

## Word Order and Anaphora(語順と照応)

著者	秋月 高太郎
号	88
発行年	1997
URL	<a href="http://hdl.handle.net/10097/12775">http://hdl.handle.net/10097/12775</a>

氏 名 (本 籍)	あき づき こう た 郎 秋 月 高 太 郎	(東 京 都)
学 位 の 種 類	博 士 (情 報 科 学)	
学 位 記 番 号	情 博 第 88 号	
学 位 授 与 年 月 日	平 成 10 年 3 月 25 日	
学 位 授 与 の 要 件	学 位 規 則 第 4 条 第 1 項 該 当	
研 究 科 , 専 攻	東 北 大 学 大 学 院 情 報 科 学 研 究 科 (博 士 課 程) 人 間 社 会 情 報 科 学 専 攻	
学 位 論 文 題 目	Word Order and Anaphora (語 順 と 照 応)	
論 文 審 査 委 員	(主 査) 東 北 大 学 教 授 福 地 肇 東 北 大 学 教 授 平 田 満 男 東 北 大 学 教 授 浅 野 檀 英 東 北 大 学 助 教 授 浅 川 照 夫 東 北 大 学 助 教 授 菊 地 朗	

## 論 文 内 容 要 旨

This thesis is about some linguistic forms or constructions relevant to human's information-management mechanism. This is based on the idea that human language is designed as information-managers. We take human language as a "complex" system which involves some aspects that it is difficult or impossible to reduce into more minimal items. The goal I would like to reach is not a theory of what we can say but a theory of how we can say or why we should say so. Then, we should focus on real forms or constructions of human language which are used in real contexts. Our concern is why we say so in a given context. This would be accounted for by making clear the factors that influence a speaker's choice to say something or the contexts or world surrounding the speaker and hearer.

In Chapter One we argue against the distinction between linguistic and extra-linguistic/cultural information. It is hard to distinguish linguistic information from extra-linguistic/cultural information. Linguistic information is a subset of extra-linguistic/cultural information, with fuzzy boundaries. In conversation or reading and writing, the speaker/writer always pays attention to what the hearer/reader knows. The speaker/writer has to know what types of world the hearer/reader designs in his/her brain. The felicity of an utterance solely depends on the world the speaker/writer and the hearer/reader design. However, the speaker/writer cannot know the hearer/reader's knowledge in a direct way, because human cognitive system is closed in an individual. Clark and Marshall (1981) proposed the way that the speaker accesses the evidence which the speaker and hearer can take for granted and calculates the mutual information based on the evidence. We adopt the way of calculation of mutual information proposed by Clark and Marshall, and suggest that some language forms and constructions (particles, anaphoric expressions, and "free" word-order phenomena) are accounted for in the model of information-management.

In Chapter Two we propose the model of human's memory system and information-management. As human's memory system, we adopt the multiple storage model. Human's memory system is divided into three different

levels : working memory, episodic memory, and permanent semantic memory. Working memory is the level where the information retrieved or coming from the outside-world is processed. The information processed in working memory is stored in episodic memory. The information about conceptual and procedural meaning of lexical items is stored in the permanent semantic memory. We adopt the idea that information-processing is taken as a file-management. In the model of file-management, we proposed the notion of “folder”. We assume that the speaker and hearer form a folder in their working memories in conversation. Introducing a new topic, keeping a topic, or reactivating an old topic are taken as management instructions of folders in working memory.

In Chapter Three we try to account for some particles in Japanese in the model of information-management, proposed in Chapter Two. The particle *wa* is taken as a folder-activator. The function of *wa* is to reactivate an inactive or decayed folder in working memory. The final particles *ne* and *yo* are sensitive to the calculation to mutual information. The speaker attaches *ne* when (s)he has the evidence to hold mutual information, while (s)he attaches *yo* when (s)he has no evidence to hold mutual information.

Chapter Four treats anaphoric expressions in Japanese. We suggest that anaphoric expressions in Japanese are sensitive to different levels of memory storages. Zero pronouns, *ko*, and *so* are sensitive to working memory (WM), *kare/kanajo* and *a* are sensitive to episodic memory (EM), and bare nouns are sensitive to permanent semantic memory (PSM). The WM-oriented referring expressions are unmarked ones, in the sense that they directly refer to the entity, while the EM-oriented and PSM-oriented referring expressions are marked ones, in the sense that they indirectly refer to the entity via the episodic memory or permanent semantic memory. To use a referring expression and not to use another (if it can be used) raise implications. To use the WM-oriented referring expression implies that the speaker/writer assumes that the entity to which it refers is not stored in the hearer/reader's episodic memory. To use the EM-oriented referring expression gives the hearer/reader a signal not to access his/her episodic memory. On the other hand, to use the EM-oriented referring expressions gives the hearer/reader a signal to access his/her episodic memory. To use the PSM-oriented referring expression follows from the lack of the independent WM-oriented referring expressions to refer to human/animate entities in Japanese.

In Chapter Five we focus on “free” word-order phenomena in Japanese. The basic idea is that word order is derived from the speaker/writer's adjustment of information-flow. We propose four principles of the adjustment of information-flow : Given-First Principle, From-Background-to-Foreground Principle, Salient-Order Arrangement Principle, and Focus-Marking Principle. Salient-Order Arrangement Principle is a default one : that is, if the speaker/writer has no special intentions, word order is arranged from the most salient one to the least salient one. Given-First Principle is sensitive to the state of the speaker/writer's working memory at a given time. The speaker/writer tends to pick up the information in the working memory and to locate it in the sentence-initial position. From-Background-to-Foreground Principle is to arrange information from ground to figure. Focus-Marking Principle is to mark a focus on a piece of information on which the speaker/writer puts a special intention. If the speaker/writer tries to introduce a new topic, (s)he tends to pick up an entity in the working memory at a given time, and relate it to a new topic about which (s)he wants to talk or write. This makes the “given information” located sentence-initially. On the other hand, if the speaker/writer tries to talk/write about a certain thing, (s)he keeps the folder formed in the working memory active, and the name of the folder is referred by the least cost anaphoric device (i.e., a zero pronoun). In this case, the “new information” comes first. If the speaker/writer has no special intention of the flow of information, the salient-order arrangement principle is applied by default. In this case, the pieces of information are ordered from the most salient information to the least salient information.

Chapter Six is concluding remarks. For four decades, clearer light has been shed on human cognitive systems

in cognitive science, psychology, neurology, ethology, AI, and so on. In the field of psychology or neurology, a theory of language-brain relations has been developed, and they have tried to seek a model of how language-processing is translated into neural action. It is clear that a particular function or processing of language is related to a particular area of human brain. In the field of ethology, the cognitive systems of animals other than human beings have been clearer. This means that the comparative study of the cognitive systems of human beings and other animals is possible now. This comparative cognitive study will shed new light on the innateness of knowledge of language in human beings. In the field of AI or computational linguistics, it is possible that some of the activities of human beings are simulated by computers. Now, linguists, people who are interested in "formal" or "pure" aspects of human language, do not ignore the results of these fields. I think, this thesis is a "bridge" between linguistics and non-linguistic approaches to human cognitive systems.

## 審査結果の要旨

自然言語の内部構造および言語使用・言語による情報伝達過程の分析には、伝統的に統語構造と形式論理的意味構造を重視する手法と言語の認知・機能的視点からの手法があるが、本論文は後者の立場から、主として現代日本語を資料として、談話テキスト内における不変化詞の出没とそのはたらき、照応表現の生起パターン、統語規則外の語順決定法則を明らかにしようとしたもので、全6章から成る。

第1章は序論であり、言語使用により伝達される多種多様な情報カテゴリーを、多くの先行研究を批判的に検証しながら整理している。さらに、伝統的なテキスト言語学で言及される情報の新旧の概念を精密化して、話者・聴者間の共有情報の管理という自然言語分析上の独自の視点を打ち出している。

第2章では、第3章以下での具体的な分析の基礎となる言語モデルを提示する。これは、言語活動に関わる記憶のタイプとそこで処理・貯蔵される情報の種類を規定する部分と、ファイル管理システムとしての言語情報処理部分とから成り、話者による旧情報への直接接近だけでなく、話者による話者・聴者共有知識の推定過程を説明する。これは、Clark-Marshall (1981) のモデルを参考にしたものではあるが、作業記憶に関する新たな仮説を加えたことにより、談話での主題と視点の唯一性をも説明できる、より一般性の高い言語モデルになっている。

第3章は、共有情報管理モデルにより、不変化詞としての日本語の格助詞「は」および終助詞「よ」、「ね」、「よね」の機能を説明している。前者に関してはこれまでに膨大な考察があるが、そのほとんどは対比性、総称的既知性を含め旧情報との関わりから論じたものである。それに対し、本論文では、談話の冒頭に生起する非総称的名詞句に付く「は」に着目し、作業記憶内情報の活性化という新たな機能視点を提示している。後者に関してはこれまでに目立った考察はなく、本論文では、これら不変化詞の機能を、話者が聴者との共有情報であると想定して記憶システムに保存するためのマーカーとして説明している。これらは貴重な知見である。

第4章では、前述のモデルによって代名詞などの照応表現について新たな説明を試みている。ゼロ代名詞や「こ」「そ」系の指示表現は作業記憶指向、通常の代名詞や「あ」系指示表現がエピソード記憶指向、照応的名詞表現が意味記憶指向になるなど日本語の照応表現が記憶システムの異なるレベルにアクセスすることを示している。従来論考と異なり、本論文の提示する説明は照応表現目的語の前置など語順変化の要因との関わりをも考慮に入れている点で、今後の研究のあるべき方向の一端を示す重要な考察であると言える。

第5章では、提案モデルに基づいて、日本語の統後規則外の語順原則について説明を試みている。自然言語における情報の線条化と話者の意図との視点から、「既情報先頭の原則」、「後景情報から前景情報への原則」、「認知的際立ちの原則」、「焦点付与の原則」等の語順原則を設定し、著者独自に収集したコーパスによりその妥当性を示すことに成功している。このような原則は従来から部分的に論じられてはいるが、本論文では原則間の相互関連に視点を置いた考察がなされており、自然言語の機能分析における貴重な知見を提供している。

第6章は結論である。

以上要するに本論文は、主として認知心理学的な研究の成果を採り入れることによって、従来の言語学的記述説明では取り扱われてこなかった非言語的情報をも取り込むシステムを提案し、談話テキストに関わりの深い諸現象を検証することによってその妥当性を示したもので、情報科学の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。