What is dialogue coherence?  

Bill Mann, June 2002

Dialogue coherence has been identified in various ways. In the literature there are many views of coherence in language, and of dialogue coherence in particular. To appreciate them, a distinction must be made between coherence and cohesion. Cohesion represents connectedness of text that arises from the use of particular linguistic devices such as pronouns, anaphoric reference, patterns of topic or focus and many more. In detail it is language dependent. The classic exploration of cohesion is (Halliday and Hasan 1976).

In contrast, coherence is a kind of impression that arises (or not) in a person who attempts to understand particular language use. In general it is not language dependent, in the sense that a translation of a coherent text is usually coherent, even if the cohesive devices of the source text are absent in the target language. Although not all of the views of coherence make this distinction, it is essential for comparisons.

Studies of dialogue from various points of view are numerous. There are dozens of technical fields with the word communication in their names. (Craig 1999) Many of them study dialogue. There are structural views, communication views and many others. Linguistics tends to produce structurally oriented studies, but not exclusively.

Even restricting attention to studies of dialogue coherence, there are many radically different viewpoints. Conversational Coherence: Form, Structure and Strategy (Craig and Tracy 1983) is a particularly relevant collection, now somewhat dated but representing ideas that persist in the wider literature. In this book, and in the wider literature as well, the distinction between coherence and cohesion is often not made. Studies of coherence are often really about abstract cohesive devices, in the sense of (Halliday and Hasan 1976). Some studies (Ellis 1983; Goldberg 1983) assume that coherence is produced by design, by appropriate use of cohesive devices.

In some studies coherence is equated to topic continuity or to the appropriate use of topic shifting devices (Crow 1983; Sigman 1983). In others it is seen as conformity to expressive rules. Grice is often interpreted as believing that conversational coherence is based on rule (or maxim) following. (Grice 1975) Still others see coherence as an identifiable outcome of rule governed (social or linguistic) behavior. (Goldberg 1983) Hawes defends his views against this idea. (Hawes 1983)

Some studies of coherence in dialogue assume that people pursue a tacit goal of being coherent, in addition to any other goals, when they interact. It is a benefit which they

---

1 The view represented below arose in part out of email discussions. The participation of Colin Matheson and Brian MacWhinney is gratefully acknowledged.

2 Terminology in this area is of course irregular. Conversation is another designator of intellectual focus, and interaction is often the key term. Exceptions aside, dialogue is generally taken to be a kind of interaction and conversation is generally taken to be a kind of dialogue. This memo is concerned with immediate dialogue, including face to face interaction, telephone calls and similar kinds, rather than less immediate interaction such as the exchange of letters or email. Dialogue specifically includes interaction of more than two parties.
seek. (Hopper attributes this orientation to interpreters of language, but not producers. (Hopper 1983)) Others see coherence as an obligation that is attached to interaction. It is an added duty. (Sanders sees this view as widely accepted, and defends his views against it. (Sanders 1983))

Some studies equate coherence with propositional consistency, see (Goldberg 1983) for citations. Others see coherence as a kind of summary impression that is a side effect of understanding an interaction, an understanding that is enabled by the processes that ordinarily govern interaction. (Sanders 1983)

The (Craig and Tracy 1983) book incorporates an admirable attempt to make the various approaches comparable. All of the chapter authors were given one particular 30 minute dialogue (included in the volume) and told to relate their approaches to it. In a volume summary, apropos of this memo, the editors note that

“Conversationalists’ goals must play a central role in any adequate explanation of discourse production and interpretation in conversations.” p. 22.

Global approaches to coherence, ones that attempt to address entire texts or dialogues, are often associated with some notion of genre or tradition, such as Rummelhart’s story grammars (Rumelhart 1975) or Schank and Abelson’s scripts. (Schank and Abelson 1977). These approaches do not account for coherence in dialogues where no stable tradition, shared by the participants, is in use. Hostage negotiation, medical interviews, emergency telephone calls, unfamiliar laboratory conversational tasks, telephone help services and many other conversational situations give the impression of coherence without a corresponding stable tradition or guidance from genre.

(DMT addresses whole dialogues, but without resemblance to those approaches.)

There is more recent progress in many aspects of understanding dialogue. A rich array of formal approaches has been built on the Discourse Representation Theory of Kamp and colleagues (Kamp and Reyle 1993; Traum 1994). Agency theory, along with various vigorous efforts to develop data annotation methods, are also producing insightful views of natural dialogue. Some of these approaches do not use an explicit notion of coherence, but they often use consistency criteria as a defense against incoherence.

Clearly there is no consensus on the nature of coherence in dialogue. Although comparing views is typically difficult, as researchers we always find some views more credible than others. We may be able to make certain alternatives more distinct by an analogy concerning oral dialogue. What is the status of breathing? Do people breathe in dialogue because they believe it will make the dialogue more beneficial? Or is there a duty to breathe in dialogue? Is breathing simply following a tradition, or an attempt to perform smoothly? Are there rules of dialogue that would, for example, make a dialogue ill-formed if it did not involve breathing? Or is breathing regulated by processes that interact with the speaking processes? DMT is designed following assumptions that most resemble the latter alternative.

The view being explored here rejects the genre, topic continuity, and rule following views as insufficient. Likewise, the view of coherence as a separate duty or goal of
conversation, added to all of the other social and achievement goals, is also insufficient.

Coherence represents integrity of intentions and information that, in the processes of interpretation, are imputed to the text creator – here the dialogue participants. This integrity includes consistency, sequential continuity and apparent completeness of the imputed intentions and their representation in language. The definition below is intended as a draft which demonstrates an alternative approach.

*Begin definition.*

Ontologically, there are linguistic experiences, and we can identify reading of monologues, reading of dialogue transcripts and experiencing dialogue (multi party, with self silent or speaking) as 3 of these. Each of these may be accompanied by an impression that the thing is coherent, or that it is not. (There may be a finer grain. In longer texts we can talk about coherent and incoherent parts. I grew up in New Jersey.) In each of these cases the kind of experience is comparable, so that there is no reason to regard coherence as divided into cases by the mode of language use.

So ontologically coherence is an aspect of reception of language. More specifically, it is an aspect of understanding the language use that is in view. (If we somehow receive some text that is in an unknown language, we do not experience either coherence or incoherence from it.)

It is thus a kind of personal experience, perhaps inseparable from the text (as above) to which it applies. So, my saying that a certain text T is coherent is saying that I find it coherent, or that people who know that language generally find it coherent, or something similarly experiential. Thus coherence does not inhere in texts.

However, people can often estimate how people will respond to particular texts. On that basis it makes sense to talk about a text being coherent or incoherent.

Coherence has three aspects: a separability aspect, a completeness aspect, and a consistency aspect.

Separability has to do with whether every experienced part has an evident part in the whole (or, possibly, with ambiguity, evidently has a part in the whole.) It is related to the absence of non sequiturs.

Completeness has to do with the experience that there is evidently some part that is missing from the.

Consistency does not necessarily mean logical consistency. There is as yet no consensus on a logical scheme that would clearly be adequate and usable. Rather it is an absence of gross incongruity, of cognitive clash between represented parts. Clearly this must be judged one participant at a time.

*End definition.*

Is this definition coherent up to this point to you? I doubt it. The completeness paragraph probably was understood perfectly well but it lacked completeness coherence. And the
remark about New Jersey, although it might possibly be seen as consistent with the rest, it probably did not appear to have an evident part in the whole, violating separability coherence. You can derive an entirely coherent definition from this one, we hope, by deleting the sentence about New Jersey and adding the word “text” in the right place in the completeness paragraph.

So coherence includes separability coherence, completeness coherence and consistency coherence.

For dialogue, in two-party immediate cases at least (which is the sort that is in view) the means by which the impression of coherence arises are more complex than for monologue. Much of the complexity comes from the diversity of situations in which dialogue is used. (More complexity will arise in multiparty cases.)

Situations impose strong tendencies on what occurs. As a result, what occurs can be viewed in terms of pure distinct types of dialogue. However, natural dialogues do not always exemplify single pure types. Dialogues often contain a mixture of types. Thus the notion of a taxonomy of dialogue types is inherently approximate and partial, and to that extent defective, and an account of a mixed dialogue may require mixed theories.

Coherence in dialogue is predominantly but not entirely based on attribution of intentions in a manner resembling what RST (Mann and Thompson 1988) suggests for monologue. There are certain parts that do not, in hindsight, form a part of the interactional flow of dialogue, but which occur. For example, if during a sentence from A, B says "Yes, but we..." it may not always be possible to assess underlying intentions of B's remark. Participants struggle for control, and failed initiatives leave debris in the transcript.

Dialogues begin and end, and endings are attempted, using socially familiar procedures that are, part by part, more ritualistic than intention based. "Thanks to you, too." would exemplify this. Other procedures come into management of the media: "Say that again; the kids were screaming," exemplifies this. So knowledge of such procedures comes into the judgment of separability coherence.

If we define "task oriented" as meaning that the active participants of the interaction take up evident joint intentions (in the sense of Clark, mostly implicit in Tuomela) (Clark 1996; Tuomela 2000), then coherence in these dialogues or dialogue fragments can be based on attribution of intentions and goal pursuit. For other dialogues or fragments, it is possible to estimate the superficial intentions of participants even where there is no evident joint intention.

Considering intentions and coherence, the most active concern is consistency. The language used by each individual suggests intentions that can plausibly be seen as commitments of that individual, and we expect such commitments to form a compatible collection. Commitments of one participant need not be compatible with those of another. However, when there is commitment to a joint intention, it must for consistency assessment purposes be seen as an intention of each committed participant.

There is a related matter about how coherence is experienced. If we say text T is coherent, sometimes it is useful to ask “Coherent to whom?”
A dialogue can be coherent to one observer or participant and not coherent to another.

For example, a medical interview, from the patient's point of view can be incoherent. The physician may be considering two or three diseases as potentially being the diagnosis of the patient's condition. At the same time, the patient may not know what diseases are being considered, why certain questions are asked, and what context of judgment of the meaning of the questions is relevant to answering them. Another physician, seeing the dialogue or the transcript, may understand the physician's intentions completely and regard the interview as coherent. But the patient does not know those intentions and can ascribe only apparent medical relevance to them, based on the situation. Such ascriptions are 100% assumption, not derived from the specific text. So the patient does not see the interview as coherent.

If a voyeuristic doctor asks some questions for his own amusement, the observing physician might recognize that fact and judge the interaction to be incoherent, and yet the patient might not notice any incoherence.

At a conference in 2001 Brian MacWhinney put forth, as one of the fundamental orientations to language understanding, the idea that language participants are "running lots of simulations" of each other, estimating what effects particular generated language might have, and shaping (selecting among alternatives) the language they produce so that it will have desirable and satisfactory effects. This is a view from cognitive psychology that MacWhinney is willing to apply at all scales, including the dialogue turn or utterance (MacWhinney 2002).

This view supports the assumption that we experience coherence by estimating the cognitive states of language users. Such estimates apply to language actually used, to choices between alternatives (language used and language avoided) and even to intentions to say certain things in the near future.

Obviously if it is correct and consequential in empirical studies, then it will be needed in our long term designs of how we get machines to accomplish language understanding.

This seems to me to be in effect the same thing that Mikhail Bakhtin (Bakhtin 1981; Bakhtin 1994) is saying when he calls certain texts "dialogic;" (MacWhinney agrees.) We may not be inclined to implement all of Bakhtin's views in computer programs. We ought to be prepared to implement this one, since it has a strong potential for organizing the many processes that we need for modeling language understanding, language generation and participation in activities as an agent.

For dialogue and interaction, the study of coherence complements the study of understanding, of generation or participation, of formal structure, of syntax and its entailed aspects.

Coherence as a research focus tends to illuminate the large scale interaction of the parts of a text, and thus it is potentially able to contribute to accounting for aspects of language function that have no natural upper limits on scale. These include disambiguation, in-text specialization of words and phrases, anaphora and many other functions.
Coherence as a research focus also makes it possible to illuminate the linguistic diversity of texts, with the possibility of classifying texts according to the various kinds of coherence principles that apply to them. For dialogue, we would expect “conversation” to be represented by one or more sets of principles, “task oriented dialogue” by other sets, and yet other sets of principles as well.

Thus the study of coherence has a distinctive place in linguistic studies.

References


MacWhinney, Brian (2002). (Personal Communication.)


Sigman, Stuart J. (1983). Some Multiple Constraints Placed on Conversational Topics In
