)	anyway, but anyway, in any case, now back to
(complete)	the end, ok, fine, (paragraph break)
True interruption	
	I must interrupt, excuse me
Flashbacks	
	Oops, I forgot.
Digressions	
	By the way, incidentally, speaking of, Did you hear about…, That reminds me
Satisfaction-precedes	
	in the first place, first, second, finally, moreover, furthermore
New dominance	
	for example, to wit, first, second, and, moreover, furthermore, therefore, finally

Figure 9

Exemplified in Figure 9, different cue words are often highly correlated with specific DSP relationships. When looking at conversation in terms of its intentional structure and attentional state, cue words provide great information into the discourse structure.

Impact on Subsequent Research and Literature

Although Grosz and Sidner's paper discusses discourse meaning, it neither aims to nor succeeds in providing a model to understand conversation, rather, its strengths are in its approach to discourse structure and providing a framework with which to understand discourse. Since the two looked at discourse as a combination of psychological, textual, and cognitive elements amongst others, it makes sense that this paper has been influential in a number of fields. Therefore, by providing a clear, detailed approach to understanding discourse, Grosz and Sidner were able to influence three major research fields, textual comprehension, cognitive understanding and analysis of discourse, and written and spoken understandings of communication.

Impact on Textual Comprehension

Perhaps the most straightforward interpolation of Grosz and Sidner's paper has been in the area of textual comprehension, where various researchers have utilized Grosz and Sidner's research to expand upon understanding different facets of text. One of these more notable papers is "Constructing Inferences During Narrative Text Comprehension", which aimed to describe how readers construct meaning for textual passages. According to the authors, a "knowledge-based inference" is created when the reader's long-term memory is activated to "encode meaning representation of the text" (Graesser 374). When doing this study, the authors assumed levels of representation in the text that mirrored Grosz and Sidner's model heavily, including the "surface level" (the exact words), the "textbase" (propositions and inferences), and the "situational model" (different states) (Graesser 376). In conjunction with a "discourse focus", where attention changes often, and a "local and global" coherence of the text, where the textual meaning differs

based on scope, this paper utilizes Grosz and Sidner's model heavily in a novel context. Jane Morris and Graeme Hirst, in their 1991 paper "Lexical Cohesion", applied Grosz and Sidner's model much more directly to study a lexical phenomenon they had not touched upon. Namely, Morris and Hirst wanted to find related words that contributed to the same lexical meaning, calling this "lexical cohesion". Lexically cohesive words can form "lexical chains" where related words span a topical unit of text. Using the discourse structure proposed by Grosz and Sidner, Morris and Hirst studied lexical chains to "provide context to aid in the resolution of ambiguity" and "provide a clue for the determination of coherence and discourse structure" (Morris 23).

Chain	Chain Range	Intention	Intention Range
1	1–44	1	1–44
2.1	2–12	1.1.1, 1.1.2	1–12
2.2	16	end of 1.1.3.1	16
2.3	24	end of 1.1.3.3	25
3	13–15	1.1.3.1	13–16
4	19–20	1.1.3.2	17–22
5	31–33	1.2.2	31–33
6	34–38	1.2.3	34–39
7,8	1–3	1.1.1	1–7
9	7–8	1.1.2	8–12

Figure 10

Beyond this, as illustrated in Figure 10 Morris and Hirst demonstrated that the intentional structure of discourse relates very closely to its lexical chains, where separate chains often indicate separate intentional units. Therefore, Morris and Hirst's paper built upon Grosz and Sidner's by looking at an unobserved lexical phenomenon and providing more information to the discourse's structure. Finally, in "Bursty and Hierarchical Structure in Streams", Jon Kleinberg adapted parts of Grosz and Sidner's paper to extract meaningful structure from "bursts" of textual data. When receiving masses of e-mail data, Kleinberg developed hierarchical structures out of the e-mails and developed a nested structure of bursts that "bears an intriguing relationship" to Grosz and Sidner's structural model (Kleinberg 19). As a precursor to modern

web-scraping, Kleinberg's model interpreted Grosz and Sidner's in a fairly new context. From these influential papers and many other related to textual comprehension, Grosz and Sidner provided a great platform for others to understand textual discourse in new ways.

Impact on Cognitive Understanding and Analysis

More unrelated yet still relevant to Grosz and Sidner's paper is its application to the relationship between cognition and discourse. In "Long-Term Working Memory", K Ericsson and Walter Kintsch adapt many principles of textual comprehension to develop better memory retention in the short and long-terms. Ericsson and Kintsch believe that our mental representation of text consists of the linguistic structure, propositional textbase, and situational model, a model reminiscent of Grosz and Sidner's. In addition, they believe retrieval cues, encoded associations, patterns, and more are inherently part of our mental models of text.

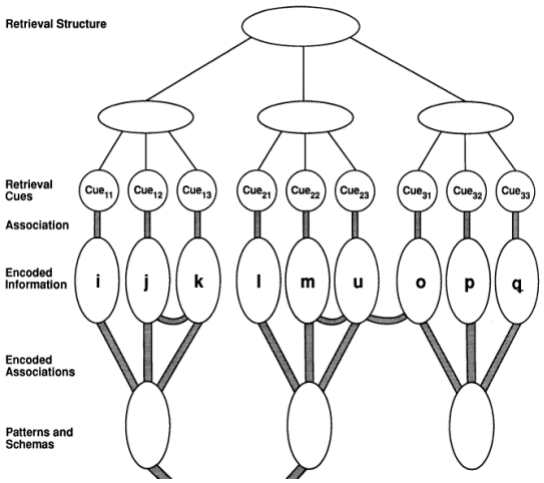


Figure 11

Figure 11 above is Ericsson and Kintsch's model of our long-term memory (LTM), where different associations like the ones seen in Grosz and Sinder's paper are represented. However, Ericsson and Kintsch created their own model, the "CI-model", which uses unique operations

like question answering and priming effects to better simulate human comprehension (Ericsson 29).

Additionally, in "Cognitive Status and the Form of Referring Expressions in Discourse", the authors find a strong correlation between the "cognitive status of the referent", or the role of the object in the statement, and the expression that should identify it, often a pronoun (Gundel 276). For example, if a speaker says, "I couldn't sleep last night. A dog kept me awake", then the referential "dog" seems to have kept the person awake, so "this" could replace "A" and provide more context into why an action happened. Therefore, this paper ties the cognitive role of an actor to an expression that could learn from it, providing a greater cognitive understanding of language and vice versa. Although Grosz and Sidner's paper deals little with the cognitive role of its model, it has seemed to be very beneficial in helping others recognize language better.

Impact on Understanding Spoken and Written Communication

Finally, in the field of written and spoken communication, Grosz and Snider's paper has been adapted to pinpoint failings in understanding components of communication and studying elements of language that are often ignored like intonation. In "Grounding in Communication", Herbert Clark and Susan Brennan suggested that "grounding" is a process where every participant in a conversation must mutually believe a certain criterion to have enough information for the following conversation (Clark 223).

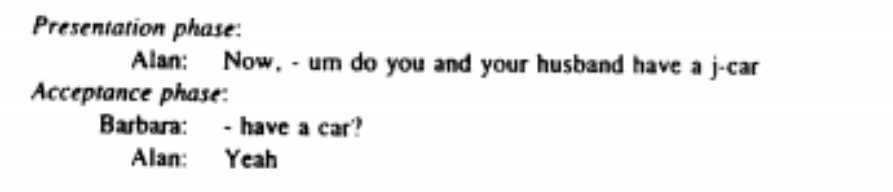

Presentation phase:
Alan: Now, - um do you and your husband have a j-car
Acceptance phase:
Barbara: - have a car?
Alan: Yeah

Figure 12

For example, Figure 12 above indicates that grounding occurs in two phases where information is first presented, then grounding occurs if it is accepted. Therefore, once everyone reaches a "common ground", everyone is acting with the same information unless "negative evidence" is presented which undermines any previous facts. This theory of a common ground is not found in Grosz and Sidner's paper, but because Clark and Brennan studied this theory in the context of linguistic structure it could be used to inform the intentional structure and attentional state of discourse.

Researchers like Ladd and Cutler, however, used this understanding of discourse structure to show that the study of "communication" lacked vital human components like intonation or rhythm. In "Intonational Phonology", Robert Ladd noted that analysis of discourse for meaning failed to include pitch or intonation. For example, he presented the sentence "Dogs must be carried." as an ambiguous phrase where intonation is needed to fully understand the sentence's meaning. In a similar vein, other researchers stated that the lack of "prosody", or rhyme and tempo, analysis makes no sense when "human sentence processing may be argued to be prosodically based terms" (Cutler 146). In this paper, the authors also cite prosody as useful when understanding syntactic and discourse structure, where changes in rhythm or tempo might signal clear breaks in thought. They even cite Grosz as a proponent of their work since Grosz ran tests and found "statistically significant associations of aspects of pitch range, amplitude, and

timing with features of global and local structure", allowing them to have better understandings of syntactic attachment and less ambiguity (Cutler 181). In looking at discourse from a different point of view, researchers like Clark, Ladd, and Cutler opened up new discussions as to what communication entails and how to more effectively study it.

Conclusion

In analyzing "Attention, Intentions, and the Structure of Discourse" by Grosz and Sidner, I hoped to shed light on this more humanistic form of computational linguistics that greatly contrasted what we saw in class. Jumpstarted by Grice in the 1960's, this look at language has only evolved as more and more researchers have looked to redefine what it means to "process" language and gain meaning from it. Although Grosz and Sidner did not attempt to fully understand the meaning of conversation, they did provide a template with which to break it down into its most logically sound parts, aiding countless other researchers and influencing other fields to look at conversation in novel ways. In our current technological climate, where AI and Machine Learning are moving at rapid speed, it seems only right that we look to more humanistic approaches to language processing. With voice assistants like Siri and Alexa that already understand human speech fairly well, I hope to see researchers take a page from Grosz and Sidner's book and work towards making these artificial systems mimic the human mind as closely as possible.

Bibliography

Clark, Herbert and Brennan, Susan. “Grounding in Communication”. American Psychological Association. Perspectives on Socially Shared Cognition. 1991.

Cohen, Philip and Levesque, Hector. “Speech Acts and Rationality”. Association for Computational Linguistics. 1985.

Cutler, Anne, Dahan, Delphine, and Donselaar, Wilma. “Prosody in the Comprehension of Spoken Language: A Literature Review”. Sage Journals. Volume 40. Issue 2. April 1, 1997.

Ericsson, Anders and Kintsch, Walter. “Long-Term Working Memory”. Psychological Review. 1995.

Graesser, AC, Singer, M, and Trabasso, T. “Constructing Inferences During Narrative Text Comprehension”. Psychological Review. 1994.

Grice, H.P. “Logic and Conversation”. Syntax and Semantics 3: Speech Arts. 1975.

Grice, H.P. “Utterer’s Meaning and Intention”. The Philosophical Review. April 1969.

Grosz, Barbara and Sidner, Candace. “Attention, Intentions, and the Structure of Discourse”. Journal of Computational Linguistics. July 1986.

Gundel, Jeanette, Hedberg, Nancy, and Zacharski, Ron. “Cognitive Status and the Form of Referring Expressions in Discourse”. Linguistic Society of America. June 1993.

Kleinberg, Jon. “Bursty and Hierarchical Structures in Streams”. Data Mining and Knowledge Discovery. October 2003.

Ladd, Robert. “Intonational Phonology”. Cambridge University Press. 1996.

Mann, WC and Thompson, Sandra. “Relational Propositions in Discourse”. Discourse Processes. November 11, 2009.

Morris, Jane and Hirst, Graeme. “Lexical cohesion computed by thesaural relations as an indicator of the structure of text”. Journal of Computational Logistics. March 1991.

Reichman, Rachel. “Plain speaking : a theory and grammar of spontaneous discourse”. Cambridge, Massachusetts. 1981.