

Narrative Coherence in Bloomfield's Menominee Texts

Sudha Arunachalam
University of Pennsylvania

1 Introduction

Leonard Bloomfield's fieldwork on Menominee in the 1920's and 30's yielded a collection of narratives told by several Menominee speakers. In his 1927 "Literate and Illiterate Speech," Bloomfield expresses surprise at the Menominee's metalinguistic judgments about interspeaker variation in the community. "The Menomini will say that one person speaks well and another badly," he writes, "that such-and-such a form of speech is incorrect and sounds bad, and another too much like a shaman's preaching or archaic" (89).

The criteria he discusses that are involved in speakers' judgments of "good" and "bad" speech are primarily of pronunciation and grammar, such as confusion of short and long vowels and use of inappropriate inflections. However, because the features of good speech are determined within the community rather than by a standard provided by orthography, many factors will likely play into a speech community's notion of "good" speech. Bloomfield thus sums the distinction between speakers to be a combination of sociocultural factors:

The nearest approach to an explanation of "good" and "bad" language seems to be this, then, that, by a cumulation of obvious superiorities, both of character and standing, as well as of language, some persons are felt to be better models of conduct and speech than others. Therefore, even in matters where the preference is not obvious, the forms which these same persons use are felt to have the better flavor. (93)

One factor in judging speaker competence, especially if storytelling is a big part of culture, may be how good the person is at narrating a story, more along the lines of pragmatic than grammatical competence. This study asks whether Menominee speakers varied along this dimension, whether some speakers were simply better at telling stories than others, and whether this factor correlated with the speaker assessments Bloomfield gathered.

Assessing what makes a good narrative is difficult; one relevant notion is coherence. Does the speaker jump around from one topic to another, making it difficult for the hearer to follow? To investigate this, we need a theoretical basis for measuring coherence in discourse. Centering Theory (Grosz, Joshi, and Weinstein 1995) provides the armature we need to investigate interspeaker variation in narrative coherence. This theory uses an algorithm which takes into account the way in which entities are referred to in discourse and yields a measure of how coherently the speaker shifts back and forth from talking about different entities. The texts I analyze are from one of Bloomfield's "best" and one of his "worst" speakers, Little-Jerome and White-Thunder, respectively. The analysis will investigate whether the two texts differ in coherence, and will suggest that judgments of narrative coherence are distinct from the judgments the Menominee community makes about good and bad speech.

In the following sections I will introduce Centering Theory and my analysis of two Menominee texts.

2 Centering Theory

Centering allows us to account for differences in perceived coherence between discourses on the basis of the form of referring of expressions and how discourse participants' attentional state is

shifted by discourse structure. The model was developed by Grosz, Joshi, and Weinstein (1995), drawing on earlier work by Joshi and Kuhn (1979), Joshi and Weinstein (1981), and Grosz and Sidner (Grosz 1977; Sidner 1979; Grosz & Sidner 1986). The version of Centering I will be using is the result of subsequent work by the original authors as well as others (e.g. Brennan, Friedman, and Pollard 1987; Walker, Joshi, and Prince 1998). First some necessary definitions will be presented.

2.1 Definitions

Each semantic entity that is part of a discourse is called a *center*. The set of centers in each utterance U_i is the set of *forward-looking centers*. The forward-looking centers for each utterance are ranked in terms of discourse *salience* (more will be said about this ranking later), and the most highly ranked forward-looking center in an utterance is the *preferred center* (Cp). The preferred center is a prediction as to what the next utterance will be about. Another special member of the set of forward-looking centers of an utterance is the *backward-looking center* (Cb); this roughly corresponds to what is generally referred to as the *topic* (Reinhart 1981; Horn 1986), or what the utterance most centrally concerns. It links the current utterance with the previous discourse. The backward-looking center of an utterance U_i is the highest-ranked center of the previous utterance U_{i-1} that is also in the current utterance. Because the backward-looking center is extremely salient in U_i , the Pronoun Rule (sometimes called Rule 1) states that if there is a pronoun in an utterance, then the Cb of that utterance is also realized as a pronoun. (The Pronoun Rule is extended to apply also to null elements in Turan 1995.) The rule reflects the observation that salient entities need not be expressed as a full noun phrase, as these entities are already in the forefront of the hearer/reader's attention, and their reference is easily determined without a great processing load.

2.2 Ranking and Transitions

The ranking of the set of Cfs is thought to be language-specific, as different languages have different ways of realizing salience grammatically. One possible hierarchy, and the one that has been used for English (Brennan, Friedman, & Pollard 1987), is based on grammatical role: Subject > Object (Direct and Indirect) > Other.

The interaction between the Cb and the Cp gives us a measure of the smoothness of the *transitions* between the current utterance U_i and the previous utterance U_{i-1} , which in turn gives us an overall picture of coherence. The transitions in Centering Theory are defined in the table below.

Table 1. Transitions from U_{i-1} to U_i

	$Cb(U_i) = Cp(U_i)$	$Cb(U_i) \neq Cp(U_i)$
$Cb(U_i) = Cb(U_{i-1})$ OR no $Cb(U_i)$	Continue	Retain
$Cb(U_i) \neq Cb(U_{i-1})$	Smooth-Shift	Rough-Shift

The transitions are based on whether the Cb ("topic") is the same from one utterance to the next, and whether the Cb of the current utterance is also the Cp of the current utterance (i.e. whether the topic of the current utterance is projected to be the topic of the subsequent utterance). When the transition is Continue, a particular entity is being talked about and will continue to be talked about in the following utterance. A Retain transition indicates that a new entity is projected to be the topic of the subsequent utterance; the speaker indicates that shift by realizing the Cb of the current utterance is a lower-ranked position. The two kinds of shift transitions indicate that the topic has changed between the previous and current utterance, but in a Smooth-Shift, we project that the speaker will continue talking about the new entity, while in a Rough-Shift, the topic is

projected to shift once again. The transitions are hypothesized to be ordered, such that Continue transitions are preferred to Retain transitions, which are preferred to Smooth-Shifts, with Rough-Shifts being the least preferred, i.e. least coherent.

The idea behind the above hierarchy is that the hearer/reader will be able to follow a discourse with all Continue transitions with less processing effort than a discourse with all Rough-Shifts, the reason being that the same entity remains in the forefront of attention throughout in the former situation. This should not, however, be thought of as indicating the ideal discourse; a discourse with all Continue transitions would be incredibly dull. (*John* did X. Then *he* did Y. *He* went to Z....) In normal discourse, Retains and Smooth-Shifts are also found, though interestingly Rough-Shifts are virtually non-existent (Di Eugenio 1998, Hurewitz 1998). The number of types of transitions, however, can be used to gauge coherence of discourse. In fact, Miltsakaki and Kukich (2000), using just a Centering algorithm, were able to predict with a high degree of accuracy the scores given by teachers to student essays. This idea will be crucial in our analysis of narrative coherence in Menominee.

2.3 Sample Analysis

The following analysis of a brief English sample, taken from Walker, Joshi, and Prince (1998), will illustrate how Centering Theory works.

- (1) *Jeff helped Dick wash the car.*
 Cb: none
 Cf: [Jeff, Dick, car]
 Cp: [Jeff]
 Transition: none (no Cb)
- (2) *He washed the windows as Dick waxed the car.*
 Cb: [Jeff]
 Cf: [Jeff, windows, Dick, car]
 Cp: [Jeff]
 Transition: Continue
- (3) *He soaped a pane.*
 Cb: [Jeff]
 Cf: [Jeff, pane]
 Cp: [Jeff]
 Transition: Continue

In the first sentence, the subject is Jeff, and so we predict that Jeff will also be the preferred center of the second sentence, which is the case; thus we have a Continue transition. In sentence 3, assuming that Jeff is the intended referent of the pronoun *he*, we again have a Continue transition. If the intended referent were Dick, the sentence would be much more difficult to understand; the hearer would require some overt disambiguation to be able to resolve the reference in this way. This fact is reflected in Centering Theory because the transition would be a Smooth-Shift, which is a less coherent transition than a Continue according to the transition hierarchy discussed above.

In the next section, we will turn to two Menominee narratives and the Centering analysis.

3 Centering Analysis of Menominee Narratives

3.1 The Texts

Centering thus provides a concrete way of measuring coherence in discourse, and a means for

answering our questions about Menominee narratives. I examine two of Bloomfield's texts, from one of his best speakers, Little-Jerome, and one of his worst, White-Thunder, to see whether there is significant variation in coherence (as presently defined) in their storytelling. Bloomfield writes about these two:

Little-Jerome: "a true bilingual. He speaks both English and Menomini with racy idiom, which he does not lose even when translating in either direction. He contrasts strikingly with the men (usually somewhat younger) who speak little English and yet bad Menomini" (1987:91).

White-Thunder: "his Menomini is atrocious. His vocabulary is small; his inflections are often barbarous; he constructs sentences on a few threadbare models. He may be said to speak no language tolerably" (1987:91).

The two narratives analyzed are folk stories of Menominee culture. They feature only a few characters, which are humans and animals. White-Thunder's narrative, "Tales of the Ancient Time," is a collection of a few stories. Each vignette is treated as a separate discourse segment for the analysis. Little-Jerome's narrative, "Me' napus and the Tree-Cat," describes a single episode.

3.2 Centering Assumptions and Ranking

For the purposes of the Centering Analysis, direct speech (dialogue) is ignored. This is because Centering Theory as laid out here cannot adequately deal with frequent speaker shifts; the dialogue comes out to be less coherent than it actually seems to be. Presumably speakers use various prosodic devices to signal to their listeners that different speakers are being quoted.

Determining the ranking is the most interesting and difficult part of doing a Centering analysis on a new language. While the ranking most commonly used for English is based on grammatical role, other factors have been found to influence ranking in other languages, such as topic marking for Japanese (Kameyama 1986) and surface word order in German (Rambow 1993). Ranking for a given language is usually determined by assessing the effects of including or excluding particular features in the analysis of carefully selected passages. Common sense is of course paramount; if a passage seems coherent but the analysis yields several rough-shifts, then the ranking should be reconsidered. For the present study, determining ranking is critical, as we must avoid circularity in determining coherence. We don't want to say that Text A is more coherent than Text B based on a ranking developed by looking only at Text A. The procedure I used was to try out different rankings based on factors I thought might play a role, and to analyze both texts, looking crucially at whether one text simply failed to be as coherent in general across different rankings. Below we will look at some of the factors considered.

The basic ranking used is grammatical role, if no other factors are relevant. Thus, following work on other languages, subjects are ranked above direct objects. The ranking of other roles does not turn out to have an effect on the analysis.

One fact about Menominee affecting ranking is that it makes a morphological distinction between proximate and obviative, the former used for the "topic of discourse, the person nearest the speaker's point of view, or the person earlier spoken of and already known" (Bloomfield 1962:38). Only one animate third person entity is proximate in any given context; all others are marked with the obviative. One of the key phenomena we are interested in is thus grammatically encoded in the language. I therefore rank proximate Cfs above obviative ones, as speakers may take more freedom with grammatical role of entities because the proximate/obviative marking will disambiguate.

Animacy is another factor considered in determining ranking. I did not find precedent in other work in Centering for counting animate and inanimate entities as distinct, but both texts

were found to be maximally coherent under the following ranking: human >> animal >> inanimate. (While a different ranking would have influenced coherence of each text individually, it would not significantly alter the differential measure of coherence between them.) Some inanimate entities are grammatically animate in Menominee, such as *osa:qsekamesow* ‘dumpling,’ and while it would be interesting to see where these entities fall on the hierarchy, the present analysis unfortunately does not bear on that issue.

Pronouns are rare in the texts; most noun phrases are null. Disambiguation is not usually difficult, however, as the verb bears marking of both subject and object. Null noun phrases are thus treated as pronouns, following Turan’s (1995) analysis for Turkish.

A summary of the ranking used is thus:

subject >> direct object
proximate >> obviative
human >> animal >> inanimate

3.3 Results and Discussion

3.3.1 Comparing Little-Jerome and White-Thunder

The results of the Centering Analysis are shown in Table 2 below. Both texts are highly coherent, as they overwhelmingly consist of Continue transitions. White-Thunder’s narrative is in fact more coherent by these standards than Little-Jerome’s. Both passages have more Continues than Retains, and more Retains than Smooth-Shifts, and neither has any Rough-Shifts, following nicely the transition ordering mentioned in section 2.2. The percentage of Retain transitions in each narrative is not significantly different ($\chi^2=2.12$, $p=0.5$), and the percentages of Continue and Smooth-Shift transitions, while significantly different from each other ($\chi^2=4.08$, 5.29 , respectively), are still quite close and significance disappears at $p=0.2$. The analysis shows, then, that the Menominee’s judgments of goodness of speech are not tied to factors captured in a Centering analysis. Both speakers were equally adept at manipulating sentence structure to reflect shifts in topic, and achieved a high level of coherence with respect to their choices of referential expression.

Table 2. Percentages of Centering transition types in two narratives

	Little-Jerome (Text 85)	White-Thunder (Text 51)
CONTINUE	61.9%	77.8%
RETAIN	23.8%	16.7%
SMOOTH SHIFT	14.3%	5.6%
ROUGH SHIFT	0%	0%

3.2 Issues in Centering

As no other Centering analyses have been done on Menominee to date, it is useful to consider theoretical and practical issues that arise when dealing with a language so different from English. This section will consider, as an example, one such concern.

A question for Centering Theory is whether information about reference that is encoded grammatically in ways other than pronouns/full NPs/nulls has the same status as these referential forms. For example, in languages like Menominee, both the subject and direct object are marked on the verb, perhaps requiring less of the hearer in terms of attending to grammatical role of null elements. A further, even finer, distinction in Menominee is the proximate/obviative one mentioned earlier. These verbal inflections may render unnecessary reliance on other factors. The

following excerpt from White-Thunder’s “Tales from the Ancient Time,” marked to illustrate the Centering analysis, illustrates the role of the proximate/obviative distinction. (All referring expressions are underlined, with nulls marked with an index, and the forward- and backward-looking centers have *Cp* and *Cb* next to them, respectively.)

- a
iniwi´n ḡ CpCb as kãqts-ãyõ´sin´kut ani´nuh ḡ Cb kãta-neqni´kut;
then that greatly-he(OBV).laughs.at.him that in.intention-he(OBV).kills.him
Then he who was intending to slay him, laughed aloud at him;
CONTINUE
- b
keqtsi´h niw ḡ CpCb uhsã´pumikin;
near [precision] he(OBV).looks.at.him.from.there+QUOT
from close by that other was observing him;
CONTINUE
- c
kn teh ḡ CpCb ḡ unãwã´nan.
not however he.sees.him(OBV)+NEG
but he did not see the other.
CONTINUE
- d
um´si´h niw ḡ Cb s nõ´htawatsin,
here [precision] [aor] he.hears.him(OBV)
The instant he heard him,
- e
iniwi´n niw ḡ CpCb às-ku´ahnet imis nipí´hih ḡ s kō´kēt.
then [precision] thither-he.jumps over.yonder in.the.water [aor] he.dives
he leapt, and dived into the water.

Without the proximate/obviative distinction, the above segment would be virtually incomprehensible, and would require a good amount of context to be ultimately deciphered. The other factors used in Centering, e.g. grammatical role, are no help here, and in fact predict incorrect reference assignment if taken alone. Note that in line c, where without the proximate/obviative markings Centering would predict the subject to remain the *Cp*, Bloomfield’s translation reads “the other,” signalling a shift in reference. Clearly, a Centering analysis must pay close attention to the particular grammatical devices used by languages in determining ranking.

4 Conclusion

The above analysis shows that narratives produced by one of Leonard Bloomfield’s “best” and “worst” speakers do not differ in coherence as measured by transition types in Centering Theory. This suggests that the Menominee dissociated different kinds of linguistic competence (roughly, pragmatic/narrative competence from phonological/morphological/syntactic competence) when making judgments about speakers’ abilities.

References
Bloomfield, Leonard. 1962. *The Menomini Language*. New Haven: Yale University

Press.

Bloomfield, Leonard. 1987 (orig. 1927). Literate and illiterate speech. In Charles F. Hockett, ed., *A Leonard Bloomfield Anthology: Abridged Edition*. Chicago: University of Chicago Press, 84-93.

Brennan, Susan, Marilyn W. Friedman, and Carl J. Pollard. 1987. A Centering approach to pronouns. *Proceedings of the 25th Annual Meeting of the ACL*, Stanford, CA.

Di Eugenio, Barbara. 1998. Centering in Italian. In Walker, Joshi, & Prince, eds.

Grosz, Barbara J. 1977. *The Representation and Use of Focus in Dialogue Understanding*. Technical Report No. 151. Menlo Park, CA: SRI International.

Grosz, Barbara J., and Candace L. Sidner. 1986. Attentions, intentions, and the structure of discourse. *Computational Linguistics*, 12: 175-204.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1995. Towards a computational theory of discourse interpretation. *Computational Linguistics*, 21/22: 203-25.

Horn, Laurence R. 1986. Presupposition, theme, and variations. *Chicago Linguistic Society*, 22: 168-192.

Hurewitz, Felicia. 1998. A quantitative look at discourse coherence. In Walker, Joshi, & Prince, eds.

Joshi, Aravind K., and Steve Kuhn. 1979. Centered Logic: the role of entity centered sentence representation in natural language inferencing. *Proceedings of the 6th International Joint Conference on Artificial Intelligence*, Tokyo, 435-439.

Joshi, Aravind K., and Scott Weinstein. 1981. Control of inference: role of some aspects of discourse structure—centering. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver, 385-387.

Kameyama, Megumi. 1986. Japanese zero pronominal binding: Where syntax and discourse meet. Paper presented at the 2nd System Development Foundation Workshop in Japanese Syntax, CSLI, Stanford, CA.

Miltsakaki, Eleni, and Karen Kukich. 2000. Automated evaluation of coherence in student essays. *Proceedings of the Workshop on Language Resources and Tools in Educational Applications*, LREC.

Rambow, Owen. 1993. Pragmatic aspects of scrambling and topicalization in German: A Centering approach. Paper presented at the Workshop on Centering Theory in Naturally-Occurring Discourse, IRCS, University of Pennsylvania.

Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*, 27: 53-94.

Sidner, Candace L. 1979. *Toward a Computational Theory of Definite Anaphora Comprehension in English*, Technical Report No. AI-TR537. Cambridge, MA: MIT Press.

Turan, Ümit Deniz. 1995. Null vs. Overt Subjects in Turkish Discourse: A Centering Analysis. Ph.D. Thesis, University of Pennsylvania.

Walker, Marilyn, Aravind K. Joshi, and Ellen F. Prince, eds. 1998. *Centering Theory in Discourse*. Oxford: Oxford University Press.