Cognitive Psychology 65 (2012) 1–38

Contents lists available at SciVerse ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

(Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension

Neil Cohn [a,*], Martin Paczynski [a], Ray Jackendoff [b], Phillip J. Holcomb [a], Gina R. Kuperberg [a,c]

[a] Department of Psychology, Tufts University, Medford, MA 02155, United States
[b] Center for Cognitive Studies, Miner Hall, Tufts University, Medford, MA 02155, United States
[c] Department of Psychiatry, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Bldg 149, 13th Street, Charlestown, MA 02129, United States

ARTICLE INFO

Article history:
Accepted 28 January 2012
Available online 2 March 2012

Keywords:
Coherence
Comics
Discourse
ERPs
Event-related potentials
Film
Visual language
LAN
N300
N400

ABSTRACT

Just as syntax differentiates coherent sentences from scrambled word strings, the comprehension of sequential images must also use a cognitive system to distinguish coherent narrative sequences from random strings of images. We conducted experiments analogous to two classic studies of language processing to examine the contributions of narrative structure and semantic relatedness to processing sequential images. We compared four types of comic strips: (1) Normal sequences with both structure and meaning, (2) Semantic Only sequences (in which the panels were related to a common semantic theme, but had no narrative structure), (3) Structural Only sequences (narrative structure but no semantic relatedness), and (4) Scrambled sequences of randomly-ordered panels. In Experiment 1, participants monitored for target panels in sequences presented panel-by-panel. Reaction times were slowest to panels in Scrambled sequences, intermediate in both Structural Only and Semantic Only sequences, and fastest in Normal sequences. This suggests that both semantic relatedness and narrative structure offer advantages to processing. Experiment 2 measured ERPs to all panels across the whole sequence. The N300/N400 was largest to panels in both the Scrambled and Structural Only sequences, intermediate in Semantic Only sequences and smallest in the Normal sequences. This implies that a combination of narrative structure and semantic relatedness can facilitate semantic processing of upcoming panels (as reflected by the N300/N400). Also, panels in

* Corresponding author. Address: Department of Psychology, Tufts University, 490 Boston Ave., Medford, MA 02155, United States. Fax: +1 617 812 4799.
E-mail address: neilcohn@emaki.net (N. Cohn).

0010-0285/$ - see front matter
doi:10.1016/j.cogpsych.2012.01.003

the Scrambled sequences evoked a larger left-lateralized anterior negativity than panels in the Structural Only sequences. This localized effect was distinct from the N300/N400, and appeared despite the fact that these two sequence types were matched on local semantic relatedness between individual panels. These findings suggest that sequential image comprehension uses a narrative structure that may be independent of semantic relatedness. Altogether, we argue that the comprehension of visual narrative is guided by an interaction between structure and meaning.

1. General introduction

Drawings have been conveying narratives through sequences of images for millennia, whether painted on cave walls, carved into reliefs, hung on medieval tapestries, or, in their modern context, appearing in comic books (Kunzle, 1973; McCloud, 1993). Compared to research on the comprehension of verbal narrative, however, few studies have examined the driving forces behind our understanding of visual narrative: what are the representations and mechanisms engaged during sequential image comprehension? And, how is meaning integrated with structure across a sequence of images?

This research addresses some of these questions by examining online processing of sequential images using both reaction time (RT) measures and event-related potentials (ERPs). In this section, we first provide an overview of experimental and theoretical approaches to the comprehension of text/discourse and sequential images, showing how principles that inform research on text and discourse can also be applied to the comprehension of sequential images. We begin by discussing coherence relationships across individual sentences and individual images. We then progress to studies that extend beyond such linear relationships, examining how global narrative structure is built across text/discourse and sequential images. After this, we consider attempts to describe such a narrative structure at a theoretical level. Finally, we highlight important gaps in the existing literature, setting the stage for the current studies.

1.1. Linear coherence relationships in language and sequential images

1.1.1. Text and discourse

In early work on text and discourse, researchers emphasized how related concepts were often drawn together into common semantic fields (Bransford & Johnson, 1972; Brown & Yule, 1983; Halliday & Hasan, 1985; van Dijk, 1977) or scripts and schemas (Schank & Abelson, 1977). Structure was thought to be imposed on these general semantic fields primarily through coherence relationships between individual sentences (Halliday & Hasan, 1976; Hobbs, 1985; Kehler, 2002; Mann & Thompson, 1987).

Over the past 40 years, multiple types of coherence relationships have been described across dimensions of reference (Haviland & Clark, 1974), temporal and event structure (Mandler, 1986; Speer & Zacks, 2005; Zwaan, 1996), space (Clark, 1972; Linde & Labov, 1975; Morrow, Greenspan, & Bower, 1987), intentionality (Lichtenstein & Brewer, 1980; Schank & Abelson, 1977), and causation (Black & Bower, 1980; Keenan, Baillet, & Brown, 1984; Trabasso, Secco, & van den Broek, 1984). Indeed, a large body of psycholinguistic work now indicates that such relationships are constructed during online discourse and text comprehension to build up a "situation model"—a mental representation of the situation described in a discourse (see Zwaan & Radvansky, 1998 for a review).

1.1.2. Sequential images

Just as for verbal discourse, most work on sequential images has focused on linear relationships between individual "panels" or frames. Comic author and theorist Scott McCloud (1993) proposed a popular model of six types of linear "transitions" between panels based on changes between characters,