The Constructive Process Theory and the Multi-Level Translation Method

NTT Information Network Systems Laboratories

Satoru IKEHARA

List of Abbreviation

L	:Language
E	:English
J	:Japanese
JL	:Japanese Language
pre-E	:pre-editing
J-to-E	:Japanese to English
CPT	:Constructive Process Theory
NL	:natural language
NLP	:natural language processing
CLs	:computational linguistics
SWD	:semantic word dictionary
SSD	:semantic structure dictionary
pvt-M	:pivot method
trs-M	:transfer method
MP	:meaning processing
MA	:meaning analysis
MU	:meaning understanding
SE	:subjecyive expression
OE	:objective expression

挨拶

○ Thank you very much, Mr. Chairman, for your kind words of introduction.

○ I am also thankful to Professor Rigsby and the Organizing Committee▼for the opportunity to visit this beautiful country▽and to get to know my Australian and other colleagues better.

○ I am particularly honored to be here▽to speak at this Anniversary of your Society.

● Today I would like to speak about CPT and MLT method.

自己紹介

① I think it would be appropriate to start by some introductory remarks about myself▼and the research efforts exerted on NLP by our Laboratories.

② Almost 20 years have elapsed▽since I joined the Electrical Communication Laboratories at NTT.

・ The first 10 years were spent in research work on▼Form. Al.Mani. and traffic theory, with emphasis on QN.

・ It is therefore 10 years▽since I started research work on NLP.

研究紹介

① NTT currently has 13 laboratories.

② Research on CLs have been conducted for long time▼at basic research laboratories in Musasino city.

③ Research on NLP were started▽at the Laboratory in Yokosuka 10 years ago.

・ I was involved in this research effort from the very beginning.

・ From a staff of three in the beginning,▼the number of researchers has grown to over 30 persons.

④ In the area of NLP,▽I have undertaken research efforts in▼
- J text to speech,
- key word extraction,
- error detection and correction of J text,
- and statistical approaches for voice recognition.

・ This research effort continues on today.

⑤ I am now engaged mainly in J-to-E MT.

言語処理と計算言語学

① While research work on CLs and NLP resemble each other,▼they are not the same and it would be adviseable to handle them separately.

② My current research efforts have been concentrated on NLP and not on CLs.

・ It is not research on how a NL can be transcribed by a theoretical model.

・ It is research on how can J, that is actually in everyday use, can be translated into understandable E.

・ My research efforts are directed toward realization of this method.

③ Upon starting research work on MT,▼a major problem was a considerable gap between CLs and NLP.

④ It can be stated that▽if we apply the results of CLs to actual every day words,▼there arise many exceptional cases and appropriate processing is limited.

X It may be that if CLs would achieve further progress and development, these problems would cease to exist.

X But can we afford to bide our time and wait for such an event before proceeding any further with our MT system ?

X Our Laboratory is an institution engaged in information processing.

X In conducting MT research, regardless of the method, the final objective is to make possible the translation of actual words.

X To await the development of CLs, would make it difficult to start MT research.

X We would be denied the budget for this purpose.

X It was no easy task for us to obtain approvals for conducting research on MT.

自然言語とは

① NL is something like a custom or habit▼that has spontaneously grown up within its own group of the L.

・ It reflects the very human society▽in which it was born.

② It is no easy task to explain each individual phenomenon▼without conflict by a unified model.

③ If Ls are a social custom, there should be a method of processing that matches them.

④ Researches of NLP require the pursuit▼of every one of these words.

我々の考え方

① I have decided to begin research work▼from the viewpoint of NLP and not CLs.

② Accordingly,▽I have sought a new approach amidst the results of conventinonal linguistics,▼promoting a new methodology.

③ The background for my research work is▼the CPT of L by Motoki Tokieda,▼who followed up on the work▽initiated in the 18th century by Norinaga Motoori.

・ Our MT research efforts have been based on this theory▼and has resulted in a MLT method▼and a test product J-to-E experimental MT system.

話の計画

① Today,▼I would like initially to introduce the outline and experimental results▼of our MT system in the half of my talk.

② In the latter half,▼I propose by describing our thinking and methods▽making comparisons with the thinking behind CLs.

Ｆｉｇ．１ 日本の現状

本論開始

① Let us start with the relation between our puporse and the ccurrent methods in Japan.

ＭＴ進歩

② As you are aware,▽MT methods have advanced▼and many commercial systems have appeared recently in Japan.

③ Our major concern, at this stage, is J-to-E translation.

日英違い

④ JL is very different from European Ls.

・ The quality of J-to-E translation is not good▼compared to French to E or E to Russian translations.

・ Current methods can be classified as a word-to-word translation,▼or literal translation.

明治の変革

⑥ JL itself has been enhanced by importation of European culture in the M-Era.

・ Therefore, JL has certain expression patterns▼that are easily translatable into European Ls.

・ These parts can be translated by current methods.

前編集

○ However,▽conventional J expressions are required to be rewritten into easily translatable J.

⑤ That is, pre-E is inevitable▼in order to apply the current MT system to practical use.

ＮＴＴの目的

⑦ The purpose of our research is to realize communication services with translation. ・ We hope to realize direct communications.

・ However,▼current methods are not satisfactory in realizing such services▼ because of the need for pre-E.

⑧ Thus, our aim is to achieve translation without pre-E.

限界

⑨ Let's look at the manner in which pre-E ia conducted▽to know about the limitations of current methods and their causes.

Ｆｉｇ．２ 前編集の項目

原文の書き方

① How should original sentences be rewritten before translation ?

② Let us look at the examples in this figure.

・ The four essential points of pre-E are shown.

単語訳し分け

③ First, every word should have only one meaning.

・ This example shows the J verb 「kakeru」▼having, in this case, six different meanings,▽pour, place, build, sit, mop, and play.

・ Such J words need to be rewritten into other words▼which correspond to a E word,▼but there are cases which present difficulties in rewriting.

形容詞の例

④ Second, structures of sentences and expressions should be simple.

・ In this example, the adjective 「utsukushii」 meaning beautiful,▼is placed in front of the pronoun (I, me or my).

・ This can resault in two interpretations,▼"beautiful me" and "beautiful daughter".

・ To avoid these ambiguities,▼modifiers should be placed immediately in front of the word to be modified,▼as in this rewritten example.

省略と補完

⑤ Third, supplementation of elliptical subjects and objects will be required.

慣用隠喩

⑥ Fourth, idiomatic expressions and metaphores need to be rewritten into other expressions.

結論

⑦ It can be stated that▼these requirements have derived from limitations of literal translations.

Ｆｉｇ．３ ２大翻訳方式

２方式

① Let's consider the translation methods.

② Conventional translation methods can be classified into two types, the pvt-M and the trs-M.

③ Among researchers of MT, there have been arguments as to which method is superior.

ピボット志向

○ Among the translation systems developed by private companies, there are many which seek to achieve the pvt-M.

・ However, there have been none that can be stated to have truly realized the pvt-M.

・ This is because universal intermediate L that can be commonly applied to many L cannot be designed.

言語とは

○ NL has primarily originated by reflecting the perpsectives and thinking of the group of people using such L.

・ It would appear unreasonable to consider a intermediate L that would be commonly applicable to every L which differs in thinking and perspective.

ピボット

○ When a translation between Ls which are very close each other is considered, or if a rough ranslation is to suffice, the pvt-M may be satisfactory.

・ But in either case, I think pvt-M has a basic misunderstanding of Ls in its background.

トランスファー

① The trs-M can be stated as being more realistic compared to the pvt-M, because it does not assume that intermediate L is universal.

② However, if intermediate L can be thought of as a meaning expression such as deep structure, the same difficulties as in the case of the pvt-M would arise.

③ The meanings of surface structure cannot necessarily be retained by deep structure.

Ｆ ｉ ｇ ． ４ 人間の翻訳

人手翻訳

○ Let us consider the manual translation method.

表現と認識

① In the expression that is to be translated, the state of the objects as recognized through the eyes of the speaker is tied with the expressions.

・ And at the sane time, the awareness of the speaker toward the object is also tied with the expression.

② To tie the speaker's recognition with the expression, a rule of words, i.e. linguistic norm, is used.

・ This norm differs with every L, such as in E and in J.

人手翻訳

○ A human translator experiences the speaker's recognition through the expressions.

・ He is required to know both the state of object and speaker's recognition thereof.

・ These contents are expressed in the framework of the target L▼on behalf of the speaker.

図式

① This figure shows▽what is objective in this path and the what is subjective in this path.

② Thus,▼translation is conducted through a re-grasping of the speaker recognition.

③ A simulation of this process is the MLT method which we proposed.

Ｆｉｇ．５ 多段翻訳方式

多段翻訳方式

① This figure shows the MLT method.

② The source L is J, and the target L is E.

③ The term "Multi-Level" indicates that▼translation is being conducted at various levels of abstraction.

パスの説明

① This system is divided up into 2 major conversion paths.

② The one is the conversion of the speaker's sense or what is subjective.

③ The other is the path for conversion of the description for the object.

・ And this path splits up into 3 more paths,▼depending on the abstraction degree in conceptualizing objects.

④ Thus, the MLT method has currently four translation paths.

従来との違い

○ The major difference between this method and the conventional methods is▼whether the speaker's recognition and the object are regarded separately or not.

Ｆｉｇ．６ 言語知識

知識が本質

○ To realize a MT system, a knowledge of L is essential.

曖昧さとの戦い

○ NLP can be stated as being a battle with ambiguity▼from the beginning to end.

・ In the translation process,▼various types of ambiguities arise,▼but this is regarded as being a lack of knowledge that is required.

知識

① To overcome such ambiguity, it is necessary to study what knowledge is lacking and to have such knowledge established as the rule or as a dictionary.

② It is important to consider the type of ambiguity and relationship▼with the knowledge corresponding to it.

方針

○ In our MT system,▼we place importance on knowledge for MP▼for the purpose of solving ambiguity of sentence structure▼and of the meaning of individual words.

我々の方法

○ We have collected 2 major types of information▼and have compiled them into a form of dictionary▽as shown in the figure.

概念化

① When the speaker expresses an object, abstraction is conducted and the object is conceptualized.

② Details will be mentioned later, but in L expressions, the concept of object is expressed separately in terms of a concept of substance and a concept of attribute.

辞書化

○ Linguistic conventions which combine concepts to expressions have therefore been compiled into 2 separate dictionaries.

単語意味辞書

○ The one is knowledge related to semantic use of words.

・ This has been compiled in a SWDic.

構文意味辞書

① The other is knowledge related to meanings of expression structures.

② The meanings of expressions are not necessarily represented by the sum of the meaning of every word in the expression.

③ In fact, it would appear thet the meanings of linguistic expressions actually used cannot be explained solely by the meaning of each word used in the expression.

④ I believe that▼expression structures need to be considered▽as units of meanings in NL.

⑤ Based on such thinking, the abstraction of expression structures centered around declinable words and compiled as units of meanings▼has resulted in the SSD.

⑥ This is one of the important points asserted in our system▼and I shall be dealing with this factor later in this presentation.

結論

○ I would like to state clearly that▼the knowledge we have discussed here is a L knowledge▼and not a worldwide knowledge such as general commonn sense nor some knowledge of specialists.

Ｆｉｇ．７ 辞書

要約

○ As discussed in the foregoing,▼we have compiled a SWD and a SSD.

意味属性体系

○ The meanings of words and their translation are listed in ordinary dictionaries.

・ But it does not serve to have the computer know▽how they are to be used.

・ For example,▼the word ″school″ is used as a ″place″▽and also as an ″organization″.

○ As a means of desscribing these knowledge,▼we have compiled the use of the meaning of words as a SA system.

知識の精度

① It is important to consider precision of knowledge description▼in relation to the degree of ambiguity.

② For example,▽an algorithm which was valid in a world of 1,000 words▼may not necessarily be valid in a world of 100,000 words.

・ Experiences indicate that▼increase of the number of words causes rapid growth of ambiguities.

③ It is a rule of science that▼quantity gives rise to changes in quality.

④ We need to keep this rule in mind, especially in NLP.

単語数

① Seeking to establish practical method, we have collected words including proper nouns normally used day to day and have compiled a dictionary of about 400,000 words for our system.

属性数

① Thus, it is necessary for SA system to work with this environment.

② We began with about 500 categories.

・ But this has resulted in that translation of verbs cannot be precisely described.

③ With the consideration of our aim and environment,▼we set up 2,800 categories for general noun SAs▼and 200 categories for proper noun SAs.

辞書項目

① Working with these SAs,▼we have compiled a SWD consisting of 400,000 words▼and a SSD consising of 15,000 sentence forms.

② The SSD is in a Valentz pattern form▼and the index words are verbs and adjectives.

Ｆ ｉ ｇ ． ８ 一般名詞意味属性

一般名詞

① This figure shows the nodes▽for the top 4 levels▼of the SA system for general nouns.

・ General nouns are branched off into concrete and abstract,▼with concrete branching off into subjects, places, and concrete objects.

② There is a total of 2,800 category names.

③ Here, I have taken one example at the deepest level.

・ The depth of category amounts to 12 levels at the deepest level.

Ｆ ｉ ｇ ． ９ 固有名詞意味属性

概要

① Next we have the SA system for proper nouns.

② Attributes for proper nouns are divided into 4 types,▼the names of persons, of places, of organizations and other.

・ These are sub-divided further.

×人名と地名

X Names of persons will stop at about the 4th level, but names of places go further reflecting on the difference in use of names of persons as opposed to places in J sentences.

X Names of places are used in the form of "nesting" such as prefecture country, ,county, towns and aza (local sub-division) thus accounting for depth of categorization.

ノードの深さ

① The number of nodes amounts to 200 and the maximum depth is 9 level.

② An example of the deepest node is shown here.

Ｆｉｇ．１０ 単語意味辞書の形式

導入

① Now, take a look at how these SAs are located in the SWD.

② This figure shows the contents of one record within the SWD.

レコード形式

① The index word is "Tokyo".

② One record consists of 350 Bytes and some 200 to 300 various types of information have been input into a record.

③ SAs of words have been registered in the final 15 fields.

属性付与レベル

① In the SWD, each word has been given a lower level of attribute.

② Conversely, with the SSD, higher level attributes have been used.

③ This is due to consideration for retaining the relationship between higher and lower level attributes.

・ Expression structures should be defined by generalized format as far as possible.

Ｆ ｉ ｇ ． １ １ 属性体系比較

比較

○ At this point,▼we would like to see the differences of description capabilities▼between of our SA systems and other similar systems.

ケース１

○ Case 1 is a typical example of a J-to-E MT system which has recently been commercialized.

・ The number of semantic categories are between 30 to 50.

ケース２

○ Case 2 is an example of the dictionary of word concept▼being planned by the Japan Electronic Dictionary Research Institute (EDR)▼established by MITI (Ministry of International Trade and Industry) of Japan.

ケース３

○ Case 3 is an example of our translation system ALT-J/E.

比較方法

① A comparison was made of the three cases.

・ And, the capabilities of describing the translation rules for verbs are evaluated.

② In the comparison, the case of ALT-J/E has been assumed to be 100%.

③ The dictionary in the case of EDR is just in the planning stage so that the values for Case 2 have been obtained from experiments conducted in our system based on similar conditions.

結果

○ The results of this comparison are as follows.

○ The capabilities of description was 31% for case 1 and 59% for case 2.

① Those results indicate that▼when precision levels of SAs are low, essential rules cannot be written.

② Our experiences also indicates that,▼in J-to-E MT,▼precision level for some 3,000 types of SAs are necessary▼for differentiation of translation for verbs.

③ However, this is not sufficient for translation of nouns.

・ I think that some other different system from the SA is required for differentiation of translation in the case of nouns.

Ｆｉｇ．１２ 単語訳し分け例

導入

○ Let me show▽what kind of translation has become possible by our system.

「掛ける」の例

① The first example is differentiation in the translation of the J verb「kakeru」.

・ This is the same example as shown in the beginning of Pre-E.

② The verb in the J sentences are all the same「kakeru」.

・ But, the translation by ALT-J/E shows▼poured, made, caused and other different words.

③ J verbs in general▽need to be translated into many different E words.

④ In the case of「kakeru」,▼there are about 15 different ways of translation for general use▼and to include idiomatic usage,▽there will be a need to consider some 80 types of translations.

「する」

○ The verb with the largest number of variations would be「suru」▼with over 300 different ways of use.

・ This verb resembles the E verb "do".

「群れ」の例

① The second example shows differentiation in the translation of the noun「mure」.

② The same word「mure」is used 8 times in the J text▼but the translations are different for each case.

③ In this case,▼the E word varies depending on what is referred to.

X According to a native E speaker who was interviewed, the word "bevy" here is used only in the case of beautiful women.

X It was also pointed out that "mod" sounds strange.

X It would appear that we need to change our rules somewhat.

X Here, it would suffice if our audience can appreciate that differentiation needed in the translation of nouns has also progressed to certain level.

Ｆ ｉ ｇ ． １ ３　自動書換と補完例

例題

○ The next is an example of an automatic rewriting function of J sentences.

第１の例：自動書換

① Here is a proper J sentence.

② The literal translation of this sentence is= (英語を読む)

・ I think this E may be understood,▽but it does not sound like proper E.

② The reason for this can be seen clearly▼as having excessive number of verbs.

日英の違い

○ J is a L of ″circumstantial logic″,▼and it necessarily uses many verbs.

・ In contrast, E is a L of ″substantial logic″▼and is said to prefer structural expressions.

動詞の削減

① Thus, to reduce the verbs,▼a partial rewriting of the J text in the figure is undertaken.

② This rewriting is performed by the translation system automatically.

③ After automatic partial rewriting,▼we get the translation as = (英語を読む)

・ I think the quality of E has improved.

×本来の目標

×① It goes without saying that it is our goal to think through J texts that are difficult to translate and to translate these into E that is readily understandable.

×② Yet it is another alternative to change Japanese text that is difficult into one that is easily translatable automatically.

×③ This method is particularly convenient when a translation function that is already completed can be applied without having to establish the function anew.

＝＝続 く ＝＝

第２の例：補完

○ The next example is supplementation of subjects and objects.

日本語の習性

① The J convention is to refrain from mentioning▼what is already known to the listener.

② Lengthy, redundant expressions are not popular.

③ Particularly, the subject and object are omitted unless absolutely necessary.

・ Because∇these can generally be judged by the partner▼by joshi and auxiliary verb expressions.

英語の習性

① In contrast,▼in E,∇both subject and object are generally necessary.

② There are cases▼when the subject is missing and the object is turned into the subject▼and rewriting into the passive voice.

③ But this is not always appropriate.

補完例文

① Here are two sentences which are linked to one another. (日本語を読む)

② The first sentence has the subject and object,▼but the second sentence has two subjects omitted.

③ Therefore,▼ALT-J/E has supplemented the first subject by the object of the first sentence.

④ And the latter subject is supplied by the subject of the first sentence.

まとめ

○ I have shown some of the new functions that have been realized by ALT-J/E.

結論

○ I believe that,▼with the realization of these new functions,▼the possibility of MT without pre-E▼is now real possibility.

－－

前半のまとめ

① The foregoing is the initial portion of my presentation.

・ I have introduced the MLT method and our MT system ALT-J/E.

② In the latter half of my presentation,▼I would like to mention details of our conceptual background.

③ If there should be any questions on the discussions up to this point, I shall be happy to answer them.

後半の開始

① I shall begin the latter portion of my presentation.

② I have already mentioned that▼our research efforts are being conducted from the viewpoint of NLP, not CLs.

③ Let me discuss the thinking behind our MT system.

背景

① The background to our research lies in the CPT of L.

② This theory was advocated some 50 years ago▼by a J linguist, Motoki Tokieda.

③ Tokieda took after the J grammar▼which was established some 200 years ago by Norinaga Motoori,▼and introduced the CPT from a standpoint of▼criticizing the philology of Ferdinand de Saussure.

話の順序

① First, I wish to present the differences▼between the concepts of CLs and of CPT.

② I shall follow by discussing▽how the CPT is related to our MLT method.

Ｆ ｉ ｇ ． １ ４　論理学の系統

本論開始

① Let's consider the background of the CLs, first.

② In considering logics,▽there has been two types of logics since the olden times.

弁証法論理学

① The one is dialectics.

・ Dialectics was advanced by Hegel.

② This type of logics asserts that▼actual existence should be thought of as a meaning theory.

・ Hence, it deals with contents.

形式論理学

① The other type is formal logics.

・ This type of logics has appeared in the era of ancient Greece.

② This logics insisted on purity and handling of an ideal world.

・ And it cut itself off▽from the realistic world.

記号／計算論理学

○ Symbolic logics and computational logics are derived from the latter, formal logics.

自然言語

① There appears an effort to apply CLs to NL.

② But NL is a L▽that deals with the realistic world.

③ This would seem to be the cause of various problems.

・ For example,▼Frege's principle, which is compositional semantics and constitutes a major assumption in CLs,▼is not valid in a NL.

----------Fig.15

Ｆｉｇ．１５ 日本語と欧米語の違い

① Let us look at the differences between the J and Ind-European Ls.

欧米語

② The Ind-European Ls have sentence patterns such as 5 patterns▼which are relatively independent from contents.

・ These Ls can be classified as a suitcase type L.

・ Structualism was developed▼from the viewpoint of importance of forms and structures of expressions.

・ Chomsky took notice of the importance of contents to overcome the limitation of structualism.

日本語

③ In contrast,▼JL is relatively free from forms▽and does not have a rigid sentence structure.

・ J can be classified as a 「furoshiki」(wrapping cloth) type L.

・ Forms are dependent on contents.

・ In other words, the shape of contents appears in a form of expression.

④ Thus, contents are important in J processing.

日本語文法

⑤ In J, there are 20 different types of grammars.

・ And four of them are well known.

○ Here, I would like to take up two▽out of the four.

・ One is Hashimoto grammar.

・ This grammar is strongly influenced by European philology.

・ And this grammar was adopted for use in school▽by the Ministry of Education of Japan.

・ Another grammar is Tokieda grammar▽which was derived from Norinaga Motoori.

時枝文法の考え

⑥ As nature is a complexed body of processes▼that develops with conflicts as the motivating force,▼so also is L.

⑦ Tokieda established the CPT of L▽based on this idea.

---------- Fig. 16

Ｆｉｇ．１６ 言語過程説とチョムスキーの違い

① This figure shows the differences▼between the CPT and Chomsky's concept.

言語過程説

② CPT asserts that▼a L should be considered as a complex body of processes▼ comprised of objects, speaker recognition and expressions.

③ Objects reflect recognition.

・ This process is explained by the reflection theory.

④ #Recgnts are combined into expressions using norms,▽namely the rule of Ls.

・ #Ling-norms are national rules▼which have been fostered in each community.

・ These norms are referred to as a grammar▼in the broad meaning of the word.

チョムスキー

⑤ In contrast,▼Chomsky interpreted L in the form of a dual structure,▼a surface structure and a deep structure.

⑥ The existence of deep structure has been assumed▼as the meanings of surface structure.

・ This is certainly a strange theory▽where because meanings are assumed for the purpose of explaining meanings.

・ It would further appear that there would be difficulty in explaining the existence of deep structure.

最近のチョムスキー

・ More recently,▼he appears to be attempting explanation based on mystic concepts.

・ He says▼L cannot be explained without assmuming that▼humans are born with a basic linguistic ability.

因果律

⑦ Well, results are achieved from causes▼and these results, in turn, become the causes for the next series of results.

自然と言語

・ Nature is a composition of such processes▼and it is the conflicts and discrepancies which activate these processes.

・Conflict is energy. === ・I believe this is the thinking followed by science.

知的産物

・Language is a spiritual by-product of humans.

・There is room for various theories to arise▼based on how the human spirit is to be considered.

---------- Fig. 17

Ｆｉｇ．１７ 対象世界の認識

○ Let us now see▽what happens▼when we regard L as a causal sequence▽between object, recognition and expression.

対象世界

○ The world of objects consists of a speaker and other objects.

・ Objects are composed from substances, attributes and relations.

③ The speaker recognizes the object world▼and at the same time, recognizes things about himself.

・ In recognition of objects,▽conceptualization takes place.

自己分裂

○ There are two occasions regarding recognition of self.

・One is the occasion where▼speaker represents himself directly in the expression without conceptualizing himself.

・On other ocassions,▼he conceptualizes himself and represent it as an object.

○ When he conceptualizes himself,▼the phenomena of self disunion take place.

・ A separate imaginary self other than the actual self appears.

・ And the self conceputualizes the actual self.

主体的｜客体的

○ Recognition related to subjects and objects is represented▼by two types of forms in linguistic expression.

・ One is a SE and the other is an OE.

④ SEs represent speaker emotions and intentions.

・ These expressions are expressed in the JL▼by post positionals and adverbs.

⑤ OEs represent conceptualized objects.

・ And these expressions are generally expressed by nouns and verbs.

英語の場合

⑥ This relation is somewhat different in E.

・ Inflections usually express SEs in E.

---------- Fig.18

Ｆｉｇ．１８ 主体的表現と客体的表現の例

① Let us take a look at this example.　　== 日英単語対を読む==

② This sentence means▼"He will also want to go to America".

③ "He, America, go" are OEs.

④ "also, to, want, will" are SEs.

ポールロワイヤル

・ A similar view is witnessed in the grammar of Port Royal in France▽300 years ago.

⑤ The difference is an important factor in the translation method.

・ I plan to mention this later in relation to the MLT method.

Ｆｉｇ．１９ 意味論

● Here, let us proceed to the problem of meanings.

① Recently, the importance of semantic processing in NLP is▼being strongly emphasized.

・ Much efforts have been devoted to semantic or meaning processing,▼but it is felt,▽in many cases, the definition of meaning is ambiguous.

○ I would like to point out here what we regard as meaning.

意味論の分類

② Let us consider the structure of Ls.

・ If we assume that meaning is a substance,▼it is one of either "object", "recognition" or "expression"▼and the fourth element of "interpretation".

形式と意味

・ If we say, at this point, that▼meaning is born out of forms,▼we would arrive at a "formal semantic theory".

・ But this would be ideological.

③ Next, if we say that forms are born out of contents,▼this would result in a "object semantic theory" or "recognition semantic theory".

対象と意味

④ If we assume that objects and recognitions are meanings,▼how should we explain the sentence that has been erroneously written.

・ This would result in a sentence that is incorrect▽in which something that is opposite is written,▼yet stating that the meaning is correct.

⑤ Further,▼the object and recognition can never remain the same indefinitely and will undergo changes.

・ Then, the meaning of the sentence will change independently from expressions.

時枝説

⑥ Motoki Tokieda chose eclecticism and took meanings to be listener's reactions.

三浦説

・ Tsutomu Miura took after the Tokieda grammar,▼but improved the theory and advocated the "relation meaning theory".

⑦ He proposed that▼structural processes from object to recognition are combined to make an expression.

・ He explained the meanings of linguistic expression as the relationship▼ between object, recogntion and expression.

⑧ This kind of relationship is concrete and specific.

・ As long as the expression exists,▼so will the meanings.

・ When expression is cancelled out,▽so will all relationships,▼and the meanings are also lost.

状況意味論

⑨ The concept of regarding relationships as meaning▼resembles recent situation semantics,▼but actually, the two are entirely different.

⑩ In situation semantics,▽the meanings of expressions▼and the meanings related to the location in which the expression is placed▼are completely mixed up.

・ The Miura grammar features a clear distinction▼between linguistic expression and expression of locations.

---------- Fig. 20

Ｆｉｇ．２０ 意味処理

○ How do we go about MP▼from the viewpoint of "relation meaning theory".

○ Meaning is the relationship between object, recognition and expression.

意味の要約

○ SE involves the emotions and intentions of the speaker,▼whereas OE expresses the status of the object▼as viewed from the eyes of the speaker.

・ Thus, MU would involve re-tracing such relationship▼and to relive and experience the status of object and recognition.

２つの意味処理

④ Here, let us think of MP in two steps.

意味解析

⑤ The first step would be the process▼which designates the rules or conventions used in linguistic expression.

⑥ Conventions regarding Ls are complex and liable to be construed in many ways.

・ The listener must identify the convention used by the speaker.

・ I would wish to define this act of identifying the convention as "MA".

意味理解

⑦ The second step would be the process of identifying the recognition of the speaker▼and the status of the object that is tied in with the expression.

・ Here, we would like to refer to this action as "MU".

まとめ

⑧ Therefore, MP consists of two processes,▼"MA" and "MU".

知識

⑨ The knowledge that is required for MA would be linguistic knowledge,▼that is, conventions or norms regarding Ls.

・ In contrast,▼MU would require knowing the status of objects▼and this would require general knowledge, that is, worldwide knowledge and knowledge in specialized fields.

⑩ Let us go over the next example. ---------- Fig. 21

Ｆｉｇ．２１ 意味解析例

① This diagram is an example of MA.　　=日本語と英語を読む=

上の文

○ The meaning of the sentence is ----------.

○ Let us think about the convention involved in these words.

高い

○ The meaning of the word ″takai″ is expensive, high, noble, loud among others▼ but here, it is used with the meaning ″expensive″.

油

○ ″Abura″ has various conventions as shown in the figure,▼but here the convention expressing the meaning of oil is used.

言語解析とは

○ Thus,▼MA is identification of the convention▼which the speaker actually used among numerous conventions∇that exist in the L.

・ MA cannot be achieved▼by merely staring at the words one by one.

下の文

① Let us look at the lower sentence.　==日本語を読む＝＝＝

② Look at this portion of the sentence.　背の高い meaning the ″back is high″.

・ In E, one word ″tall″ is used to convey this meaning.

・ But looking at back and high separately▼will not bring the concept of tall.

③ The same can be said of the other expression.

・ Sell and oil jointly means idle away one's time in this case.

④ Thus,▼conventions regarding words can never be decisive▼if they are thought out word by word.

まとめ

② I said that▼the meanings of expressions are the relationship between objects, recognition and expressions.

③ Structrures of objects reflect themselves into speaker recognition.

・ These structures are also combined to expressions.

構造と意味

① This means that▼the structure of an expressinon is a part of the meaning.

・ Structures cannot be separated from meanings because structures also represent meanings.

② Thus, in MA also,▼there is the need to think the relationship between structure and meanings.

・ More specifically,▼it is important to grasp the structure of expressions as units of meaning.

---------- Fig.22

Ｆｉｇ．　２２　意味理解

○ Next, let us consider about MU.

① As discussed in the foregoing,▼MU involves the identification of the speaker's recognition▼and the status of object as tied in with the expression.

理解

② To know the speaker's recognition and the status of object,▼the listener must be equipped with a certain world in his mind.

・ This world must have certain elements in common with the world▼which the speaker is depicting.

・ To restructure a world▽corresponding to the expressions of the speaker,▼he will be required to link up the elements of the speaker's expressions▼and the elements of his own world.

③ In other words,▼MU can be considered as linking the elements of the linguistic expressions▼and within the world model of the listener.

・ This action of linkage will structure a new portion in the listener's world.

知識

④ In contrast to liguistic knowledge for MA,▼MU will require worldwide knowledge.

⑤ Compared to linguistic knowledge,▽worldwide knowledge is massive▼and considerable amount of difficulties will be involved in research.

・ We are therefore conducting our reseach effort based on the following concepts.

---------- Fig. 23

Ｆｉｇ．２３ 意味解析とＭＴ

① We shall discuss MP in MT by dividing it into MA and MU.

意味解析とＭＴ

② The relationship between our MT research and MP is as follows.

知識

○ In MT, the contents of source text to be translated are diversified▼and it would be difficult indeed to prepare a worldwide knowledge.

ＭＴの方針

⑤ MT can be acceptable▼even if the computer cannot entirely understand the contents of the text▼as long as the translation results can be finally understood by humans.

⑥ Replacing the linguistic conventions used in the JL▽by E conventions▼and leaving it up to the judgment of the average person.

・ This is the basic stance assumed by our research efforts in MT.

ステップ

○ First,▼we want to realize a translation based on MA.

○ Next,▼we will extract the expressions and concepts▽which could not be translated by MA▽and attempt a MU with a limited world model.

意味理解は

○ In contrast,▼with areas such as telephone number information and database inquiries,▼the target world can be contracted into a comparatively small range.

・ We are structuring world models and conducting research in MU for such fields.

---------- Fig. 24

Ｆｉｇ．２４ ＭＬＴ方式概念

ＭＬＴの構成

① The method of translation based on the foregoing MA▼is the MLT method.

② I have previously mentioned the need to observe differences▼between SE and OE.

③ And, I have also pointed out the necessity of consideration▼about the meaning of structures.

④ The abovementioned translation method has been structured based on these two factors.

第１の点

⑤ The first point is considered as follows.

・ J and E are very different in presenting SEs.

・ Therefore,▽SEs are segregated from OEs▼and converted into E.

第２の点

⑥ The second point is resolved by abstraction of expression structures.

・ Three levels of abstraction have been proposed.

・ The first is related to a level of idiomatic expressions.

・ The second concerns specific structures such as loosely coupled words.

・ The last concerns most loosely coupled structure that can be represented by general rules.

主体／客体の扱い

⑦ From the foregoing,▽MLT method consists of two major structures.

○ From the viewpoint of linguistic knowledge,▼linguistic conventions are divided into those▼which are related to SEs and those related to OEs.

⑧ Of the two,▼the knowledge related to SEs are smaller in scale compared to OEs▼and have been solved in terms of rules or programming.

・ But the OEs being massive,▼we have decided to compile them in the form of a dictionary.

⑨ We shall next explain the method of summarizing linguistic knowledge▼related to objectve expressions.

---------- Fig. 25

Ｆｉｇ．２５ 言語知識

世界の構成

① I have already explained that▼the target world of linguistic expression is structured by the object▼consisting of substances, relations and attributes▼ and by the subject which is the speaker himself.

・ Here,▼the objectified speaker himself is identical with substances.

・ Thus,▼knowledge for OE can be considered as consisting of three types,▼ substances, relations and attributes.

名詞

② In terms of linguistic expression,▼substances and relations can be expressed by nouns.

・ Thus,▼the knowledge concerning use of nouns has been summarized in a SWD.

属性

③ Regarding attributes,▼there are a dynamic type and a static type▼which are represented by verbs and adjectives, respectively.

・ Thus, these have been summarized in a SSD.

まとめ

④ As mentioned in the earlier portion of my presentation,▼linguistic knowledge has been summarized based on these ideas.

×シソーラスとの違い

X As an item similar to the SWD, there is thesaurus.

X In E, the thesaurus by Roget is famous and there are several examples in Japan.

X These have perceived the similar$_{ity}$ between the meanings of individual words and have made classifications accordingly.

X Our word semantic dictionary features the viewpoint on which the word is based and used and has thus added a semantic role.

---------- Fig. 26

Ｆｉｇ．２６ ＮＴＴの現状

① I have presented an outline of the J to E MT system, ALT-J/E▼and the related conceptual background.

現状

② The current state of NLP research at our research laboratory▼can be summed up as follows.

形態素解析

③ First, morphological analysis▼which constitutes a considerable stumbling block in the JL.

・ I think the complexity is not less than that of syntax analysis or meaning analysis.

④ Technology involved has been almost completed.

・ A remarkably high level of precision and accuracy has been achieved.

構文解析

・ Syntax analysis and MA are being undertaken in the course of MT research.

・ High levels have been achieved,▼but I do not believe,▼it is satisfactory yet to be applied to newspaper article translation.

意味解析

④ Research on MU has been initiated in question answering.

・ This will require more time and effort.

---------- Fig. 27

Ｆ ｉ ｇ ． ２ ７ 今後の課題

① The main issues confronted by MT research▼and future research targets are as presented in the figure.

当面の課題

② Current subjects are . . . . ==No.1 を読む==

・ Our major research efforts are being concentrated in these areas.

将来の課題

③ Our future study efforts will be concentrated on . . . . ==No. 2を読む==

・ Preparations for these studies are now under way.

おわりの挨拶

○ Today,▼I have mentioned the thought and method of our MT system.

○ I am hoping that these ideas will become one of the bases of non-literal translation.

・ The problem of non-literal translation has begun to attract attention▼as witnessed by the fact that▼it was taken up at the IJCAI Workshop in Sydney last August.

○ If sharing our experiences in dealing with this system▼would be of interest to you,▼it would be our pleasure to do so.

○ This concludes my presentation,▽and once again,▼my heartfelt appreciation for this opportunity.

・ I thank you.

コメント

inscription → description ??

consider, concieve, perceiveの違い

図２０はここからスタート（４０分番組の時）

● Here, let us proceed to the problem of meanings.

・I can't discuss details of meaning theory here.

・However, we have adopted Miura's theory which assert meanings are relationship between object, recognition and expression.

・This theory is called "relation meaning theory".

⑨ The concept of regarding relationships as meaning▼resembles recent situation semantics,▼but actually, the two are entirely different.

⑩ In situation semantics,▽the meanings of expressions▼and the meanings related to the location in which the expression is placed▼are completely mixed up.

・The Miura grammar features a clear distinction▼between linguistic expression and expression of locations.

図２４の１文差し替え

I have not discussed the details of structural meaning of expression.

However, I must point out here the necessity of consideration about the meanings of structure as the second point of our translation method.