

Louwerse, M.M. & Graesser, A.C. (in press). Macrostructures. In
E.K. Brown (Ed.) *The encyclopedia of language and linguistics*.
Amsterdam: Elsevier

Macrostructure

Max M. Louwerse and Arthur C. Graesser

University of Memphis

Address for correspondence:

Max M. Louwerse

Department of Psychology /

Institute for Intelligent Systems

University of Memphis

202 Psychology Building

Memphis, TN 38152

Phone: (901) 678-2143

Fax: (901) 678-2579

Email: mlouwers@memphis.edu

Max M. Louwerse

Department of Psychology /

Institute for Intelligent Systems

University of Memphis

202 Psychology Building

Memphis, TN 38152

Email: mlouwers@memphis.edu

Arthur C. Graesser

Department of Psychology

University of Memphis

202 Psychology Building

Memphis, TN 38152

Email: a-graesser@memphis.edu

Abstract

Macrostructures are global textual structures that form the global meaning of the text. They are created by applying macrorules (deletion, generalization and construction) to a sequence of propositions derived from the text. The result is a set of macropropositions that form the gist of the text. Empirical evidence for macrostructures comes from summarization and recall of text. Macrostructures were introduced by Van Dijk (1972, 1977, 1980) and further developed into psychological notions of discourse comprehension by Kintsch and Van Dijk (1978) and Van Dijk and Kintsch (1983). These theories have been refined and computationally implemented in the Construction-Integration model (Kintsch, 1988, 1998).

Macrostructure

Macrostructures are structures that organize texts globally, just as microstructures that organize locally. Given that texts are not just concatenations of sentences, texts need to be structured both locally (connections between clauses and sentences) and globally (larger fragments of discourse, e.g. paragraphs). Syntactic rules, the meaning of the words of the sentence and general heuristics of discourse form the microstructure of a text. These microstructures organize sets of interrelated propositional representations of the phrases, clauses and sentences of the text. Macrorules translate these sequences of propositions into a smaller set of more general propositions by deleting propositions that are less important for the overall meaning of the text, by generalizing propositions into supersets and by constructing new text units that replace the meaning of the old set. Macrorules operate recursively, so that macrostructures that are formed by macrorules may be subject to another cycle of macrorules, further generalizing the gist of the text. Macrostructures are therefore abstract semantic descriptions of the semantic content of the text, similar to the text's global meaning and theme and providing global coherence.

The term macrostructure for global principles of text organization was first proposed by Bierwisch in 1965 for narrative structures in literary texts. In 1968 Harris discussed a similar idea of global text structuring. In that same year the *Morphology of the folktale* by Vladimir Propp was translated from Russian (original 1928). Propp argued that Russian fairy tales share a particular narrative structure. Around that same time narratologists like Greimas (1966), Bremond (1964), Labov & Waletzky (1967), Lévi-Strauss (1960) and Todorov (1971) proposed similar narrative grammars. According to these grammars, stories are like sentences in that their narrative structures are structures analogous to

syntactic structures. The development of the concept of macrostructures should be seen against the background of the developments of narrative structures.

Macrostructures became an established term in text linguistics after being further developed by Van Dijk (1972). Van Dijk's text linguistic approach was very much based upon theoretical linguistics and Chomsky's (1957, 1965) generative-transformational grammar. Chomsky (a student of Harris) argued that sentences have a recursive capacity. Each sentence has a deep structure that is interpreted by the semantic component of the grammar. Syntactic transformations relate the deep structure to the surface structure of the sentence. Some of the elementary transformations consist of adjoining, moving, deleting, copying constituents. Van Dijk argued that text grammars also have these deep and surface structures. The equivalent to the sentential surface structures are microstructures; the equivalent to the sentential deep structures are macrostructures. As with the sentential surface structures, microstructures have underlying rules to represent the underlying semantic representation of the sentences. As with the sentential deep structures macrostructures have an abstract semantic character and are specified by macro-semantic rules operating on the microstructures. Although the concept of macrostructures remained the same, the direct link to sentence structures faded in later work (Van Dijk, 1977, 1980).

Around the time of Van Dijk's introduction of text grammars in (text) linguistics and poetics, Kintsch (1974) argued that cognitive psychology should not focus on isolated sentences only, but has to focus on texts. Kintsch proposed that the representation of texts in memory is a network of interrelated propositions. These propositions are units

of meaning roughly corresponding to phrases or clauses. A proposition consists of a predicate that modifies one or more arguments. A sentence like “The teacher explained the concept to the students” can be represented propositionally as (EXPLAIN, TEACHER, CONCEPT, STUDENT). It contains one predicate (EXPLAIN) and three arguments (TEACHER, CONCEPT, STUDENT). Arguments are generally the nouns of a clause, but can also be prepositional phrases and even references to other propositions. For instance, the second and third propositions in a text conjoined by the connective ‘or’ can be represented as (OR, 2, 3).

Kintsch’s psychological ideas on texts and representations of meaning in memory and Van Dijk’s linguistic ideas on text grammars resulted in an impetus in research on discourse comprehension, first resulting in Kintsch and Van Dijk’s (1978) influential model of text comprehension. The goal of the model is to explain coherence in text. According to the model the aim of text comprehension is the formation of micro- and macrostructures. In the comprehension process meaningful text units are transformed into propositions. Take for instance the following extract of a text used in Kintsch and Van Dijk (1978) and its propositional representation.

A series of violent bloody encounters between police and Black Panther Party members punctuated the early summer days of 1969.

1 (SERIES, ENCOUNTER)

2 (VIOLENT, ENCOUNTER)

3 (BLOODY, ENCOUNTER)

4 (BETWEEN, ENCOUNTER, POLICE, BLACK PANTHER)

5 (TIME: IN, ENCOUNTER SUMMER)

6 (EARLY, SUMMER)

7 (TIME: IN SUMMER, 1969)

Coherence is achieved by an overlap of the arguments in these propositions. The comprehender links the arguments of propositions in order to form a coherence graph. In the graph below, proposition 4 is linked to proposition 1, 2, 3 and 5 by the argument ENCOUNTER. Similarly, propositions 6 and 7 are linked to 5 by the argument SUMMER. This graph (Figure 1) is the representations of the microstructure of the text.

FIGURE 1 ABOUT HERE

However, there are working memory constraints to linking propositions. According to the model only a limited number of propositions can stay in one processing cycle. The processing cycle is a temporal period during which propositions are linked. Propositions that stay in multiple processing cycles, and stay longer in working memory, are more memorable. Propositions enter working memory when they are selected by a leading-edge strategy: recent propositions and/or those higher in the hierarchy. Imagine that the processing cycle could hold only four propositions. In this example, according to the leading-edge strategy propositions 4, 5, 7 and 3 would move on to the next processing cycle.

The macrorules deletion (those propositions whose deletions do not change the meaning of the text are deleted); generalization (sequences of propositions that can be replaced by supersets are generalized); and construction (sequences of propositions that can be replaced by a single proposition are constructed into a single proposition) are carried out on the microstructure to obtain the macrostructure. The above extract is too short to provide a complete analysis, but one can imagine the following macrorules being applied to the microstructure:

Deletion:

(BLOODY, ENCOUNTER) & (BETWEEN, ENCOUNTER, POLICE, BLACK PANTHER) → (BETWEEN, ENCOUNTER, POLICE, BLACK PANTHER)

Generalization:

(TIME: IN SUMMER, 1969) → (IN, EPISODE, SIXTIES)

Construction:

(BLOODY, ENCOUNTER) & (BETWEEN, ENCOUNTER, POLICE, BLACK PANTHER) → (FIGHT, POLICE, POLITICAL MOVEMENT)

The long-term recall of the text is based on the propositions in the macrostructure of the text. In the above example a macroproposition like (FIGHT, POLICE, POLITICAL MOVEMENT) is a good approximation of the gist of the text, could be used as the title for the text and is likely to be remembered best.

Microstructures and macrostructures as discussed so far can be seen as rules and structures of text grammars. Van Dijk and Kintsch (1983) added to these abstract semantic representations of the text the flexibility and fallibility of the user, thereby moving away from the text itself to knowledge/text interaction. One of the problems with the Kintsch and Van Dijk (1978) model and related theories is that text representation is entirely based on the text. The role of the reader's knowledge and the strategies the reader uses in the comprehension process are limited. In addition to a multileveled propositional textbase, Van Dijk and Kintsch argue for a situational model that incorporates the reader's interpretations – both correct and incorrect – of the text. The situational model can for instance explain individual differences in interpretations of texts, differences in the meaning of translations, and the fallibility and flexibility of memory. It also allows for the text being grounded. In the formation of macrostructures sociocultural contexts (for instance, cultural information regarding the situation and the type of interaction) play an important role. In addition, to the textual macrostructures Van Dijk and Kintsch (1983) therefore proposed pragmatic superstructures, schematic structures similar to the rhetorical structure of the text. An example of such a superstructure is a narrative schema or an argumentative schema. Knowing the superstructure of a text allows the reader to anticipate the global organization (macrostructure) of the text.

Kintsch's construction-integration model (1988, 1998) continued and extended the Kintsch and Van Dijk (1978) and Van Dijk and Kintsch (1983) process model. In the CI model the process from text to mental representation consists of two phases. In the construction phase an approximate mental model is locally constructed based on the text

and the reader's background knowledge. In the integration phase this tentative mental model stabilizes by filtering out irrelevant and redundant information. An activation process spreads around the network of propositions, boosting strong links between propositions and dampening weak links in order to get a well-structured mental representation. The representation resulting from the Integration process is then stored in the long-term memory. As with the Kintsch and Van Dijk (1978) and Van Dijk and Kintsch (1983) models, the distinction between microstructure and macrostructure is orthogonal to the distinction between textbase and situation model. The text yields a textbase. The textbase is integrated with the background knowledge of the reader and yields the situation model. This means that two global features characterize the mental representation of a text: its macrostructure and the situational and world knowledge.

Various studies have shown that the macrostructure can significantly predict recall and summarization results. Those propositions that end up in the macrostructure are remembered better than propositions in the microstructure (Graesser, 1981; Kintsch and Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon & Keenan, 1975; McKoon and Ratcliff, 1980; Meyer, 1975).

Because macrostructures are higher level properties of sequences of propositions, it is hard to identify linguistic manifestations of macropropositions and their macrostructures at the surface level of the text. However, some indicators of macropropositions can be identified. Titles of texts, summaries and topical sentences, often at the beginning or end of a paragraph generally indicate macropropositions. Furthermore, given the properties of coherence graphs, clauses that tend to be conjoined

by connectives are good candidates for macropropositions, as are sentences that contain demonstratives and pronouns.

Propositional macrostructures cannot be computed automatically from the text. One of the reasons for this is that the macrorules that can be applied to the microstructure remain underspecified. Kintsch (1998, 2002) has however shown how macrostructures can be derived from the text computationally using latent semantic analysis. The meaning of a sentence is represented as a vector in a high-dimensional semantic space. Those vectors that relate most to the overall text (and have the highest typicality scores) can be identified as macropropositions.

References

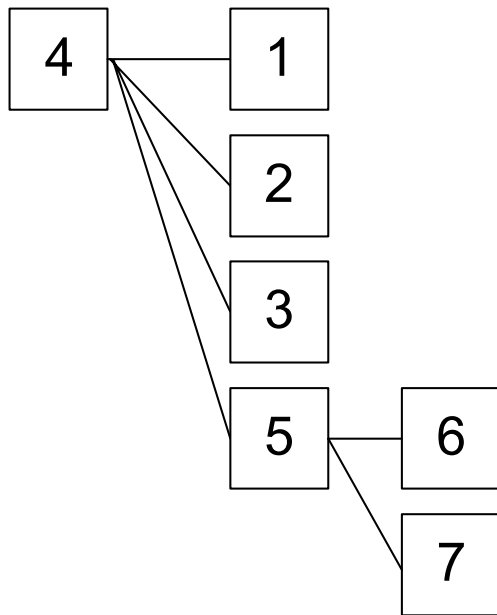
- Bierwisch, M. (1965). 'Poetik und Linguistik'. In H. Kreuzer & R. Gunzenhäuser (Eds.), *Mathematik und Dichtung* (pp. 46-66). Munich: Nymphenburger.
- Bremond, C. (1964). 'Le message narratif'. *Communications*, 4, 4-32.
- Chomsky, N. (1957). *Syntactic structures*. The Hague, Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Graesser, A.C. (1981). *Prose comprehension beyond the word*. New York, Springer Verlag.
- Greimas, A. (1966). *Sémantique structurale: Recherches de méthode*. Paris: Larousse.
- Harris, Z. S. (1968). *Mathematical structures of language. Interscience tracts in pure and applied mathematics*, 21. New York: John Wiley & Sons.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1988). 'The role of knowledge in discourse comprehension: a construction-integration model'. *Psychological Review*, 95, 163-182.

- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W. (2002). 'On the notions of theme and topic in psychological process models of text comprehension'. In M. Louwerse & W. van Peer (Eds.) *Thematics: Interdisciplinary studies* (pp. 157-170). Amsterdam: Benjamins.
- Kintsch, W., Kozminsky, E., Streby, W., McKoon, G., & Keenan, J. M. (1975). 'Comprehension and recall of text as a function of content variables'. *Journal of Verbal Learning and Verbal Behavior*, 14, 196-214.
- Kintsch, W., & Keenan, J. M. (1973). 'Reading rate and retention as a function of the number of propositions in the base structure of sentences'. *Cognitive Psychology*, 5, 257-274.
- Kintsch, W. & Van Dijk, T.A. (1978). 'Toward a model of text comprehension and production'. *Psychological Review* 85, 363-394.
- Labov, W. & Waletzky, J. (1967). 'Narrative analysis: Oral versions of personal experience'. In J. Helm (ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle: University of Washington Press.
- Lévi-Strauss, C. (1960). 'La structure et la forme'. *Cahiers de l'Institut de Science Économique Appliquée*, 99, 3-36.
- McKoon, G., & Ratcliff, R. (1980). 'Priming in item recognition: The organization of propositions in memory for text'. *Journal of Verbal Learning and Verbal Behavior*, 19, 369-386.
- Meyer, B.J.F. (1975). *Organization of prose and its effect on memory*. Amsterdam: North Holland.

- Propp, V. (1968). *Morphology of the folktale*. (original work published 1928). Austin: University of Texas Press.
- Todorov, T. (1971). *Poétique de la prose*. Paris: Seuil.
- Van Dijk, T.A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Dijk, T.A. (1972). *Some aspects of text grammars. A study in theoretical linguistics and poetics*. The Hague: Mouton.
- Van Dijk, T.A. (1977). *Text and context. explorations in the semantics and pragmatics of discourse*. London: Longman.
- Van Dijk, T.A. (1980). *Macrostructures*. Hillsdale, NJ: Erlbaum.

Figure Caption

Figure 1. Example of coherence graph of extract of text (Kintsch & Van Dijk, 1978).



Keywords

Macrostructure, microstructure, Van Dijk, Kintsch, text comprehension, discourse processes, construction-integration model, text linguistics, theme, proposition, text grammar, narrative structure, text analysis, stylistics, macrorules

Cross-references

thematics, reading processes in adults, psycholinguistics: overview, computers and text, proposition, text linguistics, text/rhetoric, text analysis, narrative: linguistic and structural theories

Biography



Max M. Louwerse is presently an assistant professor in the Department of Psychology and the Institute for Intelligent Systems at the University of Memphis. Dr. Louwerse received a Ph.D. degree in Linguistics from the University of Edinburgh, Scotland. His research interests are in interclausal relationships, discourse markers, questions answering systems, mixed initiative dialog, themes, narrative structure and various other aspects of discourse processing. He is currently involved in research projects on developing coherence measurements in texts, intelligent tutoring systems, question answering systems, automated speech recognition and automated dialog management. He translated and introduced Propp's *Morphology of the Folktale* in Dutch and edited together with Willie Van Peer the interdisciplinary volume *Thematics. Interdisciplinary studies* (2002).



Art Graesser is presently a full professor at the University of Memphis in the Departments of Psychology and Computer Science, serving as a co-director of the Institute for Intelligent Systems and director of the Center for Applied Psychological Research. Dr. Graesser received his Ph.D. in 1977 in psychology from the University of California at San Diego. His research interests are in text comprehension, inference generation, conversation, reading, knowledge representation, question asking and answering, tutoring, intelligent tutoring systems (AutoTutor), education, computational discourse, and human-computer interaction. He has served as an editor of the journal *Discourse Processes* for 10 years, is one of the founders of the Society for Text & Discourse, and is an editor of the 2003 *Handbook of Discourse Processes* (2003).