Coherence, Reference, and the Theory of Grammar

Andrew Kehler

Contents

1

Introduction

In *The Psychology of Science*, the eminent psychologist Abraham Maslow remarked, as a comment on the mechanistic tradition of research in behaviorism, "I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail" (Maslow 1966, pp. 15-16). This maxim (and many variations of it) has since become widely cited, for reasons that are not difficult to see: It embodies a universal tendency that professionals working in various fields – including research, engineering, business, law, teaching, and yes, carpentry – find themselves succumbing to in the normal course of pursuing their trade. After all, when you lack access to the proper tools for a task, you are basically left with three choices. First, you can go out and acquire a suitable tool. This is usually the most difficult of the three choices, since you may have to search far and wide, or even invent the necessary tool yourself. Second, you can simply ignore the problem or task at hand and pursue something else. Lastly, you can try to use a tool that you *do* have. This is often the convenient middle road chosen, even if the fit to the task, and the result of applying the tool to it, leaves something to be desired.

I would perhaps not be going too far out on a limb if I suggested that linguistic theory is not immune to Maslow's maxim. We are all familiar with studies in which a researcher has seemingly addressed an interesting and recalcitrant linguistic fact by attempting to retrofit it to an existing set of tools, with a result that is not unlike how a pistachio might look were one to try to split it with a hammer instead of a nutcracker. This type of approach not only typically fails to shed much light on the phenomenon at hand, it often leads to the contribution of unwarranted assumptions about the object of study itself: If the tool does not fit, and one continues to maintain the assumption that it should, then the only conclusion one can draw is that the object of study must not be as it appears. It is perhaps natural to view a potentially deep and complex

phenomenon through the lens provided by a familiar (even if comparatively narrow) set of linguistic tools, but it can have dire ramifications for the state of our understanding. Indeed, a continual cycle of analyses based on a fundamentally flawed premise can keep entire bodies of literature in a state of stagnancy. The stagnancy results not from a lack of cleverness on the part of the linguists who wield their tools, but because the tools are inherently not, in and of themselves, capable of explaining the data.

My goal in this book is to introduce – or rather, reintroduce – a set of tools which I claim has a relevance to linguistic theory that has largely gone unacknowledged. The tools pertain to the methods by which hearers establish the *coherence* of multi-clause sentences and discourses. Discourse coherence is an area that to this point has typically been studied in relative isolation, as a postcursor to the production and processing of the syntactic and semantic structures of individual sentences. I intend to show, however, by way of developing a theory of coherence and then using it as a crucial component in analyses of five diverse linguistic phenomena, that coherence is not only a useful tool for analyzing mainstream linguistic problems, but also a necessary one.

What is Coherence?

When we comprehend a discourse, we do not merely interpret each utterance within it. We also attempt to recover ways in which these utterances are related to one another. To see this, consider the rather unremarkable passage given in (1).

(1) John took a train from Paris to Istanbul. He has family there.

In most discourse situations, we will likely infer that John's having family in Istanbul is the reason for his taking a train there. While this inference is not explicitly stated, it is a natural one to draw under the assumption that the utterances bear some relationship to each other, that is, that the discourse is coherent.

We can compare passage (1) in this respect with passage (2), from Hobbs (1979).

(2) ? John took a train from Paris to Istanbul. He likes spinach.

Most people find this version to be notably odd, yet, like (1), the sentences that comprise it are both well formed and readily interpretable. Interpretation instead goes wrong during one's attempt to infer a connection between them. After all, what does going to Istanbul have to do with liking spinach? In asking this, we are questioning the coherence of the passage.

As Hobbs (1979) points out, with a little thought one might come up with a scenario in which passage (2) would become coherent. For instance, one could conjecture that perhaps the spinach crop failed in France, and Turkey is the closest country in which spinach is available. Under this assumption, one can now infer a cause-effect relationship analogous to the one we identified for passage (1), and as a result the passage is more natural.

The fact that hearers infer such relations when interpreting passages like (1), and even go so far as to contemplate additional assumptions that would license such inferences for passages like (2), illustrates that the need to establish coherence is basic to our natural language understanding capacity. Just as we attempt to identify syntactic and semantic relationships when presented with a sequence of words in an utterance, we attempt to identify coherence relationships when presented with a sequence of utterances in a discourse. The establishment of coherence is hence a powerful mechanism that allows us to communicate, and conversely understand, considerably more meaning than that conveyed by individual sentences alone. In this sense, the meaning of a discourse is greater than the sum of the meanings of its parts.

Having argued for the centrality of coherence establishment to language interpretation, we would naturally like to have a theory that characterizes the possible ways in which successive utterances can be connected to form a coherent discourse. Several researchers have in fact attempted such a characterization, in which a set of connections is enumerated as a list of *coherence relations* (Halliday and Hasan 1976, Hobbs 1979, Longacre 1983, Mann and Thompson 1987, Polanyi 1988, Hobbs 1990, inter alia; see Hovy (1990) for a compendium of over 350 relations that have been proposed in the literature). I will likewise present a list of relations here, but one in which these relations are seen to arise from a fundamental cognitive distinction. This distinction was first articulated by the philosopher David Hume in his *Inquiry Regarding Human Understanding* (1748), who makes the following general statement concerning the types of connections that can hold between ideas.

"Though it be too obvious to escape observation that different ideas are connected together, I do not find that any philosopher has attempted to enumerate or class all the principles of association—a subject, however, that seems worthy of curiosity. To me there appear to be only three principles of connection among ideas, namely *Resemblance*, *Contiguity* in time or place, and *Cause* or *Effect*."

Hobbs (1990, pp. 101-102) was the first to point out that Hume's principles could be used as a basis for categorizing coherence relations, but he did not pursue such a categorization in depth. I will in fact argue for such a categorization. Indeed, the position I take is a strong one: That Hume's categories comprise a small set of basic types of cognitive principles that, when applied to the domain of discourse interpretation, give rise to such relations.

Upon having offered my theory of coherence, I will then utilize it as a fundamental component within analyses of five diverse and well-studied linguistic phenomena. I will describe how Maslow's maxim has raised its head in the literatures of these areas, showing that each can be characterized by two contradictory properties: (i) an implicit assumption that the data can be explained with a uniform set of tools (e.g., solely by syntactic rules, semantic mechanisms, or uniform discourse-level strategies), and (ii) a set of data that would seem to defy this assumption. In each case, I argue that the data can be explained with a cross-modular theory that interfaces a relatively straightforward account of the properties of the linguistic phenomenon in question with the effect of discourse-level interpretation processes used to establish coherence. I synopsize the book in greater detail in the following section.

Overview of the Book

I will begin by presenting my neoHumian categorization of coherence relations in detail in Chapter 2. Therein a core set of relations is proposed along with a specification of the constraints that each imposes. For instance, the *Explanation* relation will be seen to capture the type of connection that we established for passage (1) and sought to establish for passage (2).

Explanation: Infer P from the assertion of S_0 and Q from the assertion of S_1, where normally $Q \rightarrow P$.

A clear pattern emerges from these definitions that accords with Hume's principles in terms of two criteria: (i) the type of arguments over which the constraints of each relation apply, and (ii) the type of inference processes that are used to establish these constraints.

In Chapter 3 I address the linguistic phenomenon of VP-ellipsis, illustrated in example (3).

(3) George likes his mother, and Al does too.

I focus particularly on the fundamental question that serves as the starting point for any analysis of VP-ellipsis: the level of linguistic representation at which it is resolved. We will see that analyses that operate at

the level of syntax, which generally require parallel syntactic structure between the antecedent and elided verb phrases, predict the unacceptability of examples like (4), in which the antecedent clause has been passivized.

(4) # This problem was looked into by John, and Bob did too. [look into the problem]

Analyses that operate at a purely semantic level of representation do not predict this unacceptability. Examples like (5) are in fact acceptable, however, as predicted by semantic, but not syntactic, analyses.

(5) This problem was to have been looked into, but obviously nobody did. [look into the problem]

I show that the seemingly contradictory data exhibits a pattern that correlates with the type of coherence relation holding between the antecedent and elided clauses. The account specifies the interaction between two independently-motivated sets of properties: the syntactic and referential properties of VP-ellipsis, and the properties of the processes for establishing each type of relation. This interaction will be seen to predict the pattern found in the data.

Chapter 4 addresses the gapping construction, illustrated along with its ungapped counterpart in examples (6a-b), taken from Levin and Prince (1986).

(6) a. Sue became upset and Nan ∅ downright angry.
    b. Sue became upset and Nan became downright angry.

Gapping is similar to VP-ellipsis in that material has been elided from within a clause. In this case, however, only bare constituents remain in the clause, none of which is a stranded auxiliary. The behavior of gapping strongly suggests that it is primarily a syntactically-governed phenomenon, and hence most previous approaches have addressed it at that level. However, these approaches fail to predict an interesting fact about examples (6a-b). Example (6b) can be understood two ways, depending on the type of connection that is inferred between the two clauses: Either Sue and Nan became emotional independently, albeit perhaps in response to the same external stimulus (the 'symmetric' reading), or Nan's becoming angry was *caused* by Sue's becoming upset (the 'asymmetric' reading). Example (6a), on the other hand, has only the first of these readings. I will show that the ability to gap in such cases depends not only on the syntactic properties of the clauses, but also on the type of coherence relation that holds between them. As in the case of VP-ellipsis, I will show how this dependency is predicted by the interaction between two independently-motivated sets of properties:

the syntactic and referential properties of gapping (which differ in important respects from those of VP-ellipsis), and those of the inference processes underlying the establishment of coherence relations.

In Chapter 5 I address extraction from coordinate structures, with particular reference to the Coordinate Structure Constraint (CSC) originally proposed by Ross (1967).

> In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

Ross proposed the CSC to account for the ungrammaticality of sentences such as (7).

(7) * What book did John buy and read the magazine?

Counterexamples to the CSC are well-attested, however. For instance, Ross himself notes that examples such as (8) are acceptable, in which extraction has occurred "across-the-board", that is, out of all conjuncts.

(8) What book did John buy and read?

Furthermore, Ross also points out that example (9a) is acceptable despite the fact that extraction occurs out of only the second conjunct, and Goldsmith (1985) and Lakoff (1986) note that examples (9b-c) respectively are acceptable despite the fact that extraction occurs out of only the first conjunct.

(9) a. Here's the whiskey which I went to the store and bought.
   b. How much can you drink and still stay sober?
   c. That's the stuff that the guys in the Caucasus drink and live to be a hundred.

I will show that this data patterns with my neoHumian categorization of coherence relations. I will then demonstrate how this pattern results from the interaction between independently motivated conditions on extraction and the constraints that need to be met in establishing each type of coherence relation.

Chapter 6 addresses the problem of pronoun interpretation. In sentences (10a-b), adapted from an example from Winograd (1972), the pronoun *they* is typically understood to refer to *the city council* and *the demonstrators* respectively. This difference is presumably due to the fact that these reference assignments make the scenario described in each passage most plausible, especially considering the fact that the syntactic conditions are the same in each case.

(10) The city council denied the demonstrators a permit because...

a. ...they *feared* violence.

b. ...they *advocated* violence.

However, in example (11), informants universally interpret the (unaccented) pronoun *her* to refer to Hillary Clinton, even though the more semantically plausible referent is Margaret Thatcher. Similarly, example (12) tends to generate a garden-path effect, in which hearers initially identify John as the referent of the pronoun instead of Bill, despite the fact that subsequent information suggests the latter as the referent.

(11)   Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships her.

(12)   John can open Bill's safe. He made a promise to get the combination changed soon.

Once again, I will show that these examples and others discussed in the pronoun interpretation literature display a pattern with respect to my neoHumian categorization of coherence relations. I then demonstrate how the inference processes underlying the establishment of these relations interact with the linguistic properties of pronouns – particularly their tendency to signal immediate interpretability with respect to a currently salient referent – to predict this behavior.

Chapter 7 addresses the problem of tense interpretation. Previous approaches have attempted to attribute the forward movement of time normally inferred between successively-described events in a narrative to the meaning of tense itself, typically the simple past. Such approaches fail to account for the fact that the simple past is compatible with any temporal ordering between events; for instance, the two events in examples (13a-d) are understood as displaying forward movement of time, backward movement of time, identical times, and no implied ordering, respectively.

(13)   a. Max slipped. He spilt a bucket of water.

b. Max spilt a bucket of water. He tripped on his shoelace.

c. Max spilt a bucket of water. He spilt it all over the rug.

d. Max spilt a bucket of water. John dropped a jar of cookies.

I present this and additional data that is problematic for two types of approach to tense interpretation. I then provide an account that combines a theory of tense with the constraints imposed by coherence relations that correctly predicts this data, as well as having additional advantages over other analyses.

Finally, Chapter 8 concludes with some final thoughts and suggestions for future research directions.

As the foregoing phenomena are rather varied, the reader may be approaching this book with an interest in only a subset of them. I have tried my best to organize the book so that only Chapter 2 is a prerequisite to reading any of the individual analyses in Chapters 3, 5, 6, and 7. Chapter 4 is also dependent on the central parts of the analysis of VP-ellipsis described in Chapter 3. However, I cannot resist the temptation to encourage the reader to read the book in its entirety, since I would argue that the fact that the same theory of coherence can be used to address outstanding issues in a diverse range of areas adds to the strength of the underlying argumentation in each particular case.

It has also been my intention to present the analyses in as theory-neutral a manner as I possibly could. Whereas many underlying theoretical assumptions unavoidably remain, my main goal is to convince the reader that coherence establishment processes must be accounted for in analyses of the linguistic phenomena addressed herein, and by extension, of other interclausal phenomena that have yet to be analyzed in these terms. I therefore sought to impose as few theoretical obstacles between the reader and this message as possible, by including only those concepts that I found necessary to make the analyses concrete. I hope that the insights expressed herein will find their proper influence in a wide range of contemporary linguistic frameworks and perspectives.

Who Might Find this Book of Interest?

This book draws significantly on work in a broad range of theoretical traditions, bringing together insights from formal, functional, cognitive, and computational linguistics. I believe that this book will be equally of interest to practitioners in each of these areas, and furthermore that it demonstrates ways in which the work of theorists in each area can be seen to tie together. With respect to areas of language processing, this book bears heavily on work in syntax, semantics, pragmatics, and discourse processing. The influence of Hume's work on my theories also make the book of interest to philosophers of language, as it illustrates a set of concrete linguistic applications of a broad and influential piece of philosophical thought. The analyses are also rooted heavily in cognitive science and artificial intelligence – particularly with respect to how we perceive our world as coherent and how we focus our attention while doing so – making the book of interest to practitioners of these fields. Finally, many of the theories expressed herein contradict prevailing assumptions in psycholinguistics, and at the same time are empirically testable themselves. As such, the work will be of interest to psycholinguists working in all areas of language processing.