Modeling Categorization Dynamics through Conversation by Constructive Approach

Takashi Hashimoto
Lab. for Information Synthesis
Brain Science Institute, RIKEN
2-1, Hirosawa, Wako, Saitama, 351-01, JAPAN
tel: +81-48-467-9663 fax: +81-48-467-9693
takashi@brain.riken.go.jp
http://www.bip.riken.go.jp/irl/takashi/

May 13, 1998

Abstract

Categorization dynamics as clustering of words in word relation is studied by a constructive approach which is suit for inquiry of evolutionary linguistic with dynamic view of language. Word meaning is represented by relationship among words to some extent and the relationship should be derived from way of usage of language. Based on this usage-base view, an algorithm to evaluate word relationship is defined. Using the algorithm, cluster structure and its dynamics of words are shown in a model with communicating artificial agents. The relevance of clustering with categorization is discussed.

Submission Type: regular paper

Topic Areas: Evolutionary Computational Linguistics

Author of Record: Takashi Hashimoto (takashi@brain.riken.go.jp)

Under consideration for other conferences (specify)? submission to other conference(s):none

Modeling Categorization Dynamics through Conversation by Constructive Approach

Abstract

Categorization dynamics as clustering of words in word relation is studied by a constructive approach which is suit for inquiry of evolutionary linguistic with dynamic view of language. Word meaning is represented by relationship among words to some extent and the relationship should be derived from way of usage of language. Based on this usage-base view, an algorithm to evaluate word relationship is defined. Using the algorithm, cluster structure and its dynamics of words are shown in a model with communicating artificial agents. The relevance of clustering with categorization is discussed.

1 Introduction

There are two ways of viewing language: structurally and dynamically. The structural view is a static one in which language structure, for example syntax (Chomsky, 1957) or pragmatic rules (Grice, 1975; Sperber and Wilson, 1986), offer idealized approaches to language. The alternative view is dynamic. It concentrates on the actual use of language rather than abstract notions of how language ought to be. It is possible to better understand the value of the second approach by thinking of metaphor.[1] Metaphorical expressions are creative or dynamic precisely because they can "bend" or "break" the rules of conventionally structured language. By producing or understanding metaphorical expressions, especially creative metaphors, our categorical system should change. We can not say valid or not valid for creative expressions, since the expressions are so novel that it is not valid for a conventional language structure. We should consider which we accept or not the expressions. If we accept them, our internal structure changes and language structure might also become to be modified. The whole system of such dynamical processes is considered as language in this view.

Evolutionary linguistics throws a bridge across these two viewpoints from the latter to the former. Global, i.e. common in individuals, structure described in the former view should be emerged as results of the process of using language and adaptation of its users. Note that the global structure in language should not be static. Change in internal structure by using language induces dynamics of the global structure

This standpoint implies that dynamics of language should be understood in the context of *complex systems*. In complex systems, it is important that global and local levels are dynamically coupled and by the dynamical coupling not only emergence of global structure but its dynamics such as collective behavior, diversification, differentiation and level formation are brought through interaction among elements (Kaneko and Tsuda, 1997).

Dynamical stability and adaptability which are often seen in complex systems must be equipped for language systems. If a language is too rigid, its users will not be able to formulate new expressions to describe diverse experiences, and if it is too unstable, no structurization both in individual and global level will be possible and hence no communication will take place. Geeraerts (1985) explains this point as it pertains to categorization: "To prevent the categorical system from becoming chaotic, it should have a built-in tendency towards structural stability, but this stability should not become rigidity, lest the system stops being able to adapt itself to the ever-changing circumstances of the outside world."

[1] Cognitive linguists often say that metaphor, and metonymy, are important when we consider linguistic categorization (Lakoff and Johnson, 1980; Taylor, 1995).

Constructive approaches are highly advantageous for understanding dynamically complex systems (Kaneko and Tsuda, 1994). These approaches are also useful for studying evolutionary linguistics (Steels, 1997). In contrast to conventional linguistics which attempts to describe various language phenomena, emergence of global order as language-like behavior is modeled through interaction among individuals using language in the constructive approach. However, as we have asserted above, only emergence of global order is not enough. Since language is indeed an ever-changing system, not only emergence but also the dynamics of global order should be observed in constructive models. Perhaps the most important consideration in the modeling of evolutionary language system is the introduction of the dynamics of individuals so as to be able to change their internal states and their relationship to other elements and to circumstances. We can get a look at dynamical behavior in the global level.

We propose *usage-based viewpoint* as that meanings of words should be discussed in terms of how language is used (Wittgenstein, 1953). Since interrelationship among words can be employed as a representation of meanings of words to some extent, as Cruse (1986) has written, "We can picture the meaning of a word as a pattern of affinities and disaffinities with all the other words in the language with which it is capable of contracting semantic relation in grammatical context," the point of view means that a word's relationship to other words should be derived by analyzing the word's usage in the language, not from its indication. A word in a sentence is understood not only as relation with only entities mentioned by the sentence but as positioning in the whole system of language.

In this paper, we model categorization dynamics as clustering of words in word relation space by the constructive approach. We will present a way to calculate relationship among words from their usage in sentences and texts[2] in the next section.

[2]Here a text means a series of conversation.

2 Word Relationship based on its Usage

We will discuss dynamics of categorization by how the relationship among words changes through conversation. Relationship among words are considered based on ways of use of words in sentences and those in texts. The algorithm to calculate relationship between words in sentences, we call it similarity in this paper, is basically attributed to Karov and Edelman's algorithm. We revise it to be suit for the current interest. The main revision is that we introduce one higher level, texts. A text is a set of sentences. In our paper, it is a series of conversation. Correlation of appearance between words and between sentences in texts is also considered for making relationship among words. Since the relationship develops gradually with uttering and accepting sentences in conversation, they are calculated successively per each utterance and acceptance.

Karov and Edelman (1996) have presented an algorithm for word sense disambiguation based on similarity among words. A key concept in the algorithm is the mutual dependency between words and sentences. That is, similar words appear in similar sentences and similar sentences are composed of similar words. Similarity between words are calculated from similarity among sentences in which the words are used. Similarity between sentences are done from that of words which are used in the sentences.

Paradigmatic relation which is semantic similarity between two grammatically identical words are considered in part from syntagmatic relation which is contiguity of words in a sentence, by this concept. For instance, suppose there are two sentences *I read a book* and *I read a magazine*. At first, the words *read* and *book* which are in syntagmatic relation in the first sentence become to have similarity from the sentence and the words *read* and *magazine* do so from the second sentence. Through the basis of their relation with the word *read*, the words in paradigmatic relation, *book* and *magazine*, can also be have similarity.

If patterns of appearance of words in texts are similar, e.g. some words appear many times in a text but do not so in the other texts, they have high affinity. Conversely, different appearance patterns between words imply disaffinity between the words. Such correlation between words are also considered for word relationship.

The relationship between words and between sentences are respectively defined by the following formulae:

$$R_{t,n}(w_i, w_j) = \quad (1)$$
$$\begin{cases} \alpha^w \frac{\sum_{s \ni w_i} K_{t,n}(s) A_{t,n}(s, w_j)}{\sum_{s \ni w_i} K_{t,n}(s)} + \\ \quad \frac{1-\alpha^w}{2} \left(\frac{D_{t,n}(w_i, w_j)}{\sqrt{N_{t,n}(w_i) N_{t,n}(w_j)}} + 1 \right) & \text{if } i \neq j \ , \\ 1.0 & \text{if } i = j \ , \end{cases}$$

$$R_{t,n}(s_i, s_j) = \quad (2)$$
$$\begin{cases} \alpha^s \frac{\sum_{w \in s_i} K_{t,n}(w) A_{t,n}(w, s_j)}{\sum_{w \in s_i} K_{t,n}(w)} + \\ \quad \frac{1-\alpha^s}{2} \left(\frac{D_{t,n}(s_i, s_j)}{\sqrt{N_{t,n}(s_i) N_{t,n}(s_j)}} + 1 \right) & \text{if } i \neq j \ , \\ 1.0 & \text{if } i = j \ . \end{cases}$$

In the above formulae, the suffixes n and t indicate a sentence and a text , respectively. The numbering system for sentences and texts is illustrated in Fig.1. The binary relation $w \in s$ means words included in a sentence s, and $s \ni w$ means sentences using a word w. For example, if there are two sentence s_1:*I read a book* and s_2:*I read a magazine*, a set $\{w | w \in s_1\}$ is {*I, read, a, book*}, a set $\{s | s \ni I\}$ is $\{s_1, s_2\}$ and a set $\{s | s \ni book\}$ is $\{s_1\}$. The coefficients α^w and α^s are weights for similarity and correlation in relationship between words and between sentences, respectively.

The first terms in Eqs. (1) and (2) count usage similarity of words used in sentences and that of sentences containing words, respectively. Similarity between words w_i and w_j is the weighted average of affinity of sentences using the word w_i to the word w_j. Similarity between sentences s_i and s_j is the weighted sum of affinity of words in the sentence s_i to the sentence s_j. The affinity of sentences to words and that of words to sentences are defined using the relationship between words and between sentences as follows:

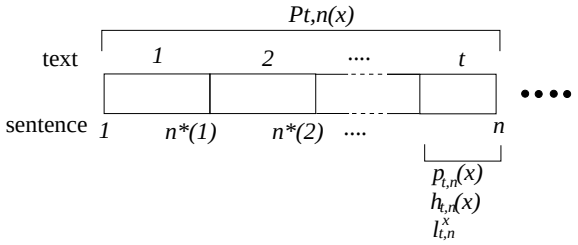

Figure 1: The numbering system for texts and sentences; and the ranges of sentence for $P_{t,n}(x)$, $p_{t,n}(x)$, $h_{t,n}(x)$ and $l^x_{t,n}$ are illustrated.

$$A_{t,n}(s, w) = \frac{\sum_{s' \ni w} K_{t,n}(s') R_{t,n}(s, s')}{\sum_{s' \ni w} K_{t,n}(s')} \ , \quad (3)$$

and

$$A_{t,n}(w, s) = \frac{\sum_{w' \in s} K_{t,n}(w') R_{t,n-1}(w, w')}{\sum_{w' \in s} K_{t,n}(w')} \ , \quad (4)$$

respectively. In these four equations (1) $\sim$ (4), weights for similarities and affinities define contribution from appearance provability of each word and sentence and length of each sentence. The weight factors are defined as

$$K_{t,n}(s) = \frac{P_{t,n}(s)}{lg(s)} \ , \quad (5)$$

$$K_{t,n}(w) = \frac{1}{P_{t,n}(w)} \ , \quad (6)$$

where $lg(s)$ is the length of a sentence s, which is defined by the number of words included in the sentence; $P_{t,n}(w)$ and $P_{t,n}(s)$, to be defined later, are the appearance provability of a word w and that of a sentence s in all sentences from the first to the nth. sentence, respectively. These weight factors mean that the more a word is used, the less informative it is; the more a sentence is used, the greater its contribution is; a word in a longer sentence is less important than one in a shorter sentence.

The second terms in Eqs. (1) and (2) count correlation of appearance of words and sentences in texts. Since the functions $D_{t,n}$ and $N_{t,n}$ are symmetrically defined for words and sentences, a symbol x is used to denote a word

(w) or a sentence (s) in the following equations. Since the correlation of words and sentences should be incrementally updated per each utterance and acceptance of sentences in order to consider dynamics of internal structure through conversation, the functions $D_{t,n}$ and $N_{t,n}$ are recursively defined in different with normal correlation coefficient:

$$D_{t,n}(x_i, x_j) = D_{t-1,n^*(t-1)}(x_i, x_j) + {l^x_{t,n}}^2 (p_{t,n}(x_i) - P_{t,n}(x_i))(p_{t,n}(x_j) - P_{t,n}(x_j)) \ , \quad (7)$$

and

$$N_{t,n}(x_i) = N_{t-1,n^*(t-1)}(x_i) + {l^x_{t,n}}^2 (p_{t,n}(x_i) - P_{t,n}(x_i))^2 \ , \quad (8)$$

where $n^*(t)$ is the last sentence in the tth. text, so $D_{t-1,n^*(t-1)}(x_i, x_j)$ is the deviation between x_i and x_j in the all sentences by the end of the text $t-1$.

The appearance provability of x in all sentences from the first to the nth. sentence $P_{t,n}(x)$ is defined as

$$P_{t,n}(x) = \frac{\sum_{t'=1}^{t-1} h_{t',n^*(t')}(x) + h_{t,n}(x)}{\sum_{t'=1}^{t-1} l^x_{t',n^*(t')} + l^x_{t,n}} \ , \quad (9)$$

and the appearance provability of x in a text t by the nth. sentence $p_{t,n}(x)$ is defined as

$$p_{t,n}(x) = \frac{h_{t,n}(x)}{l^x_{t,n}} \ , \quad (10)$$

where $h_{t,n}(x)$ is frequency, i.e. the number of appearance, of a word or a sentence and $l^x_{t,n}$ is the length of a text counted by x, i.e. the number of all words or that of all sentences, in a text t by the nth. sentence (see Fig.1).

By uttering or accepting a sentence, the word-sentence affinity (Eq. (4)) is calculated from word relationship at the previous sentence; and correlation of words and sentences are updated. Then, these formulae are calculated in the order of Eqs. (2) $\rightarrow$ (3) $\rightarrow$ (1). Before the first sentence, word relationship with itself is 1.0 and that with the others are 0.0, i.e. $R_{1,0}(w_i, w_j) = \delta_{i,j}$. It is the same for new words. In the first text, $D_{1,n}(x_i, x_j)$ and $N_{1,n}(x)$ are 0, since there is no difference between $p_{1,n}$ and $P_{1,n}$, but the correlation of a word or a sentence with itself is always 1, i.e. $D_{1,n}(x_i, x_j)/\sqrt{N_{1,n}(x_i)N_{1,n}(x_j)} = \delta_{i,j}$.

3 Modeling for Conversation

In our model there are M **topics** ($0 \sim M$). The topics have periodic boundary, so the neighbors of a topic 0 are M and 1. A randomly selected topic with zero, one or two adjacent topics are presented for agents for a conversation. The presented topics is called **situation** and a series of conversation is called **text**.

A speaker agent selects a topic randomly from the presented situation and make a sentence for the topic. In case that an agent does not have a name for the selected topic at utterance, a word created by combining characters as a name for the topic. In the initial state, agents does not have a name for each topic.

If there is zero or one unknown word in the uttered sentence, another agent, hearer, accepts the sentence. If the uttered sentence is accepted, the agents take turns their roles of speaker and hearer. The new speaker agent makes a sentence to reply to the accepted sentence. If the uttered sentence is not accepted by hearer, they don't take turns and the speaker randomly selects a topic from the presented situation again. When 3 sentences are not accepted or the number of sentences accepted by both agents becomes 100, the text ends up. Then a new situation is selected for a new text.

To make a sentence for utterance a speaker selects a word arbitrary from the accepted sentence as a focus word. At the commencement of a text the name for the selected topic plays the role of the focus word. The speaker adopts the highest affinity sentence with the focus word from the word-sentence affinity matrix (Eq.(4)) of the agent. If the adopted sentence is the same with the accepted sentence, the second highest affinity sentence is adopted. If the sentence is the same with the uttered sentence by the speaker at the immediately previous conversation, then the third highest affinity sentence is

adopted.[3]

Agents modify the adopted sentence to utter according to the creativity rate c. There are three modification ways: add, replace, or delete a word. One of the three is selected according to the sentence modification rates, $m^s_{\rm add}$, $m^s_{\rm rep}$, and $m^s_{\rm del}$, respectively. A word for addition or replacement is selected from already known words. The word is also modified by one of the three ways: add, replace or delete a character according to the word modification rates, $m^w_{\rm add}$, $m^w_{\rm rep}$, and $m^w_{\rm del}$, respectively.

Relationship among words is updated per each utterance of a sentence for the speaker agent and per each acceptance of a sentence for the hearer agent.

4 Simulation Results

In this section, we report the results obtained from simulations of two agents conversations to see internal dynamics of individuals. The simulations are done in the following parameters: The maximum of topics M is 9. The set of characters is {a, e, i, o, u}. The creativity rate c is 0.1. The sentence modification rates, $m^s_{\rm add} = 0.2$, $m^s_{\rm rep} = 0.7$, and $m^s_{\rm del} = 0.1$. The word modification rates, $m^w_{\rm add} = 0.33$, $m^w_{\rm rep} = 0.34$, and $m^w_{\rm del} = 0.33$. The weight coefficients for similarity and correlation in Eqs. (1) and (2) are $\alpha^w = \alpha^s = 0.4$. The results described here are under the specific parameter values, however almost same results are observed in the wide range of the parameter values.

An example of typical conversations are as follows:

```
TEXT: 25 SITUATION: 0 1 2

TOPIC: 2
agent1.name[2]= ea          agent0.name[2]= i

1: ea-i                     0: ea-eii-iii-iuuo-eii
1: ea-eii-eii-eii           0: ea-i
1: ea-eii-iii-iuuo-eii      0: iuuo
1: eoi                      0: i-eoi
1: i-eoi                    0: eoi
1: i-ei-eoi                 0: i-eoi
1: eoi                      0: i-ei-eoi
1: i                        0: i-eii
1: eii-eii-eii              0: eii-eii
1: eii-eii                  0: eii-eii-ii-uuoe
1: eii-eii-eii              0: eii-eii
1: eii-eii                  0: eii-eii-ei
```

This example is the beginning part of text $t = 25$ of a simulation. The situation is composed of the topics 0, 1 and 2. The agent 1 selects the topic 2. The name for the topic 2 of the agent 1 is `ea` and that of the agent 0 is `i`. As in this example, the name for a topic can be different between agents in conversation. The agent 1 utters a sentence `ea-i` from the focus word `ea` which is the name for topic 2. The agent 0 accepts the sentence and replies a sentence `ea-eii-iii-iuuo-eii` from the focus word `ea`.

4.1 Word Cluster

By processing the word relationship matrix $R(w_i, w_j)$ by the principal coordinate analysis, one method of the multi-dimensional scaling (Gordon, 1981), the characteristics of word relationship among words is clearly seen in scattered diagrams. Fig.2 is an example of the scattered diagram at text $t = 3$. The first two coordinates obtained from the principal coordinate analysis are used for X and Y axes, respectively, to draw the scattered diagram. For all values of the parameters, such clustered word groups are observed.

We can see four clusters of words in this figure, $1 \sim 4$, $5 \sim 7$, $8 \sim 12$, and 13. Words in the two clusters, $1 \sim 4$ and $8 \sim 12$, are distributed rather small regions. Words in the cluster $5 \sim 7$ connects between these two clusters. The word 13 is rather solitary and has almost same relation with both of the two clusters in small regions.

We can get more information about structure of clusters by drawing the word relationship directly as in Fig.3. The numbers in Fig.2 correspond to words in X and Y axes in Fig.3. Words in a cluster scattered in smaller regions have relatively closer relationship with each other. Such cluster has a flat top if it is drawn in the graph as Fig.3. The words $1 \sim 2$ and $9 \sim 12$ indeed have flat tops as in the graph. We call such cluster a flat cluster. In a cluster lying in a narrow and long region, like the cluster $5 \sim 7$, neighbor

[3] Because of this setting, a period three conversation is likely to occur.

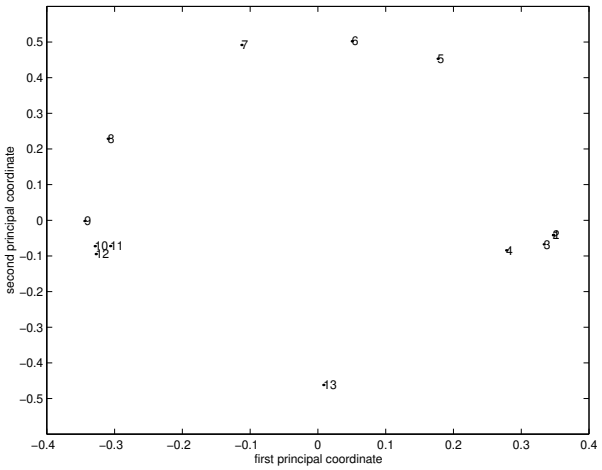

Figure 2: An example of scattered diagram of the word relation matrix $R(w_i, w_j)$ processed by the principal coordinate analysis. X and Y axes are the first and second principal coordinates obtained by the analysis, respectively. This data is at the end of text $t = 3$. Each number in the diagram indicates a word. Starting from a word in a right center cluster, each word is numbered as that words in closer position in the diagram have the closer numbers and roughly counter clockwise. Four clusters are observed.

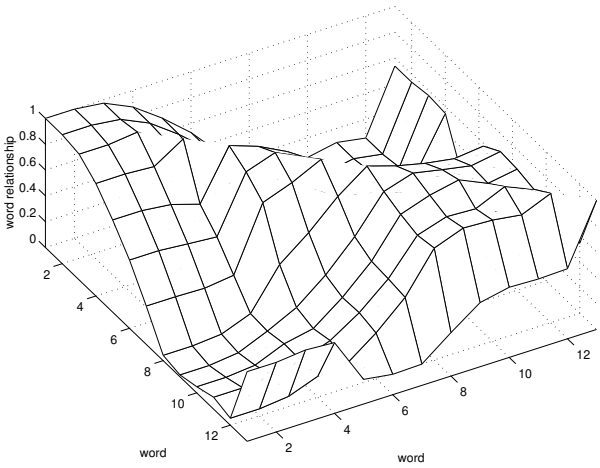

Figure 3: Word relationships$R(w_i, w_j)$ shown in Fig.2 is directly drawn in this graph. Words are listed in X and Y axes. The number for each word is corresponding to that of in Fig.2. Z axis is $R(w_i, w_j)$.

words have relatively closer relation than with other words, but words at the both ends are not so close. Namely relationship among words in such cluster gradually change from end to end. This type of cluster has a gradual decay of relation as shown in Fig.3, it is named a gradual cluster. These two types of cluster structure are typical in our simulation.

Around the beginning of simulations a few flat clusters exist rather independently. Flat clusters extend their boundaries by new words. Some sentences connect these clusters. For instance, there are two flat clusters at the end of the previous text ($t = 2$). They have almost no relation with each other. One cluster develops its boundary and becomes to have a flat top and a gradual edge as the cluster $1 \sim 4$. These two flat clusters become to have relation by some new sentences connecting them. At the end of text $t = 3$, the cluster structure grows as in Fig.2 and 3.

4.2 Dynamics of Word Relationship and Cluster Structure

Dynamics of word relationship and changes of structure of clusters are studied in this subsection. An example of dynamics of word relationship is shown in Fig. 4. This is data of a word `aa` ($R(\texttt{aa}, w)$) in the same simulation reported in the above subsection. Each line is sustained in a value with some fluctuations for some texts and sometimes shows abrupt and large changes. At $t = 21$, simultaneous and large changes occur in relationship with many words. Some words with high value in $R(\texttt{aa}, w)$ become lower and some words with small value rapidly develop their relation with the word `aa`. Such dynamics, taking place of word relationship, are sometimes observed in many word's relationship and in a wide range of α^w and α^s. Since the nonlinearity of the similarity term in $R(w_i, w_j)$ (the second term in Eq. (1)) is weaker than that of the correlation term, the taking place dynamics in $R(w_i, w_j)$ are rare if α^x is too large.

Let us see this simultaneous and large change more precisely. We describe the dynamics of $R(\texttt{aa}, w)$ per each utterance/acceptance of sen-

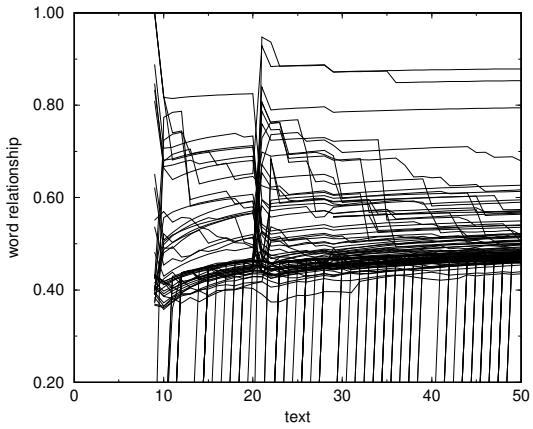

Figure 4: An example of transition of word relationship per each text is depicted. This graph shows dynamics of relationship of the word aa with the other words (Y axis: $R(\mathtt{aa}, w)$) at the end of each text (X axis) for an agent in the same simulation with Fig.2 and 3. Since the word aa appears at $t = 9$, the graph starts from that text. Lines arising from X axis indicate appearances of new words. Each line often fluctuates. Relationship with many words shows simultaneous and large changes at $t = 21$.

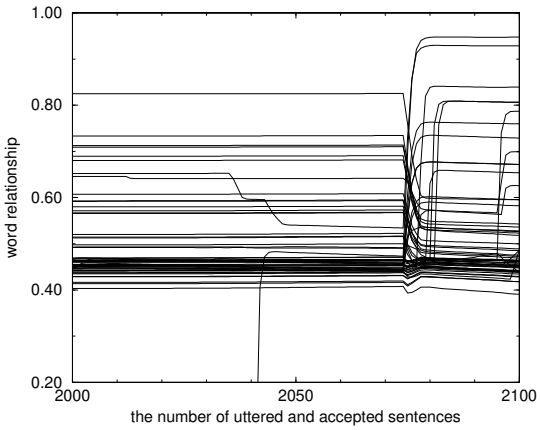

Figure 5: Transition of word relationship of the word aa with the other words per each uttered/accepted sentence in the text $t = 21$ is depicted. X axis is the number of uttered/accepted sentences (n) and Y axis is the word relationship $R(\mathtt{aa}, w)$. Between $n = 2070$ and $n = 2080$, many words switch their relationship with the word aa rapidly and simultaneously.

tence in Fig.5. The simultaneous and large change is observed between $n = 2070$ and $n = 2080$. It is clear from this figure that this change is not caused by an appearance of new word.

Uttering a new sentence "aa-o-o" at $n = 2075$ induces this taking places of relationship. The word aa is not a new word. But it is not used with a word o for a long time (used only two times as "aa-o" in the text 9 and "o-o-aa" in the text 10). A rare or new usage may change relationship of a word with other words largely. Words used in this text ($t = 21$) develop their relationship with the word aa and relationship with words used in other texts but not used in this text deduces. Ubiquitously used words don't change by appearance of this rare way of use.

We have reported here a rapid change by not new but rare usage of word. Rapid changes by appearance of new words are, of course, observed. There are cases where taking places of word relationship progresses more slowly.

To see transition of cluster structure by the large change of word relationship, we inquire how words move in a scatter diagram. The sentence-word affinity matrices $A(s_i, w_j)$ at $n = 2074$ and $n = 2077$ (162 sentences $\times$ 66 words) are combined to one matrix ($162 \times 66 \cdot 2$) and it is processed by the principal component analysis (PCA). The first and second principal components are used to draw the scattered diagram Fig.6. Corresponding words between two sentences are connected by arrows. Although sentence-word affinity $A(s_i, w_j)$ is not completely the same with the word relationship $R(w_i, w_j)$, rough structure of clustering shown in scatter diagrams of $R(w_i, w_j)$ is kept in those of $A(s_i, w_j)$ processed by PCA.

From this diagram, it can be seen that words in a cluster moves in the almost same directions before and after this change, so the whole structure of clusters does not change largely. But the

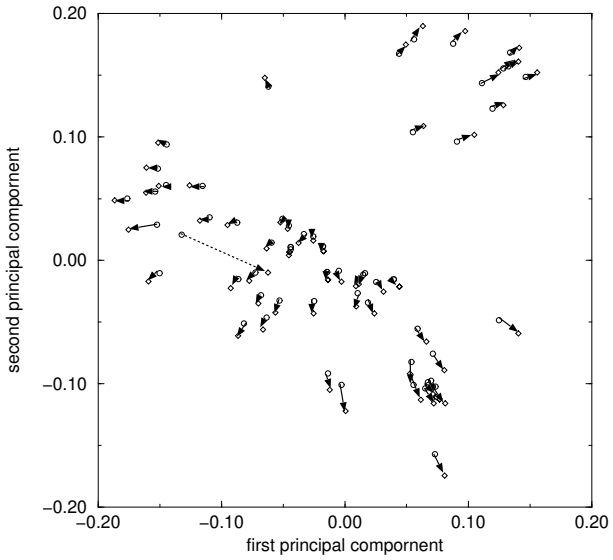

Figure 6: This figure shows dynamics of cluster structure between the number of uttered/accepted sentences $n = 2074$ and $n = 2077$. It is the scattered diagram obtained from the principal component analysis of a matrix made by combining the sentence-word affinity matrix $A(s_i, w_j)$ at both uttered/accepted sentences. X and Y axes are the first and second principal components, respectively. Symbols $\circ$ and $\diamond$ mark words at $n = 2074$ and $n = 2077$, respectively. The corresponding words between two sentences are connected by solid arrows. The broken arrow means the move of the word `aa`. We can observe that the word moves in different direction with the other words and changes its belonging cluster after utterance of the 2077th. sentence.

word `aa` moves in the different way with words in the cluster at $n = 2074$ and it changes its belonging cluster by this taking places of word similarity. Large changes in word relationship of a word does not restructure the whole system of clusters, however this change effects on word relationship step by step and also on uttering sentences. Thus the entire clustering structure slowly changes.

5 Summary and Discussion

We have proposed a viewpoint that word meaning is represented by relationship among words and it should be based on usage of language. By defining an algorithm to evaluate word relationship in keeping with this viewpoint, dynamics of clustering of words through conversations are shown in a constructive model of artificial agents with the word relationship matrices as their internal structure. Different with the previous work (Hashimoto, 1997), clusters do not fall into fixed and uniform structure by introducing a concept of text as an upper level of sentences.

The clustering can be regarded as categorization of words by the agent through conversation, since words in a cluster have stronger relation with each other and weaker relation with words out of the cluster. There are two major types of cluster structure, flat and gradual. A boundary of flat cluster is so sharp that it is clear whether an entity is a member of the cluster or not. This type of cluster is like a category in which membership is rigidly determined by necessary and sufficient conditions as scientific notions. In contrast to the flat cluster, a gradual cluster has graded change in relationship from large to small. In actual simulations these two types of cluster usually present in combination. Namely, a cluster is likely to have flat top and graded boundary. If we think of is as a category, a flat top part corresponds to central members of the category and words having small relation with the central ones are peripheral members. This structure is like a *prototype category* (Lakoff, 1987; Taylor, 1995). To what extent words are included in a category is a matter of gradient.

The structure of cluster has characteristics of dynamical stability and adaptability. By uttering or accepting sentences in which usage of words are new or rare, directly concerned words move their positions in clustering structure, however the whole structure is not suddenly transformed but slowly changes.

How cluster structure links with situations and topics, and what structure and its dynamics are seen in a global level, such as common clustering structure and common names for topics, have not reported here because of shortage of space. These topics are discussed in

(Hashimoto, 1998).

Acknowledgments

I am grateful to Mr. Iba for his advice about statistical analysis. This work is supported by Special Postdoctoral Researchers Program at Institute of Physical and Chemical Research (RIKEN).

References

N. Chomsky. 1957. *Syntactic Structure*. Mouton, The Hague.

D. A. Cruse. 1986. *Lexical Semantics*. Cambridge University Press, Cambridge.

D. Geeraerts. 1985. Cognitive restrictions on the structure of semantic change. In Fisiak J, editor, *Historical Semantics*, pages 127–153. Mouton de Gruyter, Berlin.

A. D. Gordon. 1981. *Classification*. Chapman and Hall, London.

H. P. Grice. 1975. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics, Vol. 3: Speech Acts*, pages 41–58. Academic Press, New York, NY.

T. Hashimoto. 1997. Usage-based structuralization of relationships between words. In Phil Husbands and Inman Harvey, editors, *Fourth European Conference on Artificial Life*, pages 483–492, Cambridge, MA. MIT Press.

T. Hashimoto. 1998. To appear in Multi-Agent Systems and Agent-Based Simulation (MABS98).

K. Kaneko and I. Tsuda. 1994. Constructive complexity and artificial reality: an introduction. *Physica*, D75:1–10.

K. Kaneko and I. Tsuda. 1997. *Fukuzatsu-kei no Kaosu-teki Sinario (in Japanese)*. Asakura Shoten, Tokyo.

Y. Karov and S. Edelman. 1996. Similarity-based word sense disambiguation. Technical Report CS-TR 96-06, Weizmann Institute.

G. Lakoff and M. Johnson. 1980. *Metaphors We Live By*. The University of Chicago Press, Chicago.

G. Lakoff. 1987. *Women, Fire, and Dangerous Things*. The University of Chicago Press, Chicago.

D. Sperber and D. Wilson. 1986. *Relevance: Communication and Cognition*. Blackwell, Oxford.

L. Steels. 1997. The synthetic modeling of language origin. *Evolution of Communication*, 1(1):1–34.

J. R. Taylor. 1995. *Linguistic Categorization – Prototypes in Linguistic Theory*. Oxford University Press, Oxford.

L. Wittgenstein. 1953. *Philosophische Untersuchungen*. Basil Blackwell.