| Title | Discourse and paragraph : Visions and revisions |
|---|---|
| Sub Title | 談話とパラグラフ：談話構造における写像の問題 |
| Author | 大堀, 俊夫(Ohori, Toshio)<br>高橋, 悦子(Takahashi, Etsuko)<br>山田, あき(Yamada, Aki)<br>柳谷, 啓子(Yanagiya, Keiko) |
| Publisher | 慶應義塾大学藝文学会 |
| Publication year | 1986 |
| Jtitle | 藝文研究 (The geibun-kenkyu : journal of arts and letters). Vol.48, (1986. 3) ,p.28- 15 |
| JaLC DOI | |
| Abstract | |
| Notes | |
| Genre | Journal Article |
| URL | https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00072643-00480001-0231 |

# Discourse and Paragraph—
# Visions and Revisions

Toshio Ohori
Etsuko Takahashi
Aki Yamada
Keiko Yanagiya

0.　The present paper attempts (1) to clarify some of the problems involved in current theories of discourse analysis, and (2) to re-interpret the concept of " paragraph " as a necessary prerequisite for an inquiry into certain functional domains of discourse processes, namely those dealing with the mapping problem of discourse production / comprehension.

What follows is thus primarily a formulation of problems, accompanied by a survey of previous studies on the paragraph in such disciplines as composition, anthropology, psychology, and linguistics. These discussions shall be the preliminary to our subsequent research.

1.　In the study of discourse structure, several levels of analysis are generally recognized.　According to Kintsch (1985: 231), they are " the linguistic parsing of the text, followed by the construction of atomic propositions that represent its meaning elements; next, these n,eaning elements are organized into a coherent textbase, which represents the full meaning of the text; from this textbase, the macrostructure of the text is derived, representing its essence or gist."　Such are the " representations of the text proper ", and this process " results in the construction of a situation model, which is not a representation of the text itself but of the situation referred to by the text." (ibid.) Among these—and doubtless many other alternatives—we are perhaps

justified in isolating two levels as common denominators of discourse analysis. One is micro-structure, the actualization of lexico-grammatical subsystems which usually take the form of a series of sentences; the other is macro-structure, what Kintsch calls the " gist " of discourse (for the earliest definition of these terms, cf. van Dijk, 1972: 5-6; the recent version by the same author can be found in van Dijk, 1985: 115).

Here, what we shall call the *mapping problem* of discourse analysis arises. Informally it is defined as the issues dealing with the interrelation between micro- and macro-structures, i.e. in which way are both structures relevant to each other? A host of models explicitly or implicitly capture some aspects of this problem, but now let us look at it in more detail. Our formulation of the mapping problem is as follows:

(1) What levels should be postulated between micro- and macrostructures?

(2) To what extent do such levels have autonomy? Are they on a continuous scale (e.g. mediated by successive chunking) and hence only defined relative to each other, or do they conform to distinct organizing principles?

(3) As a specifically linguistic problem, in which way does microstructure by itself contribute to other upper level processes?

For any theory of discourse analysis to be adequate, we require that it answer the questions formulated above. Among them, from our perspective, the third is of particular importance. The scheme of Kintsch quoted above is one possible answer to the questions (1) and (2), presumably representing models developed in cognitive psychology and artificial intelligence. We admit that this way of modeling is most comprehensive and explicit in their conceptual design. But in view of the question (3), it seems to underestimate the importance of linguistic structures. In Kintsch (1985) and other related works at least, the output of the parsing component is text-plus-annotations, out of which propositions are constructed. These annotations include, as far as we see from his examples, only the clues necessary for proposition con-

struction to the exclusion of functional notions. Yet what if the micro-structure provides the clues for the addresser's control over the information flow of discourse? This is not a groundless suggestion. Here, we refer to one notion that seems lacking adequate treatment in cognitively oriented models, namely topicality. In the parsing of English sentences, this notion would not require special treatment in many cases, since English is a subject-prominent language and topics are usually identifiable with subjects. But consider the cases of such topic-prominent languages as Chinese, or topic- and subject-prominent languages such as Japanese. If the parsing component, mainly designed for English and other typo-logically similar languages, is to deal with such cases, then its design must change drastically.

Another example is from English, borrowed from Bolinger (1977: 11). There, the following pair is discussed:

(4) The forecast says it's going to rain.
(5) The forecast says that it's going to rain.

Of these, if one steps into a room and wants to drop a casual remark about the weather, (4) is all right, but (5) is rather odd. Bolinger explains this as " if we look at situations where speakers are volunteering information, where no question has been asked and no answer is implied, but what is being said comes out of the blue, it is unnatural for the word *that* to be used." (ibid.) If this observation is correct, we must deliber-ately incorporate such a fact (in this case the presence of *that* as an indicator of the addresser's unconscious control over the status of the information in the nominal clause) into the model of discourse processes. Our argument does not invalidate the cognitive models, but they should be supplemented by the subtler annotation of functional properties of micro-structure.

Now, our hypothesis is that there is at least one intermediate level up to which micro-structure is relevant. In other words, we can char-acterize this level in terms of the formal cues available from micro-structure. Further, we assume that the unit identified on this level is relevant in the macro-processes. This leads to a restricted version of

the mapping problem, which can be reformulated in the following way:

(6) What are the organizing principles of this level, and what is the nature of the entity acting as a unit of this level?

(7) In which way does micro-structure contribute to this medial level process?

By taking such a strategy, we do not lose the generality attained in previous models, and at the same time take the characteristics of each particular language sufficiently into account. We consider it sets the discourse analysis on the firmer ground and specifies the proper concern of linguistic theory.

The " unit " of this intermediate level, we consider, can be identified with *paragraph* with certain modification. A new term could have been coined, but we followed Longacre (1979) in adopting this relatively established term since we expect the studies devoted to the paragraph will provide analytic concepts and insights that may be useful to our research.

Before doing the survey of previous studies, a few possible counter-arguments should be answered, concerning the indeterminancy and incoherence of the definition of paragraph. Above all, it is essential to distinguish two concepts of paragraph, the *orthographic paragraph* (henceforth o-paragraph) and the *semantic paragraph* (s-paragraph). The former is the common-sense conception of paragraph, delimited by typographic conventions. The latter concerns the chunk of information as conceived in psychology (hence an alternative term would be *cognitive paragraph*). This can be analyzed regardless of the medium (though features that count are different), so some even adopt the term *paratone* for spoken discourse.

For o-paragraph, all that linguistic analysis can do is to leave it out. As Herbert Read saw half a century ago, o-paragraph is largely a matter of prose rhythm, and as far as it cannot be defined with scientific rigor, we do exclude it from the present consideration. On the other hand, concerning the s-paragraph, several monographs (largely independent of one another) have been published. One such study is Koen, Becker,

and Young (1969). Their experiment provided subjects with both normal English and nonsense prose materials, and the subjects apparently agreed with one another in their judgments of paragraph boundaries in both cases. This suggests that the division of discourse into units is psychologically real, and further, such units (i.e. s-paragraphs) are largely dependent on the formal features of discourse. Another, more recent series of studies, Sakuma (1979, 1984) also supports the postulation of such a unit. There also, the judgments of the subjects (313 for English, and 209 for Japanese) did not agree with the *original* indentations when they divided the non-indented materials. Nevertheless, both English and Japanese subjects agreed *among each other* to a considerable extent. All these results establish the o-paragraph and s-paragraph as distinct entities, whose agreement is but one possibility among many. The object of our investigation is thus restricted to the latter.

2. In this section, we shall survey studies on the paragraph in their specifics. Remarks from more general standpoints will be found in the next section.

2. 1 *Composition.* We should note that works on composition are primarily prescriptive. Their concern is—as the paragraph is always already visually identifiable by orthography—the prescription of strategies to facilitate the processing of discourse, i.e. how one should realize his articulation of thought (as related to some aspects of the s-paragraph) by means of o-paragraphs. From this comes the common-sense definition of the o-paragraph as " a development of a single idea."

Since our standpoint is descriptive and not prescriptive, we shall merely mention some notions developed in Christensen (1967), a seminal essay in this field. In this essay, following the 19th century logician Alexander Bain, Christensen redefines topic sentence as analogous to the " base clause " of a complex sentence. " It is the sentence on which others depend. It is the sentence whose assertion is supported or whose meaning is explicated or whose parts are detailed by the sentences added to it." (ibid.: 57–58) He further sets up two organizational schemes,

two-level and multi-level structures, by analogy with coordination and subordination in grammar. His presentation led to various reactions, but we shall only raise two points here. First, as has been repeatedly pointed out, it is hardly retainable that one should put a single topic sentence at the beginning of each o-paragraph. What may be identified as a topic sentence is often at the end of the o-paragraph, is split into the initial and final positions, or is not manifest at all (cf. Braddock, 1974). The topic sentence should rather be seen to comprise one of the writer's options. For readers, it is not something " out there ", but is a result of active participation. Second, the concept of topic sentence to some extent overlaps what others later called discourse topic. Current discourse analysis practically displays the content of a unit of discourse in the form of proposition, without much acknowledgment to Christensen and other works. That topic sentences are at times scattered or that they are not apparent at all in every o-paragraph may invalidate the traditional prescription in composition, but the fact is independent of the hypothesis that there is possibly a proposition-like representation for each s-paragraph.

An additional note, however, should be made to the extensive study by Markels (1983, 1984), in order to do justice to the status quo of paragraph study in composition. One criticism made there as pertaining to previous approaches is that they lack the consideration of form. Markels basically retains the traditional definition of o-paragraph as a development of a single idea, but tries to substantiate it with due regard to the surface syntactic form of constituent sentences. The key concept is " recurrence chain ", which in turn leads to the definition of cohesion as " (a) a dominating item recurrence present in or inferable from all sentences in a paragraph, and (b) a pattern or totality manifested by that item recurrence." (Markels, 1984: 21) This chain could be multiple, an improvement on traditional models like Christensen's, which can only handle a single-term chain (cf. ibid.: 88 f). This orientation is basically well-advised, but we find there remains a lot to be improved as regards the linguistic categories and functions employed therein, despite the frequent reference to the adjacent disciplines (as linguists,

we cannot accept that " . . . the sentence ' John bought a ball and the ball was red ' and the sentence ' John bought a red ball ' have the same deep structure but different surface structures " (Markels, 1984: 13)).

2. 2   *Linguistics*.   Modern linguistics has long limited the upper limit of analysis to the sentence.   Hence both s- and o-paragraphs have quite often been neglected.   One earlier exception is the studies by the fieldworkers of tagmemics.   This is quite natural because fieldworkers almost always have to deal with languages without orthography, yet they also have to segment the whole discourse into parts, which is of vital necessity in the field.   In this sense, the basic research strategy in tagmemics is to find out the formal markers of s-paragraphs through informant tests.

Numerous monographs written by tagmemically oriented researchers fully illustrate clear patternings of discourse, even though they lack the system of o-paragraphing.   A number of findings resulting from this program are taken up in Grimes (1975).   There it is also claimed that one of the general principles underlying the " partitioning " of discourse is the orientation of place, time, participants, etc., a principle widely accepted not only by tagmemic researchers, but also by those working in other frameworks.

Now we take up Longacre (1979, 1980) as representative of the systematics for paragraph analysis in tagmemics.   As usual in tagmemics, Longacre claims the grammatical structure of discourse is " partially expressed in the hierarchical breakdown . . . into constituent embedded discourse and paragraphs and in the breakdown of paragraphs into constituent embedded paragraphs and sentences." (Longacre, 1979: 115) His major contention is that " we are able to construct a system of paragraphs that does not compare unfavorable with constructing, for example, a system of clauses in a given language " (ibid.: 116).   He says an analysis of s-paragraph should be based on the notion of functional slots and its fillers, which may be a single sentence or a set of sentences that make up an *embedded paragraph.*

Longacre, after listing a set of closure markers and thematic unity markers (those marking paragraph topic), presents some basic parameters

as an descriptive apparatus of s-paragraph structures.   There are three parameters: binary versus n-ary, movement along a parameter (with the values of conjoining, temporal relations, logical relations, elaborative devices, and reportative devices), and weighing considerations. (ibid.: 121) The first is structural, the second notional, and the third rhythmical.

One remarkable achievement in tagmemics is the specification of the correlations between the s-paragraph markers and discourse genres. This suggests the direct contribution of micro-structure to macro-structure, namely the identification of the " plans " of discourse.   On the other hand, it is curious that there are few generalizations from the typological perspective, in spite of the numerous demonstrations from a wide variety of languages.

Apart from tagmemics, there seems to be no " school " of discourse analysis that is known for its consistent interest in the paragraph.   However, the works of a number of scholars deserve attention, some of which are taken up in the remainder of this section.   One is John Hinds, who in a series of papers explored some aspects of the paragraph, though to varying degrees.   Hinds, as his title suggests, explores the " paragraph level constraints " (1977: 78) on pronominalization.   He clearly distinguishes between s- and o-paragraphs, the former alone being the focus of analysis.   He also admits a hierarchical organization in discourse, dividing discourse into s-paragraphs, and s-paragraphs in turn into *segments*, " the sentences of which are related more closely to each other than to other sentences in the paragraph." (ibid.: 82)   Within a segment, there is one, and only one *peak sentence* where " a full noun phrase occurs, while it is within nonpeak sentences that pronouns occur." (ibid.: 82)   Within a segment, sentences that contain pronouns occurring before the peak build suspense or set the state, and those after elaborate or explain the information contained in the peak in the case of informative paragraphs.

The merit of Hind's study lies in his explicit correlation of the syntactic device of pronominalization with the addresser's control over the distribution of information: " . . . an author is free to organize basic information in any way he feels will highlight or dramatize the points

he wishes stressed." (ibid.: 84)　In other words, the choice between full NPs (notably names) and pronouns gives rise to different groupings of sentences into segments.　For example, in an obituary article which tells of an artist's death, two sentences relate his achievements.　Hinds distinguishes three possibilities of organization, among which both sentences take full NPs, i.e. they are highlighted.　This way of explanation, in our interpretation, casts light on a certain portion of the mapping problem.　But its weakness, as well as strength, is that the analysis is largely restricted to the contrast between pronouns and names. What is questionable is the status of definite full NPs.　The explanation Hinds provides is far less persuasive than those for other issues.　At the same time, such a restriction of scope allows him a systematic account of the phenomena under investigation.

Hinds (1979) is an expansion of the 1977 paper reviewed above, with the elaboration of the labels both for the function a segment performs within a paragraph and for that a sentence performs within a segment, according to different genres.　But they still seem to remain notional and do not differ much from traditional definition.

Another series of papers concerns topic continuity in Japanese, of which Hinds (1984) provides a synopsis.　Here he develops his thesis for Japanese that a participant is denoted by NP marked by *ga* for the first occurrence, by *wa* for the second, and by zero thereafter.　This chain can be violated in cases of orientation shift or rhetorical underlining (cf. the definition of segment).　These points do not seem novel, but his study on the topic continuity in Japanese discourse, together with his study on English pronominalization, would certainly provide a starting point for the contrastive analysis of discourse.

Makino (1979) is another study explicitly devoted to the paragraph. His primary aim is, as its title suggests, to explain the legitimacy of paragraph as a linguistic unit in descriptive terms (ibid.: 283).　Though not a crucial flaw, Makino does not distinguish between o- and s-para- graphs explicitly.　His first argument seems to be directed to the concept of o-paragraph.　That is, the discourse constraint on pronominalization by o-paragraph, as Hinds discussed.　But his main body of discussion

is devoted to the user's ability to divide the discourse into s-paragraphs.

Makino's experiment is similar to Sakuma's in that he provided the subjects (20 both for Japanese and English) with non-indented materials, and told them to give the most appropriate division into parts. But his experiment differs from Sakuma's in some respects. First, Makino's materials are chosen more carefully to provide the texts that perform the equivalent functions in both English and Japanese (e.g. recipe, medical diagnosis, etc.). Second, he did not tell the subjects to divide the whole material into parts by their own decision, but he made up several possible paragraphings beforehand, and let the subjects judge the appropriateness of each alternative. The percentages of the subjects' judgments about the acceptability of paragraphing ranged from 100% to 16%.

In his account on the organizational principles of s-paragraph, Makino mentions the local cohesion, marked by *wa*, lexical repetition, zero, etc. He observes that " it is next to impossible to paragraph a discourse chunk into two paragraphs when local cohesion in terms of phoricity is of high degree." (Makino, 1979: 286) He also affirms that there are several possibilities for o-paragraphing as long as it does not violate the hierarchical organization of s-paragraphing. But from our viewpoint, it might be necessary to separate the consideration of s-paragraphs (which Makino attributes to the " deep semantics ") and their formal cues from o-paragraphs.


3. At this point, we shall give a few general comments on the studies surveyed here, together with a tentative framework within which the s-paragraph should be investigated systematically. One common feature of the the above studies is an attempt to account for the hierarchical organization of paragraphs.

Another feature that such studies share is that they lack the exhaustive and systematic treatment of the formal cues that contribute to the identification of paragraphs. By systematic treatment we mean the establishment of the correlations (or more strongly, hierarchy) among the elements. That is, not all the elements have equal standing in the discourse process

in question. The medial level processes involve the multi-dimensional factors which interact with each other. The previous approaches, which are item-centered, do not provide sufficient account on this point, even though many of them find that the formal cues, as well as the content, play an important role in paragraphing. Our contention is that certain elements are given more preference than others, that they differ in their scope, and that such difference may be explained with due regard to other phases of linguistic structure in a principled way.

As is clearly seen, this orientation is in harmony with the tendency to treat languages so as to reveal the particularities of their structures as well as their non-randomness. Below, we shall elaborate on this a little further.

Givón (1981, 1983, 1984) introduces the concept of *functional domain* into typological studies, in preference to function, " because syntactic functions most commonly are not atomic and totally discrete." (Givón, 1981: 164) Such domains " are most commonly *clines*, upon which a number of more-or-less distinct points may be plotted along a functional *continuum*. These functional domains seem to be reasonably universal. But different languages may identify different discrete points or ' sections ' along a domain via syntactic structures that *code* these points or sections." (ibid.) Such a concept allows us to turn from the item-centered approach to the function-centered approach in the analysis of paragraphs. Thus the s-paragraph may be operationally re-defined in the following manner:

(8) S-paragraph is a discourse unit that stands between micro- and macro-structures, and is characterized in terms of the set of features $\{f_1, f_2, f_3, \ldots, f_n\}$ that belong to other functional domains as well. If all the features take default values, a proto-typical paragraph is realized.

Some of our basic standpoints may be derived from this. First, the s-paragraph is not what is definable in terms of necessary and sufficient conditions. As has been pointed out, the indeterminancy involved in the identification of the paragraph is not due to the inadequacies of

methodology, but to the nature of the object itself. To capture this, the adoption of prototypes seems most plausible. Next, the fact that features characterizing the s-paragraph belong to other functional domains implies that the s-paragraph is a domain intersecting others. In Givón's term, the functional domain of paragraphing is most approximate to that of " larger thematic organization ", which signals " various levels of the higher overall thematic organization of the discourse, in ways that transcend either individual participants or smaller chunks of sequenced information." (Givón, 1984: 35)

Givón further provides the following hierarchy, a reasonable framework within which the previous studies may fall:

(9)  THEME > ACTION > TOPICS / PARTICIPANTS

Many of the formal cues previously identified properly fit into this scheme: recurrence chain (Markels), orientation of participants (Grimes), and the constraints on pronominalization (Hinds) relate to the TOPICS / PARTICIPANTS continuity, tense / aspect / modality (cf. orientation of time) to the ACTION continuity, and connectives to the THEMATIC UNITY. The three levels stand in the relation of implicational hierarchy, i.e. " Most *commonly* it [N.B. the thematic paragraph] also preserves topic and action continuity. However, topics / participants may change within the discourse without *necessarily* changing either action continuity or theme continuity. Action continuity may change without *necessarily* changing thematic continuity." (Givón, 1983: 8)

In conclusion, we may claim to have demonstrated a more adequate framework for investigating the restricted area of the mapping problem. The next step to be followed is the detailed analysis of the formal cues for s-paragraphing from the functional-typological perspective.

### Bibliography

Becker, A. L. (1965) "A tagmemic approach to paragraph analysis," *College Composition and Communication* 16. pp. 237–242.

Bolinger, D. (1977) *Meaning and Form*. London: Longman.

Braddock, R. (1974) " The frequency and placement of topic sentences in expository prose," *Research in the Teaching of English* 8. pp. 287–302.

Christensen, F. (1967) "A generative rhetoric of the paragraph," in *Notes*

*toward a New Rhetoric*. New York: Harper. pp. 52–81. orig. in *College Composition and Communication* 16 (1965).

——, Becker, A. L., Rodgers, P. C., Miles, J., and Karrfalt, D. H. (1966) " Symposium on the paragraph," *College Composition and Communication* 17. pp. 60–87.

Freedle, R. O. and Carroll, J. B. (eds.) (1972) *Language Comprehension and the Acquisition of Knowledge*. Washington D.C.: Wiley.

Givón, T. (ed.) (1979) *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press.

—— (1981) " Typology and functional domains," *Studies in Language* 5. pp. 163–193.

—— (ed.) (1983) *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.

—— (1984) *Syntax: A Functional-Typological Introduction*, vol. I. Amsterdam: John Benjamins.

Hinds, J. (1977) " Paragraph structure and pronominalization," *Papers in Linguistics* 10. pp. 77–99.

—— (1979) " Organizational patterns in discourse," in Givón (ed.) pp. 135-157

—— (1984) " Topic maintenance in Japanese narratives and Japanese conversational interaction," *Discourse Processes* 7. pp. 465–482.

Hoey, M. (1983) *On the Surface of Discourse*. London: George Allen and Unwin.

Kintsch, W. (1982) "Aspects of text comprehension," in Le Ny, J. F. and Kintsch, W. (eds.) *Language and Comprehension*. Amsterdam: North-Holland. pp. 301–312.

—— (1985) " Text processing: a psychological model," in van Dijk (ed.) pp. 231–243.

Koen, F., Young, R. and Becker, A. L. (1969) " The psychological reality of the paragraph," *Journal of Verbal Learning and Verbal Behavior* 8. pp. 49–53.

Li, C. N. (ed.) (1976) *Subject and Topic*. New York: Academic Press.

Longacre, R. (1979) " The paragraph as a grammatical unit," in Givón (ed.) pp. 115-134.

—— (1980) "An apparatus for the identification of paragraph types," *Notes on Linguistics* 15. pp. 5–22.

Makino, S. (1979) " Paragraph, is it a legitimate linguistic unit?—A case study from English and Japanese," in Brown, L. L. and Steinmann, M. (eds.) *Rhetoric 78: Proceedings of Theory of Rhetoric: An Interdisciplinary Conference*. Minneapolis: University of Minnesota Center for Advanced Studies in Language, Style, and Literary Theory. pp. 283-296.

Markels, R. (1983) " Cohesion paradigms in paragraphs," *College English* 45.

pp. 450–464.

—— (1984)  *A New Perspective on Cohesion in Expository Paragraphs*.  Carbondale: Southern Illinois University Press.

Padučeva, E. V. (1974)  '' On the structure of the paragraph,'' *Linguistics* 131.  pp. 49–58.

Pike, K. L. and Pike, E. G. (1983)  *Text and Tagmeme*.  London: Frances Pinter.

Rodgers, P .C. (1966)  ''A discourse-centered rhetoric of the paragraph,'' *College Composition and Communication* 17.  pp. 2–11.

Sakuma, M. (1978)  '' Gendai America-jin no paragraph-ishiki,'' (The American sense of the paragraph—an investigation of reader's response to contemporary expository prose)  *Ningen Bunka Kenkyu* 2.  pp. 97–108.

—— (1984)  '' Yomite no danraku-kubun to bunsho no kouzou-genri,'' (Reader's paragraphing and the structural principles of discourse)  *Gengo* 13-3.  pp. 106–115.

Stern, A. (1976)  '' When is a paragraph? ''  *College Composition and Communication* 27.  pp. 253–257.

*Text* 3–3 ( =Special issue on formal models of discourse) (1983)

van Dijk, T. A. (1972)  *Some Aspects of Text Grammars*.  The Hague: Mouton.

—— (1985)  '' Semantic discourse analysis,'' in van Dijk (ed.).  pp. 103–136.

—— (ed.) (1985)  *Handbook of Discourse Analysis, vol. 2: Dimensions of Discourse*.  New York: Academic Press.

Young, R. and Becker, A. L. (1966)  '' The role of lexical and grammatical cues in paragraph recognition,'' *Studies in Language and Language Behavior*.  Progress Report No. 2.  Ann Arbor: Center for Research on Language, University of Michigan.