Using picture sequences to study referential accessibility

Choonkyu Lee  
Utrecht Institute of Linguistics OTS

Abstract. Discourse researchers have studied various accessibility factors arising from linguistic context, and psycholinguists have recently begun to discuss the role of competition between potential referents in the visual scene even without explicit linguistic mention. Dimensions of discourse content such as story time and space, however, still have not been investigated thoroughly. In this paper, I discuss some methodological possibilities and considerations for investigating various accessibility factors including content dimensions, with a focus on wordless image stimuli. A narrator’s mental representation of the structure of story content—such as story time and space—is one of the determinants of the narrator’s use of different types of referring expressions in organizing narrative discourse effectively.

Keywords. Referential accessibility; Narrative discourse; Situational content; Image stimuli; Methodology

1. Introduction

When referring back to a character that has already been introduced earlier in narrative discourse, the speaker has several options in choosing the type of referring expression, including a repeated name (e.g., Billy), a definite description (the boy), and a pronoun (he). Researchers have noted that this referential choice by the speaker is sensitive to various factors of referential accessibility (e.g., Ariel 2001). Explicit references such as repeated proper names (as rigid designators) are associated with discourse entities that are low in accessibility in working memory, and impoverished forms such as pronouns (which only indicate grammatical gender and/or number) are associated with discourse entities that are high in accessibility, while definite descriptions, which have more explicit content (such as a role label) than pronouns but are at the same time anaphors usually requiring an explicit linguistic antecedent, lie in the middle of the accessibility hierarchy.

Many determinants of accessibility have to do with the nature of the preceding linguistic context. For example, referential distance (e.g., Givón 1992) is defined in terms of the number of intervening clauses between an antecedent and an anaphor; that is, the longer the intervening linguistic material, the lower the accessibility of the target discourse entity to be re-mentioned. Givón (1992) provided summary statistics based on existing text across languages to demonstrate the relationship.

Centering, defined in terms of the topicality of the preceding clause (e.g., Brennan 1995), is also associated with the likelihood of a pronominal anaphor. Brennan (1995) provided a review of comprehension studies showing a processing advantage for
pronouns referring back to the center, compared to other referential types, which incurred a processing cost. She also provided new production data to demonstrate the impact of the information structure in an immediately preceding clause on the referential choice for the anaphor in the next clause. (It is important to note that the target content in her production study was a videotaped basketball game, which is a continuous event throughout with no temporal gaps and only minor spatial shifts.)

Further, the presence of other discourse entities that were mentioned explicitly in preceding discourse creates referential competition (e.g., Arnold and Griffin 2007). In Arnold and Griffin’s (2007) storytelling task in which participants described a pair of scenes, participants were less likely to use a pronoun to refer back to a character in the second scene after having mentioned another character in the first scene, even though there was a gender mismatch between the two characters and thus no potential referential ambiguity with a pronoun.

Each of these accessibility factors that depend on the preceding linguistic context has received empirical support and seems to play a partial role in a narrator’s referential choice and a listener/reader’s coreference resolution. There are, however, factors that go beyond linguistic text, such as perceptual availability giving rise to referential competition (Fukumura et al. 2010); conceptual knowledge allowing bridging inferences without an explicit antecedent (Haviland and Clark 1974); and mental representations allowing ‘conceptual anaphors’ with mismatching grammatical features (Gernsbacher 1991). Furthermore, there is another class of factors to consider in referential choice, namely, situational dimensions of the discourse content, such as topic time, space, and overall theme, whose influence on referential choice has been understudied.

In Section 2, I review previous studies demonstrating the impact of content dimensions on a reader’s perception of story continuity as well as a narrator’s referential choice. In Section 3, I discuss some widely used narrative elicitation methods with regard to their advantages and disadvantages for investigating referential choice and information structure. I make concluding remarks in Section 4.

2. Narrative content dimensions

In narrative discourse, there are two levels of structural dimensions. One has to do with the text, i.e., the discourse topic time, space, and individuals that are indicated explicitly in the discourse. The other has to do with the target story content, with its own story time, space, and protagonists, among others. For example, there may be a long temporal gap or spatial shift from one scene/point in a story to the next, which a narrator may or may not choose to indicate as such in discourse structure. A particular narrator’s linguistic representation of story content is only a limited one, but given a reasonably faithful representation, the listener/reader can use the linguistic markers of discourse topic time, space, and individuals to build a situation model of the target content, namely, the story world (e.g., Grimes 1975, Vonk et al. 1992).

One of the earliest studies pointing to the importance of the structure of the situational content on a narrator’s referential choice is Clancy’s (1980) study of narrative discourse in English and Japanese. She used Chafe’s (1980a) “pear film” to elicit narratives, and described the impact of episodic boundaries on narrative organization.
(further discussed in Section 3). She found narrators’ tendency to mark shifts in topic space, perspective, or episode by nominal reference, rather than pronominal anaphora, in referring back to characters. Chafe (1980b) also noted that the cognitive difficulty in resetting the orientation of a narrative may depend on multiple factors of situational content – e.g., space, time, and people – although he did not link it directly to referential choice.

Besides global narrative coherence in terms of overall thematic continuity, the subdimensions of time, space, causality, and intentionality make independent contributions to a reader’s perception of story content continuity (Magliano et al. 1999). Magliano et al. (1999) reported online sentence fit judgments (‘How well does the sentence fit into the context of the story?’) and story reading speed data from story reading experiments which suggested that participants kept track of discontinuities along temporal, spatial, causal, and intentional dimensions separately in updating their situation model.

From the narrator’s perspective, the structure of the story content dimensions – e.g., how much time passed from one scene to the next in the target content – should be taken into account in organizing her/his narrative in order for the communication to be effective. Anderson et al. (1983) conducted a passage continuation experiment and found that the duration of a temporal gap in the story timeline as indicated by a temporal adverbial (ten minutes later / seven hours later) influenced the likelihood of mention of minor characters. In other words, for characters that are not of primary importance in the story plot, the narrator is likely not even to mention them when asked to continue the story after a long time shift in story time that goes beyond the typical duration of the topic event, whereas the same characters are more likely to be mentioned after a shorter interval in story time. Also in Natural Language Generation, McCoy and Strube (1999) tested their intuition that topic time in discourse serves as a discourse structuring device for application to automatic reference generation. In their reference generation algorithm for deciding between a pronoun and a definite description, they included a binary ‘time change’ parameter (whether there is a time shift or not in the discourse topic time) in addition to other factors such as referential distance and referential competition. Using references to persons in three New York Times articles as the gold standard, they found that using just the time change parameter led to 72.5% accuracy in producing the correct forms of referring expressions (where chance-level accuracy is 50% because they tested for two types of referring expressions, pronouns vs. definite descriptions). Comparing (a) a model which includes all their parameters except time change to (b) another model which includes all parameters as a minimal pair, they also observed performance improvement from (a) 78.5% to (b) 84.7%. In sum, representing a change in discourse topic time as a model parameter can help generate a naturalistic referring expression in re-mentioning a discourse entity.

Vonk et al. (1992) conducted an original production study on referential choice in relation to narrative content continuity. In a passage continuation task in Dutch in which participants had to write a sentence with a particular feeder word for a discourse referent, pronominal feeders led to more thematically continuous sentences than thematically discontinuous ones (as judged by independent judges with the referring expressions omitted). ‘Full NP’ feeders (proper names and definite descriptions) led to the opposite pattern, with more thematically discontinuous sentences. To complement the sentence
continuation data, Vonk et al. (1992) also conducted a self-paced reading experiment with a secondary probe recognition task, and found that readers were slower in recognizing a full NP anaphor in a probe (which was over-specific for the single protagonist in each text) than a pronominal anaphor. The authors argued, based on these findings, that the type of an anaphoric referring expression is tightly linked to a narrator’s subsequent narrative planning, and to a reader’s perception of thematic continuity. In addition to these production and processing experiments in which the type of a referring expression was determined by the researcher and presented to the participant, Vonk et al. (1992) conducted an elicitation task using cartoon strips with no verbal descriptions, which is reviewed in the next section.¹

3. Elicitation methodology

Various kinds of visual stimuli without verbal descriptions have been used in linguistic research to elicit narratives. Some of the pioneering studies used a silent film (e.g., the pear film, Chafe 1980a) or a wordless picture book (e.g., Mayer’s [1969] Frog story; see Bamberg 1985; Berman and Slobin 1994). Researchers have continued to use these stimuli extensively across many different cultures (e.g., Erbaugh 2001; Strömqvist and Verhoeven 2004; MacWhinney 2000) to study a wide range of topics, ranging from sociolinguistics to acquisition. Silent films and picture books with no verbal descriptions are excellent material for eliciting narratives to study discourse processes, including the dynamics of accessibility of discourse referents, because they have a coherent storyline and a global thematic structure, as well as varying degrees of continuity/discontinuity between scene transitions at a local level. Adult participants, in particular, readily treat the picture sequences as parts of a coherent whole, providing a narrative with transitional phrases (The next morning...) rather than a set of disconnected scene descriptions.

3.1. Silent films

In early work in this direction, Clancy (1980) used the ‘pear film’ to study narrators’ referential choice in 20 English and 20 Japanese narratives, in relation to the amount of intervening linguistic text and referential interference from other discourse referents between an antecedent and an anaphor. In addition to the impact of intervening material between an antecedent and an anaphor, Clancy (1980) noted ‘unusual’ shifts from inexplicit to explicit forms of reference within very short intervals that could not be explained by referential ambiguity (examples in Clancy [1980: 171-173]). She observed that these shifts to nominal references followed an ‘episode boundary’ where a slight thematic change occurs in the narrative, and argued that these shifts may be either a narrative device to signal narrative structure to the audience or a reflection of the narrator’s own mental representation, in which an episode boundary has deactivated discourse referents from the previous episode (see also Bestgen 1998).

¹ Chafe (1980b), Ariel (2001), and Landragin (2007) are excellent reviews of additional factors in information structure and accessibility including phonology, perception, speech/dialog processes, and the amount of descriptive content in the referring expression, although they make little or no mention of the impact of content structure on referential choice.
Clancy (1980) also made some important observations that have methodological implications, in her discussion of character introductions in the opening scene. She noted that a few Japanese narrators used zero anaphora from the very beginning in introducing the pear man in the film (or, "simply plunged into the narrative without actually introducing the pear man" [Clancy 1980: 145]) based on their assumption that the interviewer must already be familiar with the story. Moreover, Clancy (1980) described narrators’ shifts in perspective between the story world and the ‘real world,’ in which narrators made evaluative remarks about the content or quality of the movie as reflections to themselves or to the interviewer. What this suggests is that, for any investigation of a linguistic phenomenon that is sensitive to the common ground between interlocutors, the physical presence of an experienced researcher may have undesirable consequences and should be considered carefully, even if the task is a largely one-sided task with little verbal communication. Unless the goal of the study is specifically to address this mode of narration, with shifts in the target audience or perspective and in the presence of an experienced audience, unnecessary personal monitoring can be avoided in the task design. For dialog studies in which personal interaction is a prerequisite, it makes sense to have at least a confederate or another participant as a partner, and for monologic narrative tasks, it is reasonable to use a confederate as the audience or just remove the physical audience component altogether by letting the narrator imagine an audience. These alternatives address the need of making the narrative task pragmatically reasonable for the research objectives. If the task design in Clancy (1980) is adopted, however, it is important to keep in mind that discourse-given/new status is often established separately in different worlds or perspectives (Clancy 1980).

Another important insight in Clancy’s (1980) discussion as well as other accounts of accessibility (e.g., Ariel 1990; Gundel et al. 1993) is that discourse-given/new status is gradient and dynamically changing, with old information becoming newer over time under working memory constraints and with referential competition from other discourse referents. Clancy (1980) proposed that more than a single mention may be required to introduce a character and establish it as old information, especially in languages like Japanese where zero anaphora is a major option for inexplicit reference but may be too inexplicit early on for a newly introduced character.2

3.2. Picture books and strip cartoons

Vonk et al. (1992) studied the impact of thematic continuity in a story on referential choice, using wordless comic strips to create experimental conditions of thematic continuity vs. thematic shift without verbal intervention. There were two versions of six strip cartoons, each with three or four pictures. One version had thematic continuity, and the other had a thematic shift in the final picture (as confirmed in pre-test rating data). The authors asked 30 native speakers of Dutch to write a coherent story for the pictures, and found a tendency (approaching significance) toward pronouns in thematic continuity versions and more balanced proportions of nominal references and pronouns after a

---

2 See Clancy (1980) for discussion of important crosslinguistic findings in English and Japanese. For example, English pronouns, as the main option for ‘inexplicit reference,’ pattern more similarly to Japanese zero anaphora than English ellipses do.
thematic shift. When two independent judges read these written narratives and judged the thematic continuity/shift at the critical sentence, those with a pronoun in the critical sentence led to significantly higher responses of ‘thematic continuity’ (81%) compared to those with an explicit nominal reference (19%) (Table 6. Vonk et al. 1992: 314).

As mentioned above, Mayer’s (1969) Frog story and other wordless picture books in the series are also excellent material for studying referential choice in narrative discourse (see Berman and Slobin 1994; Strömqvist and Verhoeven 2004). Although there is little discussion of referential choice in Berman and Slobin’s (1994) volume, there is more discussion in Strömqvist and Verhoeven’s (2004). In particular, Hickmann (2004) discussed the relationship between content structure and information structure as reflected in the use of discourse markers and accessibility markers from a developmental perspective. We have also used Mayer’s wordless picture books in our recent work on the impact of content dimensions on referential choice (e.g., Lee 2012). Some of the most useful characteristics of Mayer’s picture books for studying the impact of content dimensions include the wide range of variation in content dimensions (e.g., short vs. long temporal intervals, and smooth vs. abrupt spatial shifts between consecutive scenes in the story) and the presence of multiple characters in most of the scenes, including a boy, his family, and his somewhat anthropomorphic pets, which often elicit personal pronouns in narratives. While these picture books may look too childish for adult participants, it is important to note that it is the target audience which is expected to be young children, and not necessarily the narrators (see Marchman 1989 for a cross-sectional developmental comparison including an adult group). In fact, this is an additional reason why it is undesirable to have an experienced researcher with the narrator during in an elicitation session. Ideally, it would be more naturalistic to have a child audience during the session, or one can have the narrator produce her/his narrative with a target audience of young children in mind (Lee 2012).

3.3. Other materials

The decisions regarding elicitation methodology would, of course, depend on the research questions one wants to address. Although some researchers are only interested in dialog processes, discourse without real-time personal interaction constitutes a large part of our linguistic activities as well. For discourse, some are more interested in written text, others in oral narrative, and others in cross-modal comparison. For example, in a study of speech processes related to the discourse-given/new distinction, Bard et al. (2000) analyzed dialog during a collaborative navigation task with maps containing differing landmarks between the interlocutors (see also Wilkins 1993).

Further useful visual stimuli for elicitation can be found on the Story-builder website (Sardinha 2011) and the L&C Field Manuals and Stimulus Materials website (login required; e.g., Kita’s [1995] animation).

---

3 Arnold and Griffin (2007), Fukumura et al. (2010), and Serratrice (2013) also used picture sequences of cartoon characters and toy figures to study the production of referring expressions in a similar passage continuation task, although the independent variables of interest in these studies were referential competition, grammatical number, or animacy, rather than continuity in content dimensions.
4. Conclusion

In this paper, I have reviewed some elicitation studies with only nonverbal stimuli that address narrators' referential choice, and pointed out some important findings, caveats for data interpretation, and methodological considerations.

In investigating the impact of a single content dimension separately from other possible confounds (e.g., teasing story time apart from space, protagonist, etc.), it would be ideal to manipulate each dimension while holding all others constant for experimental control. This would necessitate customized illustrations rather than naturally existing picture books, but this is reasonable for the story time dimension, which can be controlled more precisely with explicit visual signals of time such as a clock. For the dimension of space, however, it is physically impossible to manipulate just space while holding all other dimensions constant, so there are limitations to this alternative.

References


