TILBURG ◆ ◆ UNIVERSITY

Tilburg University

The inventio phase of document design

Spooren, W.; Renkema, J.

Published in:
Document Design. Linking writers' goals to readers' needs

Publication date:
1999

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Spooren, W., & Renkema, J. (1999). The inventio phase of document design: On deciding what to communicate. In A. Maes, H. Hoeken, L. G. M. Noordman, & W. Spooren (Eds.), *Document Design. Linking writers' goals to readers' needs* (pp. 163-173). Tilburg University.

The inventio phase of document design
On deciding what to communicate

Wilbert Spooren & Jan Renkema

ABSTRACT

In document design studies, the phase of inventio (deciding what information to communicate) is often neglected. In this contribution we intend to remedy this situation, by combining the Quintillian inventio system with the Gricean notion of relevance. First we discuss the difference between writer-relevance (a writer's goals) and reader-relevance (a reader's expectations). Next we employ an analysis of leaflets and standard letters to put this distinction to use in practice: Domains can be explored systematically using topical questions that are grouped into patterns, depending on the nature of the domain. Finally we discuss a number of implementation problems for using the system in everyday practice.

1. INTRODUCTION

The main pages of the Unix operating system, the help files, are known for their user-unfriendliness. Specifically for outsiders it is difficult to find one's way around. The text in (1) is the first part of the help information for a command called "dir ", which gives the reader the format of directories.

(1)
NAME
dir - format of directories

DESCRIPTION
A directory behaves exactly like an ordinary file, save that no user may write into a directory and directories must be read using the getdirentries(2) system call or the directory(3V)library routines. The fact that a file is a directory is indicated by a bit in the flag word of its inode entry; see fs (5).

A directory consists of some number of blocks of DIRBLKSIZ bytes, where DIRBLKSIZ is chosen such that it can be transferred to disk in a single atomic operation (512 bytes on most machines).

[...]

Each DIRBLKSIZ byte block contains some number of directory entry structures, which are of variable length. Each directory entry has a struct direct at the front of it, containing its inode number, the length of the entry, and the length of the name contained in the entry. These are followed by the name padded to a 4-byte boundary with null bytes. All names are guaranteed null-terminated. [...]

What is noticeable about this text is that it gives all sorts of information that novices cannot understand, and that it doesn't give all sorts of information that novices would seem to need. What are "getdirentries(2) system calls " or "directory(3V) library routines"? Why don't we hear about the circumstances under which one needs a command like "dir "? Another problem is that information is given when we do not expect it. E.g., why is the very first information that we get the fact that directories behave like ordinary files? What we would expect first is information about what the command does. We think that the problem here is a problem of *content*: The author has used a typically self-oriented strategy, in which he or she has dumped all sorts of information in the text without regard of the informational needs of the reader.

In this contribution we intend to address the problem of content selection in document design; that is, the problem of deciding what to communicate. This is an important part of the document design process, that nevertheless has received remarkably little attention in the literature on writing texts. Handbooks of writing (e.g., Hairston, 1981; McNab McGrimmon, Trimmer & Sommers, 1984) usually restrict themselves to the advice to use something like the Quintillan topical questions or perform a brainstrom. The situation is not radically different for psychological literature on writing. For instance, the influential Flower & Hayes model of text production (1980), pays much attention to the generation phase, a flow chart of which is presented in Figure 1.

Figure 1: Generating phase in Flower & Hayes model
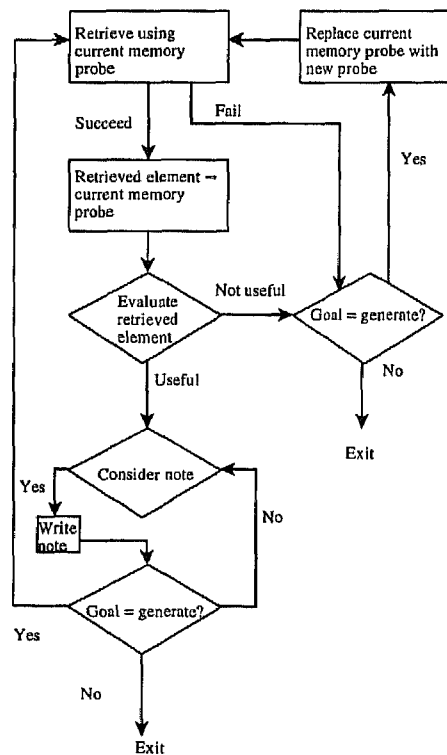
In this model the basis for generating content is searching through memory on the basis of a memory probe and judging whether or not the element retrieved is useful. If so, the element is transformed into a note and eventually into a text. The crucial point here is the notion 'usefulness': What is it that makes an element useful for a text?

We are under the impression that these issues are left to the imagination of the writer by the handbooks on writing. We will first discuss relevance as a usuable principle of content finding. Then we will say something about text types and content, leading to some heuristics for finding content for informative texts. We will discuss two case studies in which the heuristics were put to use, one case of document *creation* and one case of document *revision*. And finally we will sketch some conclusions.

2. RELEVANCE

What we want to suggest is that writers can optimize their content finding by following systematic heuristics, guided by a principle of relevance. Relevance has widely been recognized as an important, albeit difficult to formalize, conversational principle, for instance in the work of the linguists Sperber and Wilson (1996): They have suggested to consider information relevant if it has maximal contextual effects with minimal processing effort. This view on relevance is limited to the receptive side of the communication process. We want to suggest that the relevance principle of Sperber and Wilson can be looked at from the speaker's point of view (speaker relevance) *as well as* from the hearer's point of view (hearer relevance) (in case of written documents this would of course be writer relevance and reader relevance).

What does it mean for information to be writer relevant? Information is writer relevant if it contributes to fulfilling the writer's goals. Information is reader relevant if it contributes to the informational needs of the reader.

In what follows we will use writer relevance and reader relevance as criteria in content selection.

3. TEXT TYPES AND TEXT PATTERNS

Whether or not information is relevant at a particular point in a document is at least determined by the following factors:

Writer relevance: does the information contribute to reaching the writer's goal?
Reader relevance: does the information fulfil the reader's information needs?
Text type: is the information dictated by the type of the text?

In previous work, we have argued (Spooren, 1994) that text types can be looked upon as fixed linkages of writer goals and reader expectations. To the extent that this is correct, text types can be seen as the intermediary between writer relevance and reader relevance:

writer-relevance ➡   text-type   ⬅   reader-relevance

At present there is no satisfactory text type taxonomy. The ones available mix heterogeneous classification criteria of writer intentions, linguistic characteristics, discourse structure, medium choices etc. Yet there are some constants, that are probably going to show up in the 'ultimate' text type taxonomy. For instance, Werlich (1982) suggests the following text types:

Text type	Characteristic
Descriptive texts	informative goal, 'tenseless '
Narrative texts	informative goal, past tense
Explanatory texts	informative goal, causal links
Argumentative texts	persuasive, causal links
Instructive texts	directive

The main defining characteristic of text types is writer's goal: Through my text I want the reader to know something (informative texts), to have an opinion concerning something (argumentative texts), to be able to do something (instructive texts). We need additional criteria to make finer distinctions, and as Wehrlich suggests, one might consider features like tenselessness versus past tense as the main distinctive between descriptive and narrative texts.

This is not the place to propose an adequate text type taxonomy. The point we want to make is that text types come along with more or less fixed set of questions that are to be answered in the text. Below we give a number of such question sets, the source of which is Steehouder et al. (1992), who use them to evaluate the *structure* of documents; we employ them for generating and evaluating the *content* of documents. In argumentative texts one typically finds an evaluation pattern, in which the main questions are:

(1) Questions for evaluation patterns
What are the relevant properties of the topic under discussion?
What are the positive aspects?
What are the negative aspects?
What is the overall judgment?

and in an instructive text we typically find an action pattern, for which the relevant questions are:

(2) Questions for action patterns
What is the goal of the action?
What are the conditions for performing the action?
How is the action to be performed?
How are sub-actions to be performed?
How is the result of the action checked?

That is to say, the claim is that in order for a text to be writer-relevant, the content of the text has to pay attention to these questions. If this is not the case, then either this is licensed by considerations of reader relevance (the reader already knows the missing information because of his or her background) or the text will lead to reading problems. In the remainder

of this paper, we will substantiate this claim on the basis two case studies, both of which concern informative texts.

4. TEXT PATTERNS FOR INFORMATIVE TEXTS

Texts can be looked upon as communicative acts in which the writer anticipates questions of the reader. An answer to such a question can generate a new question, which needs to be answered, etcetera etcetera, ad infinitum. We have just argued that text types can be characterized by the primary sets of questions that readers expect to be answered by such a text. Which particular set is relevant for a text depends primarily on the writer's goal. Here we see how writer-relevance determines the content of a text. For various informative texts the following sets of questions are relevant.

(3) Prime communicative act: announcing a measure
What exactly is the measure?
Why is the measure needed?
How will the measure be carried out?
What are the effects of the measure?

(4) Prime communicative act: description (of an object/entity)
What does the object/entity look like?
In what respects does the object/entity ressemble/differ from other objects/entities?
What components does the object consist of?
How can the object/entity be defined?
How did the object/entity come about?
For whom is the object/entity intended?

(5) Prime communicative act: positive reaction to a request (approval)
What exactly is the approval?
On what arguments is the approval based?
What are the conditions under which the approval holds?
What are the consequences of the approval?

(6) Prime communicative act: negative reaction to a request (rejection)
What exactly is the rejection?
On what arguments is the rejection based?
What are the consequences of the rejection?
What are the conditions under which the rejection holds?
How can the reader lodge objections against the rejection?

In order to establish these sets of questions we used the following procedure: Set (4) is based on Dutch handbooks on text structuring (Steehouder et al., 1992; Pander Maat, 1994). Sets (3), (5) and (6) are based on a survey carried out under 44 undergraduate students from Tilburg University. They were asked to indicate on a 10-point scale how essential a particular component is for three complex communicative acts: announcing a measure, responding positively to a request, and responding negatively to a request (1: maximally essential, 10: minimally essential; 0: not relevant). They were asked for instance:

How essential are the following components for *announcing a measure* (1=most essential, 10= least essential)?

___   background of the measure
___   other measures resembling the present one
etc.

For each component it was established how many respondents found it relevant and what the mean essentiality score was. The questions mentioned in (3), (5) and (6) correspond to the components mentioned frequently and with a low essentiality score.

Writers can find content for their text by systematically answering these questions and if the answers suggest new topics, with corresponding new question sets, the secondary questions will have to be answered, and so on. This will lead to large trees of generated content. These trees need to be pruned. Criteria for pruning are considerations of writer- and reader-relevance:

(6) Criteria for pruning content
information is pruned if
□   it is not sufficiently directly related to the writer's goal (writer-relevance)
□   it is information that is already available or of no use to the reader (reader-relevance)

Case study 1: Generating content
How can this system of content finding be demonstrated in actual use? A school board from a Roman-Catholic school in a village in the vicinity of Tilburg was in need of a leaflet to inform the parents of two- and three-year olds about the characteristics of the school, so that the parents were able to make a motivated choice for sending their child to the school.

Here are some results of an initial analysis of client goals, reader characteristics and contextual restrictions:

(7) Analysis of client goals, reader characteristics, and contextual restrictions
     Client 's goal: reader knows the characteristics and strong points of the school, so that if the reader is interested, (s)he can ask for an extended brochure.

Reader characteristics:
□   readers vary in background and reading experience
□   most of the readers have a Roman-Catholic background
□   most of the readers have some background knowledge about the school, because it has been in town for quite a while

...

Context information:
□   the leaflet will be distributed at the open day or sent upon request
□   the leaflet is complementary to a lengthy brochure and a 'Calendar of activities '
□   length of the leaflet: no more than two pages A4
□   no budgetary restrictions

As the first goal of the client was to inform the audience, it was decided to make description the *prime communicative act. This lead to a first version of the content selection*, which was evaluated by discussing it with the client and by interviewing two couples who were considering to send their child to the school. The round of feedback lead to the conclusion that some of the questions in the Description pattern were less important for the writer's goal, because the information was already available in other documents and too detailed for the main goal to acquaint the reader with the strong points of the school. Moreover, some important strong points appeared to be missing (which was considered bad from the point of view of the secondary goal of persuading the reader to ask for the extended brochure). For instance, information concerning the exact location of the school in a small and safe street, away from busy traffic, was present in the first version, whereas information concerning the catholic identity was less prominent. The feedback round lead to the conclusion that the location information could be discarded with because of considerations of reader-relevance (all inhabitants of this small village have detailed background knowledge about the whereabouts of the school), whereas information concerning the identity was promoted because of considerations of writer-relevance (profiling the school against the, non-religious, competition).

All in all this lead to the following content selection (only the questions that lead to further extensions of the content are in the overview):

(9) Excerpt from content analysis
Topic: [xxx]-school [description pattern]

What is the [xxx]-school? A primary school
Where is the school located? in [xxx]
What is the size of the school? 130 pupils
What are other salient characteristics?

...

  The *cultural-religious basis* of the school [description]
   What is the cultural-religious basis of the school?
    Roman-Catholic [description]
     What does 'Roman-Catholic ' imply?
      1. attention for ideological and social values within the society
      2. raise children as responsible and tolerant persons, on a Christian basis, with respect for each other
      3. translate a Christian view to everyday practice
  For whom is the school intended?
   Children of four years and older who live in [xxx]
    What are the activities of the school?
    1. class activities [description]
    2. other activities [description]
   How is the school organized?
    1. school board [description]
    2. parent council [description]
    3. teachers ' council [description]
  How can one obtain additional information? [action]

This selection of content was then used for the further process of text design, and lead ultimately to a folder used by the school. What the example shows is that the propagated system can be used to select content in a process of developing a new document and to discuss this important part of the document design process with the client.

5. CONTENT ANALYSIS AND TEXT REVISION

Our second example is of a somewhat different nature, in that we used the system as a basis for analysis and revision of existing texts. The case we want to discuss concerns a project carried out in 1995-1996 at Tilburg University for the Dutch equivalent of the IRS. The IRS uses a large body of standard letters. These letters do not function properly and that is why IRS asked for an evaluation of the comunicative value of those standard letters. In this project content experts and communication experts worked together to evaluate the usefulness of the letters. They did that on the basis of a structured checklist for text quality. This lead to commentaries on a number of dimensions, such as text type, tone, content, structure, formulation and presentation. These commentaries were the basis for rewriting the letters. All in all more than 200 letters were analyzed, all concerning income tax and lottery tax.

What we did was analyze the commentaries made by the experts for a sample of those letters. The sample consisted of 20 letters, 10 letters in which measures were presented, and 10 reactions to a request. Of those reactions five were positive and five were negative. Below an example of such a letter is given, in this case a measure, namely a decision to impose a tax for a passenger car or motorcycle, in combination with an analysis of that letter, in terms of content.

Example letter BNA00102
                        Re: retroactive collection of taxes on passenger cars and motorcycles

                        Dear [GENDER] [NAME],

occasion for the measure	You have failed to declare taxes on passenger cars and motorcycles (TPM) for the period [...]. You have paid none of the money due. Therefore, you are now subject to retroactive collection, assessment number [...].
specification of the measure	You are required to pay the amount of Dfl. [...]. This sum consists of [1: taxes still owed on passenger cars Dfl. [...] /2: taxes still owed on motorcycles Dfl. [...]]
period of application of the measure	You must pay the full amount before [...].
procedure to comply with measure	You have again received a pre-printed giro credit slip that you can use for paying the assessment. If a different method of payment is used, please mention the assessment number and the abbreviation TPM. You can pay in cash at the post office. If you do so, you are required to hand in the giro credit slip.

how to lodge an objection	This decision is subject to appeal. If you do not agree with this decision, you must send a notice of objection to the Tax Department unit that has rendered this decision, within six weeks of date of notification. The postal address is given above this decision.
details of procedure to lodge an objection	A notice of objection will be considered on time if received within the term set. If sent by mail a notice of objection will be considered on time if it is postmarked before the end of the term and received no later than a week after the term limit has expired.

Yours sincerely,
Tax Department/[SIGNATURE]

We have analyzed these letters with respect to content, that is we have taken the questions pertaining to measures as a basis, as level 1 content; additional questions and their answers are taken as level 2 content etc. I.e. on the basis of the question set for measures we expected this type of letters to obey to the following scheme:

(9)   Hypothetical content of measure letters
Level 1   Level 2

What is the occasion for the measure?
What exactly is the measure?
                    [description pattern]
                    reach of the measure
                    period of application of the measure
Why is the measure needed?
What are the effects of the measure?
                    1. if positive [description]
                    Why these effects?
                    2. if negative
                    How can the reader lodge an objection?
                    [action pattern]

Level-1 information is closest to the goals of the writer and the reader expects it to be present because of the text type. It scores therefore high on both writer-relevance and reader-relevance. Lower level information scores lower on writer- and reader-relevance. Hence we expected that faulty level-1 information inevitably will lead to comments of the experts, whereas they were expected to be more lenient to lower level information. To what extent is this the case?

We analyzed the 42 comments the experts gave. These comments fall into two types: Either information is considered superfluous or it is indicated that necessary information is lacking. If we subdivide the comments into level 1 and level 2 comments this leads to the following picture:

	Comments	Superfluous	Necessary but lacking
Level 1		1	22
Level 2		8	11

Table 1: *Number of comments as function of comment type and level of information (case study 2).*

There is a strong relation between the type of comments and the level of the information: At level 1 the comments are predominantly that necessary information is lacking, whereas at level 2 both types of comments occur ($\chi^2 = 6.61$, p = .01).

The main point is that our approach corresponds significantly to the comments of the experts and therefore gives us a means for systematically talking about such an elusive matter as the content of a document.

6. CONCLUSION: IMPLICATIONS FOR DOCUMENT DESIGN

Of course these are not the amount and character of data to test the validity of the approach. Whatever data we have are merely indicative for the usefulness of the approach. What exactly is that usefulness?

☐ The approach provides us with a means of getting a grip on an elusive aspect of the document design process, namely content selection, that is usually out of the reach of document designers.

☐ In the process of creating a new document, it gives us a means of generating feedback from the client on content selection

☐ In text revision, it gives us a means of talking about and revising the content of existing documents

On a closing note, there is an interesting parallel between the lack of attention for the inventio phase of document design and the way classical rhetoric has been perceived through the ages. Classical rhetoric originally focused on the inventio phase. With time rhetoric has become associated almost exclusively with the elocutio phase, the phase of formulating, and with rhetorical tricks to seduce an audience. In a way, this paper is a plea for a reversal of this trend and for giving content selection its original prominence in the process of document design.

REFERENCES

Flower, L. & Hayes, J.R. (1980). The dynamics of composing; making plans and juggling constraints. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive Processes in Writing* (31-50). Hillsdale, NJ: Erlbaum.

Hairston, M. (1981). *Succesfull Writing: A Text for Advanced Composition.* New York: Norton.

McNab McGrimmon, J., Trimmer, J.F. & Sommers, N.I. (1984). *Writing with a Purpose.* Boston, etc.: Houghton Mifflin.

Pander Maat, H. (1994). *Tekstanalyse: een Pragmatische Benadering* [Text Analysis: A Pragmatic Approach]. Groningen: Martinus Nijhoff.

Sperber, D. & Wilson, D. *(1996). Relevance : Communication and Cognition* (3rd ed.). Oxford etc.: Blackwell.

Spooren, W. (1994). Teksttypen - tekstsoorten: tekstwetenschappelijke opmerkingen over de bruikbaarheid van teksttypologieën [Text types - text sorts: discourse study remarks on the usability of text typologies]. *Spiegel, 12* (2), 59-78.

Steehouder, M., Jansen, C., Maat, K. Staak, van der J. & Woudstra, E. (1992). *Leren communiceren* [Learning to communicate] (3rd, compl. rev. ed.). Groningen: Wolters-Noordhoff.

Werlich, E. (1982). *A Text Grammar of English.* Heidelberg: Quelle & Meyer.