

Referring Expressions and the Local Coherence of Discourse in a Parallel Corpus of English and Japanese Map Task Dialogues

Etsuko Yoshida
Mie University/University of Edinburgh
1577 Kurimamachiya-cho, Tsu, Mie
514-8507 Japan
tantan@human.mie-u.ac.jp

Abstract

This paper examines how discourse entities are linked with the center (Cb) in the Centering Model in the construction of coherence in discourse, by looking at how the choice and distribution of referring expressions are correlated with the center transition patterns. Based on the analysis of parallel dialogue data in English and Japanese, the proportion of center transition patterns is investigated. More specifically, despite the grammatical differences in the form of references between the two languages, the ways of discourse development in both sets of data show distinctive similarities in the process in which topic entities are introduced, established, and shifted away to subsequent topic entities. There is little doubt that full NPs are the main conduit for the topic center, and continue to be used both within and across discourse segment boundaries, while the (zero) pronoun can carry the topic center only in limited stretches of discourse and is likely to be discontinued at the end of the discourse segment.

1. Outline

In this paper, I examine how discourse entities are linked with the center (Cb) in the Centering Model in the construction of coherence in discourse, by looking at how the choice and distribution of referring expressions are correlated with the center transition patterns. This may suggest that there is a universal feature in which noun phrases contribute to the center transition of Cb in developing the topic entities in dialogic discourse.

Centering is the model in which discourse

coherence is predicted by the transitions of discourse entities and the local focus of attention within the discourse segment (Grosz et al. 1995, Walker et al. 1998). I first analyse the distribution of center (Cb) transition patterns and the types of referring expressions in each center transition pattern. Then I clarify the preferred combination of transition sequence patterns by investigating the relationship between the types of referring expressions and the way of constructing topic chains in particular transition sequence patterns.

2. Example

Let us look at the following extract in CONTINUE transition in English. The chain of NPs tends to be combined with the chain of pronouns and zero pronouns: *this grass* is immediately replaced with PRO *it* and ZERO and established as a Cb in CONTINUE transition until the new entity *a house* is introduced. This new entity can predict the change of Cb in RETAIN transition, but the current Cb still continues as a noun *the grass*.

(1) (English: L1eq4c2)

TA109: And then you're turning up towards this grass CON

TB110: And going round it CON

TA111: Yeah round the left hand side [\emptyset] and then over the top [\emptyset] CON CON

TB112: Right

TA113: And then stop when you get to the/

TB114: Edge of it CON

TA115: Edge of it CON

TB116: Okay

TA117: And then do you have a house with on it?

RET

Just up from the grass on the right **CON**

TB118: No

In Japanese, it is possible to see that temporary chains of zero pronouns do occur, but the topic chains of NPs tend to link the entities in CONTINUE transition with these zero entities. Thus, the entities that are a current topic are consistent, but the types of expressions can be replaced depending on the different context of the situation in discourse.

3. Summary and Further Research

It has been clarified that despite the difference in grammatical forms of referring expressions between English and Japanese, there are similar patterns in the way the discourse entities are introduced, established as a focus of attention, and are shifted away to the new entities in discourse development. It can be clearly seen that the chain of NPs can contribute to discourse coherence in both English and Japanese, but their distribution shows that NPs have different roles in different types of Cb transition patterns. That is, NULL and SHIFT are the transitions in which the entities are introduced as a full NP and the expressions can be extended according to the understanding of the entities for the participants. In CONTINUE transitions the entities contribute to the topic chains of NPs, mainly as bare nouns in Japanese in the same way that pronouns contribute to topic chains in combining NPs in English data.

It is still difficult to evaluate the results as reliable, but the findings themselves are interesting. They suggest that the center of attention is maintained by the chain of NPs rather than (zero) pronouns, and that the chain of NPs is correlated with the local and global focus of discourse coherence. There is little doubt that full NPs are the main conduit for the topic center, and continue to be used both within and across discourse segment boundaries, while the (zero) pronoun can carry the center of topic only in limited stretches of discourse and is likely to be discontinued at the end of the

discourse segment.

These observations are not fully explained by existing theories of anaphora resolution and it is difficult to predict typical patterns of referential transitions in naturally occurring discourse. Two alternative perspectives appear to be particularly promising. First, it is worth noting that such repeated NPs tend to function as proper names in the discourse; second, the speaker's short-term memory is repeatedly activated by the combination of these expressions. Both phenomena call out for further research.

References

- Grosz, Barbara, Aravind Joshi, and Scott Weinstein. (1995) Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21/2, 203-225.
- Gundel, J. K., N. Hedberg and R. Zacharski (1993) "Cognitive status and the form of referring expressions in discourse." *Language*, 69, 2: 274-307.
- Poesio, Massimo, Rosemary Stevenson, Barbara Di Eugenio, and Janet Hitzeman. (2004) Centering: A parametric theory and its instantiations. *Computational Linguistics*, 30/3, 309-363.
- Prince, E. (1981) 'Toward a taxonomy of given-new information' In P. Cole (ed.), *Radical Pragmatics*. New York: Academic Press. 223-56.
- Walker M. A., A. K. Joshi and E. Prince (eds.) (1998) *Centering Theory in Discourse*. Oxford: Clarendon Press.
- Walker, M. A. (1998) "Centering, Anaphora Resolution, and Discourse Structure." In Walker M. A., A. K. Joshi and E. Prince (eds.).401-436.
- Walker, M.A. (2000) 'Toward a Model of the Interaction of Centering with Global Discourse Structure' *Verbum*.
- Yamura-Takei, Mitsuko. (2005) *Theoretical, Technological and Pedagogical Approaches to Zero-Arguments in Japanese Discourse: Making the Invisible Visible*. Doctoral thesis, Hiroshima City University.