

# 基本的な文字化の原則(Basic Transcription System for Japanese: BTSJ)の英語への応用について

宇佐美まゆみ・鈴木卓

本稿は、作業段階のものであるため、BTSJをそのまま翻訳した部分と、英語の慣習に基づきBTSJに追加したり、BTSJから削除したりした部分がある。今後、検討を重ねていく予定であるが、以下に、現段階のものを掲載する。

## Basic Transcription System for English (BTSE): A Trial Version

### 1 Introduction

This document presents 'The Basic Transcription System for English: BTSE', a transcription system for English conversation developed as a component of a multilingual platform, 'The Basic Transcription System: BTS'.

BTSE follows the general orthography (e.g. the alphabet, spelling, capitalization, etc.) of English in order to enable easy and fast reading of data. As a rule, commas and periods are also used according to their conventional usages. (See '6.1 Transcription conventions' for details).

BTSE, which is optimized for use in quantitative analysis of natural conversation, adopts the 'discourse sentence' (discussed below) as the basic unit of analysis.

The rest of this document explains the following: Identifying a discourse sentence, starting a new line, symbols used at the end of a discourse sentence, numbers assigned to lines and discourse sentences, and the other transcription conventions/symbols.

### 2 Identification of discourse sentences

In order to facilitate the coding of sentence-level phenomena such as speech level shifts, BTSE adopts the term 'discourse sentence' which means a 'sentence in discourse', and uses it as the basic unit of analysis.

'A discourse sentence' is defined as 'a sentence in conversation as interaction' and can be identified in the following way. As a rule, 'a discourse sentence' corresponds to what is considered 'a sentence' in the conventional sense. However, in naturally occurring conversations, there are utterances that are not structurally complete as a sentence but may be considered as one. These include 'one-word sentences', elliptic sentences, and utterances that end in mid-sentence. Whether these utterances are to be judged as a discourse sentence or not is determined with consideration to turn-taking and pauses. Thus, turn-taking and pauses are two important elements for the identification of discourse sentences.

Following are some examples of an utterance to be treated as a discourse sentence if followed by a pause or another speaker's utterance: a one-word sentence (Example 1), an elliptic sentence (Example 2), a sentence unfinished due to the speaker's hesitation (Example 3), and a sentence unfinished due to interruption by another speaker (Example 4). A 'backchannel' or a laugh can also be judged to be a discourse sentence from consideration of turn-taking and pauses (Example 5).

On the other hand, not all the elements that are structurally complete as a sentence necessarily are judged to constitute a discourse sentence. For example, a filler used when the speaker tries to remember something, (e.g. 'Let's see, if you walk, it'll take about 12 minutes') is to be put together with the preceding/following elements to jointly

constitute a single discourse sentence (Example 6). Direct quotations are also treated likewise (Example 7). Identical 'sentences' repeated by the same speaker, as in 'That's right, that's right' (Example 8), also jointly constitute a single discourse sentence when pronounced with no pause between them and judged to have unity, even though each of these elements is syntactically complete as a sentence.

Even when an utterance produced by one speaker is interrupted by another speaker's utterance and divided into two separate lines, it is nevertheless treated as one discourse sentence as long as it constitutes a single sentence structurally. (See '3. Starting a new line').

In the following examples, an arrow (->) indicates the discourse sentence to which each heading refers. When the heading only refers to part of a discourse sentence, that element is underlined (Examples 6, 7, and 10). See '7. Key to transcription symbols' for the other transcription symbols used in the examples which are specially developed for BTSE.

Example 1: One-word sentence

Example 2: Elliptic sentence

Example 3: Sentence unfinished due to the speaker's hesitation

Example 4: Sentence interrupted by another participant's utterance and eventually ended in mid-sentence.

Example 5: Backchannels and laughs considered as a discourse sentence due to the existence of pauses before and/or after them.

Example 6: Structurally complete sentences not considered as a discourse sentence which jointly constitute a discourse sentence with preceding/following elements (An example of a filler: wavy line)

### 3 Starting a new line

As a rule, a new line is started each time speakers change. However, when one speaker produces multiple discourse sentences consecutively, a new line is started for each discourse sentence even though the speaker does not change (Example 11). Short backchannels and laughs which are produced in a low voice (e.g. 'Hm') and overlap another speaker's utterance are put in brackets and inserted within the other speaker's utterance at the location judged to be the closest point, except in cases when they carry a particular function other than to show that the participant is listening to the speaker (Example 12).

Example 11: Multiple discourse sentences produced by a single speaker

Example 12: Backchannels or laughs inserted within another speaker's utterance

### 4 Symbols indicating the end of a discourse sentence

When a second speaker starts his/her utterance while one speaker's discourse sentence is still in progress, the first speaker's discourse sentence will be divided into two lines by the second speaker's utterance, following the rule regarding starting new lines. Thus, a single discourse sentence sometimes extends over multiple lines. Therefore, one of the following symbols indicating whether the discourse sentence is finished or not is placed at the end of each line for clarity. Following are explanations of such symbols.

#### 4.1 Symbols used at the completion point of a discourse sentence

“.” (a period) is placed at the completion point of each discourse sentence. For an assertive sentence, a period is placed by itself. For a question or a confirmation, a question mark followed by a period is used, as in “?”. “?” is also used when it is clear from the intonation or the context that the utterance functions as a question or a confirmation, even if it does not have any of the formal clues normally associated with questions or confirmations such as “Wh-” or a tag question (e.g. ‘Sorry, your name is...?’). A period is always added after a question mark so that the number of periods will correspond to that of discourse sentences.

When a laugh is judged to carry a substantial function in the interaction and therefore to constitute a discourse sentence by itself, the sound is transcribed and a period is added after <laugh>, to show that it is a discourse sentence (Example 13).

When a sentence end is judged to be produced with a hesitation, ‘...’ is added before the comma (i.e. ‘...’) (Example 14).

Example 13: A discourse sentence consisting solely of a laugh.

Example 14: Hesitation at the end of a discourse sentence

#### 4.2 A discourse sentence extending over multiple lines

When a second speaker starts his/her utterance while one speaker's discourse sentence is still in progress, two commas are placed at the end of the first segment of the first speaker's discourse sentence to show that the sentence is in progress. The second speaker's utterance is transcribed in a new line. The second segment of the first speaker's discourse sentence is then continued in the next (i.e. third) line. One of the following symbols is added at the end of the discourse sentence: “,” “?”, or “...” (See ‘5.1 Discourse sentence numbers, Figure 1’)

A same discourse sentence number is given to segments of a single sentence to indicate that they make up one discourse sentence even though the segments are in different lines. (See ‘5.1 Discourse sentence numbers’)

Thus, one of the following (set of) symbols is always added at the end of a line: “,” “?”, “...”, or “,”.

#### 5 ‘Discourse sentence numbers’ and ‘line numbers’

In BTSE, the number of discourse sentences and that of lines capture information about conversation from different perspectives. To enable analysis from both perspective, two different numbers, a ‘discourse sentence number’ and a ‘line number’, are given to each line.

##### 5.1 Discourse sentence numbers

One number is assigned to each discourse sentence so that the number of discourse sentences will be easily recognized. When a single discourse sentence is broken into more than one line due to speaker change, the discourse sentence will extend over multiple lines. In this case, the same discourse sentence number is given to these lines to show that they all belong to a single discourse sentence. In addition, serial numbers are given to those sentence segments in order to indicate their position within the discourse sentence (e.g. ‘44-1’ ‘44-2’, ‘44-3’ in ‘Figure 1’ below).

In addition to the period (‘.’) placed at the end of a discourse sentence, an asterisk ‘\*’ is placed in the ‘Discourse sentence completion’ cell. The number of sentences will match that of periods (‘.’) and that of asterisks ‘\*’. Thus, sentence completion points are double checked to eliminate errors of overlooking them or counting the same sentence more than once.

##### 5.2 Line numbers

In BTSE, line numbers are used in addition to discourse sentence numbers. One number is assigned to each line so that an utterance produced earlier will always have a smaller line number. Discourse sentence numbers may

become ‘disordered’ in a transcript as in Figure 1 (In the discourse number cells in Figure 1, 44-2 is located after 45 and 44-3 after 46). Line numbers indicate the order of utterances as they are produced.

Figure 1

#### 6 Transcription conventions and symbols

This section introduces the principles of transcription in BTSE and the transcription symbols unique to BTSE. An explanation of transcription is provided first, followed by explanations on how to transcribe various kinds of information about speech, such as non-linguistic sounds, para-linguistic information, visual information, and personal information.

##### 6.1 Transcription conventions

In BTSE, the standard English orthography (the alphabet and spelling) is generally followed in order to facilitate reading. However, to record the original speech sound as accurately as possible, the following cases are specially treated. 1. Words pronounced exaggeratedly or with a marked stress are spelled with capital letters (e.g. ‘REALLY?’, ‘It’s BIG’) 2. Words pronounced differently from their standard pronunciation or with a dropped sound/syllable are transcribed as heard and then followed by their standard spelling (e.g. ‘I like im ‘him’’).

Single-digit numerals are spelled out. Arabic numbers are used for double or more digit numerals including those used in years, phone numbers and prices, etc. (e.g. ‘My mom was born in 1956.’)

Commas are used according to their conventional usage. A comma is also used to indicate a noticeable interval in an utterance even at locations where it is not customary to use a comma (Example 20). In addition, when a single discourse sentence is produced as an inverted sentence which is considered a nonstandard form, a comma is placed within the discourse sentence and a period at the end of it (Example 21).

Example 20:

Example 21:

##### 6.2 Speech sound information

To provide basic information about the context and situation of utterances, BTSE uses special symbols to record intonation, pauses, silence, latching, hesitation, overlaps, and unintelligible sounds.

###### a. Intonation

One of the following symbols is used to record intonation, only when doing so is likely to be necessary or helpful for analysis; [↑][→][↓] (Example 23). A double question mark (‘??’) is used to indicate a rising intonation used for confirmation of understanding etc. (Example 24).

Example 23:

Example 24:

###### b. Pauses and silences

When a slight pause is felt to exist in the tempo of speech, it is transcribed as ‘/int/’. When a pause lasts for one second or longer, it is transcribed as ‘/sil 1.5/’ (the numeral indicates the length of silence in seconds). If a silence functions as a reply, it is treated as one discourse sentence and assigned one line. Otherwise, the symbol for silence is placed at the beginning of the next speaker's line to indicate clearly which speaker breaks the silence.

Example 25:

c. Latching

When there is no interval between utterances, or the interval is relatively shorter than the average length of intervals in the particular conversation, we call it 'latching' and transcribe it with two '=' s. These symbols indicate that two utterances or discourse sentences, produced by one or more speakers, are assigned two separate lines in transcription but are produced continuously without much pause between them. A '=' is placed at the end of the preceding line, and another '=' is placed at the beginning of the following line.<sup>1</sup>

Example 26:

d. Hesitation

When a hesitation can be heard, whether in mid-sentence or at sentence end, it is transcribed as '...'. If it occurs in the middle of a discourse sentence, a comma is added to '...' (i.e. '...,') before the rest of the sentence.

Example 27.

e. Overlap

Utterances produced simultaneously are transcribed as follows: The portion of speech produced simultaneously with another speaker's speech is put in brackets ('< >') in each speaker's line. What is overlapped is then followed by '{<}', while what overlaps is followed by '{>}'.

Example 28:

In conversation, a speaker's discourse sentence can be interrupted by another speaker's speech, and sometimes ends at that point as a result. When this occurs, '[I]' is added to the end of the first speaker's utterance (before the period) to indicate that the discourse sentence is ended unintentionally. '[J]' is added to the beginning of the second speaker's utterance to indicate that it interrupts and results in ending the other speaker's discourse sentence.

Example 29:

f. Unintelligible sounds

Unintelligible sounds are recorded with #####. The number of # indicates the number of syllables the segment is assumed to include.

Example 30:

6.3 Para-linguistic information

Laughs and short backchannels which overlap another speaker's speech are transcribed as follows.

a. Short backchannels overlapping another speaker's speech

Short backchannels overlapping another speaker's speech which do not have a special meaning are put in brackets and placed within the other speaker's line at a point judged to be closest to the point of production.

Example 31:

b. Laughs

Laughs or utterances produced while the speaker is laughing are transcribed as '<laugh>', '<while laughing>', or '<both laugh>' etc. When the laugh can be clearly heard as verbal sounds, it is transcribed as such ('<Ha ha ha ha>'). If a laugh functions as a reply, it is treated as one discourse sentence, but otherwise, it is placed in the utterance or at the end of the discourse sentence which includes the laugh.

Example 31.

If a laugh is produced overlapping another speaker's utterance which is in progress, it is treated in the same way as a short backchannel and transcribed as '<laugh>'.

c. Contextual information

In order to help understand the context of utterances, prosodic features (accent, pitch, volume, speed etc.) are recorded when necessary for use by analysts as background information.

Example 33: [whispering], [Inhaling], [loud voice ↑], [sound of drinking] etc.

6.4 Visual effects

The following symbols are used to highlight such elements as direct quotations and proper names for easy recognition.

a. Quotes

When a speaker directly quotes his/her own or another person's speech, the quoted part is placed in double quotation marks.

Example 34:

b. Highlighting

Elements which should be highlighted for easy recognition, such as individual letters spelled out, proper names such as a book title, or utterances directly expressing the speaker's idea or emotion, are italicized.

Example 35:

6.5 Privacy considerations

Since the content of conversation may concern the speakers' identity or privacy, two sets of transcripts are made in BTS, one which records the conversation as it is, and one which can be used for publication. In the latter, personal information which may concern the cooperators' identity or privacy, such as proper names, is not transcribed. In conversations between strangers, for example, speakers often explain their own names or discuss episodes involving their names. The following explains how such cases are to be transcribed in consideration of privacy protection. Other proper names such as place names, or school names, etc. are also treated in the same manner.

- According to how names are introduced, they are transcribed as 'JBM03 Family name', 'JBM03 Given name', or 'JBM03 Full name' etc.
- If the speaker uses another word or name etc. to explain his/her name, it is not transcribed as it is. It is to be transcribed, for example, as 'A word which is related to JBM's name'.
- When (part of) a proper name overlaps another speaker's speech, letter 'N's are used to represent the number of syllables included in the overlapped part, in order to clearly indicate which portion is overlapped. An explanation needs to be provided afterwards to suggest what kind of word(s) are actually said but omitted in the transcript.
- When different versions of the same name are mentioned to discuss personal names, they are indicated as 'Given name 1', 'Given name 2' etc.

## 7 Key to transcription symbols

Below are summarized the symbols used in BTSE, which have been introduced above. Listed first are the basic symbols used for transcription used in BTSE. BTSE, however, is a 'basic' set of transcription rules as its name indicates, and it is hoped that users add their own symbols depending on the specific purpose of their inquiry, for example by adding more detailed prosodic information. A sample case of such an addition of purpose-specific symbols is provided at the end.

- .
- .. A period indicates the end of a discourse sentence.
- „ When a second speaker starts his/her utterance while one speaker's discourse sentence is not yet finished, a pair of commas is used to indicate that the first speaker's discourse sentence is still in progress. The second speaker's utterance is transcribed in a new line.
- , 1. Commas are used according to their conventional usages.  
2. A comma is also used to indicate a short interval in speech.
- ? A question mark is used for a question. It is also used at the end of a sentence without the grammatical form of an interrogative sentence, if the sentence is produced with a rising intonation and functions as a question.
- ?? A pair of question marks is used when a rising intonation is used in mid-sentence for the purpose of confirmation, etc.
- [ ↑ ][ → ][ ↓ ] If judged necessary, intonation is recorded using one of these symbols, which represent a rising, flat, and a falling contour respectively.
- /int/ This symbol is used when a slight interval is felt to exist in the tempo of speech.
- /sil 2/ A pause longer than one second is treated as 'silence' and recorded with the length of silence in seconds as in the example. If a silence functions as a reply, it is treated as one discourse sentence and assigned one line. Otherwise, the symbol for silence is placed at the beginning of the next speaker's line to help identify clearly which speaker broke the silence.
- = = These symbols are used when there is no interval between utterances, or the interval is relatively shorter than the average length of intervals in the particular conversation. These symbols indicate that two utterances or discourse sentences extend over two lines in transcription but were produced continuously. A '=' is placed at the end of the preceding line followed by a '.', or ',', and another '=' is placed at the beginning of the following line.
- ...
- < >{<} A set of three dots is used to indicate a hesitating tone whether it is in mid-sentence or at sentence end.
- < >{<} When there is an overlap of speech, the segment of speech which is overlapped by another speaker's speech is put in '< >', and then followed by '{<}'.
- < >{>} Likewise, the segment of speech which overlaps another speaker's speech is put in '< >', and then followed by '{>}'.
- [ [ ] ] Double square brackets are used when one speaker's discourse sentence is interrupted by another speaker's utterance, and ends at that point as a result. '[[' is added to the end of the first speaker's utterance (before the period), and ']]' is added at the beginning of the second speaker's discourse sentence.
- [ ] Contextual information can be provided in single square brackets. In order to help understand the context of utterances, prosodic features (accent, pitch, volume, speed etc.) are recorded when necessary at the end of the line.
- ( ) Short backchannels without a particular meaning are put in brackets and inserted within the other speaker's utterance at the location judged to be the closest point.
- < > Laughs or utterances produced while the speaker is laughing are transcribed as '<laugh>', '<while laughing>', or '<both laugh>' etc. If a laugh functions as a reply, it is treated as one discourse sentence, but otherwise, it is placed in the utterance or at the end of the sentence which includes with the laugh.

- (< >) If a laugh is produced overlapping another speaker's utterance which is in progress, it is treated in the same way as a short backchannel and transcribed as '<laugh>'.
- “ ” A direct quote of another person's or the speaker's own speech is placed in double quotation marks.
- Italics* Elements which should be highlighted for easy recognition, such as alphabetical letters produced on their own, proper names such as a book title, or utterances directly expressing the speaker's idea or emotion, are italicized.
- ##### Unintelligible sounds are recorded with #####. The number of # indicates the number of syllables the segment is assumed to include.
- ‘ ’ Single quotation marks are used to indicate words which cannot be transcribed as they are for privacy reasons. They are also used to indicate the standard spelling of a mispronounced word.

## An example of an additional symbol used for a particular research purpose

Below is an example showing how BTSE can be used according to the purpose of study.

- // There can be more than one item to be coded in a discourse sentence or in one line of transcription, as is often the case with backchannels or laughs. In order to code these items individually, '/' is added after each item to code, and a new line is started after it. '/' thus means that a new line is started after it in order to facilitate coding, even though the line is not to be divided at that point according to the original transcription rules of BTSE.

Since these lines belong to the same discourse sentence, a same discourse sentence number is given to these multiple lines. In order to identify the order of utterances with '/', however, an additional symbol is added to the discourse sentence number (Alphabetical letters should be used for this purpose so that they will not be confused with serial line numbers of a single discourse sentence.)

Below is an example showing how '/' can be used for the coding of backchannels.

Figure 2

## 8 Input format

It is advisable to use the standardized input format described below to facilitate shared use of transcripts and to increase efficiency when more than one analyst process the same sets of data. For saving transcripts, the following standardized format should be used.

### 8.1 Data saving format

One transcript is to be saved on one sheet. <sup>ii</sup> Each transcript is saved under a file name which contains the speaker identification codes, the transcriber's name, and the date (Year00-Month00-Date00).

Example 41: BF01-SM03 Julia Ryan 050628  
(BF01, SM03 etc. are codes used to identify speakers.)

### 8.2 Page format

Each sheet is to be saved under the following format

- a. Header: Speaker identification codes are placed in the center of the header. The exact starting time and the ending time of the transcribed portion of the conversational data are measured and recorded in brackets. When the conversation is transcribed to its ending, it is indicated as FINAL, and if the conversation continues after the end of the transcription, it is indicated as CONTINUED.

Example 43: BF01-SM03 (0'00-3'20: CONTINUED)

(This header indicates that a conversation between speakers BF01 and SM03 is transcribed from its beginning for 3 minutes and 20 seconds and that the conversation does not end at the end of transcription.)

- b. Footer: Aligned to the right, the stage of transcription and the transcriber's name is recorded, separated by a colon.

Example 44: First stage: Julia Ryan

- c. Margins: Top and bottom margins are 2 centimeters each. Side margins are 1.5 centimeters each.
- d. Font: Times, 10 points.
- e. Columns: Five basic columns are to be used: Line numbers, Discourse sentence numbers, Discourse sentence completion, Speakers, Content of utterance. 'Line numbers', 'Discourse sentence numbers', 'Discourse sentence completion' columns are filled in according to the explanations provided earlier in this document. The speaker is identified in the 'Speakers' column using an identification code. The 'content' column contains the actual transcription of the words spoken.
- f. Widths of columns: The width of each column should be 5, except for the 'content' column which should be 55. From Excel's formatting options, 'Wrap text' (in cells) needs to be selected.
- g. Colors: Different colors are used for each speaker's lines to help differentiate between their utterances.

<sup>i</sup> Spreadsheet software may automatically recognize '=' as part of a mathematical formula. When inputting data, adding an apostrophe (') in front of '=' will eliminate this problem, and only the '=' will be shown afterwards.

<sup>ii</sup> If Microsoft Excel is used, for example, one transcript is saved on one sheet. One 'book' of Excel contains three 'sheets', but extra sheets without data will only make the file heavier and therefore should be deleted. However, when more than one transcriber is working on the same conversation, or one person is processing more than one transcript, those conversations are saved under different names which include the stage of transcription, the name of the transcriber, and the date on which it is saved.

## 自然会話コーパスの分析から得られる会話教材作成の視点<sup>1</sup>

木林理恵

### 1. はじめに

本稿では、言語運用という観点から作成された自然会話コーパスである『BTSによる多言語話し言葉コーパス—日本語2』(宇佐美監修 2005)を紹介し、その問題点を踏まえたうえで、収録された自然会話の一部をデータとして用い、間接的な発話によって機能が実現されているかどうかという、コーパス作成時とは異なる角度から再分析を行う。そしてその結果から、自然会話を日本語教育に生かすための視点を探る。

### 2. 『BTSによる多言語話し言葉コーパス—日本語2』について

ここでは、本稿のデータとして使用する『BTSによる多言語話し言葉コーパス—日本語2』の概要を紹介し、このコーパスで言われている「タイプ2」について検討する。

#### 2.1. 『BTSによる多言語話し言葉コーパス—日本語2』の概要

このコーパスは、自然会話と創作会話の特徴を比較・分析することを目的として作成されたもので、「挨拶する」「依頼する」などの発話の「機能」に該当する部分を様々な自然会話資料から抽出しまとめるという形態になっている。機能は、東京外国語大学「会話モジュール」で扱われている40の機能をもとにしている<sup>2</sup>。

本コーパス作成の際には、「挨拶する」「依頼する」といったそれぞれの機能に対して、基準を設定し、宇佐美まゆみ研究室にある自然会話データから、各基準に合う談話を抽出した。その作成過程の詳細は関崎他(2004)を参照されたい。

本コーパスの特徴として、当該の機能と、当該の機能を表す典型的な言語形式<sup>3</sup>という観点から、談話を3つのタイプ(表1)に分類して収録していることがあげられる。関崎他(2004)によるとコーパス作成にあたっては、語用論の分野において、状況に応じて間接的な言い回しを選択されることや、発話の字義通りの意味とは別の意味が伝わるという現象が指摘されている(Austin 1960, Searle 1986, Brown and Levinson 1987, Thomas 1998 など)ことを考慮に入れた、と述べられている。

表1. 当該の機能と当該の機能を表す典型的な言語形式の有無という観点からの抽出談話の分類

|       |   | 当該の機能を表す典型的な言語形式 |      |
|-------|---|------------------|------|
|       |   | 有                | 無    |
| 当該の機能 | 有 | タイプ1             | タイプ2 |
|       | 無 | タイプ3             | ——   |

<sup>1</sup>本稿は、日本語教育国際研究大会(2006年8月5日-6日、ニューヨーク、コロンビア大学)にて発表されたパネル「自然会話分析と日本語教育: 自然会話の特徴を生かした会話教材開発の展開の可能性」(まとめ役・司会: 宇佐美まゆみ)における筆者による口頭発表「自然会話コーパスの分析から得られる会話教材作成の視点」をまとめたものである。

<sup>2</sup>東京外国語大学大学院地域文化研究科21世紀COEプログラム「言語運用を基盤とする言語情報学拠点」では、17言語にわたる外国語学習教材として、40の機能ごとに40のスキットを収録した会話モジュールをウェブ上で公開している。URLは次のとおりである。http://www.coelang.tufs.ac.jp/modules/

<sup>3</sup>当該の機能を表す典型的な言語形式は、予備調査の段階において当該の機能を実現していると判断して抽出した談話に多く見られた文型や表現、及び、日本語教育教材の中に見られる表現のうち、当該の機能を表すことが多いとされている文型や表現とした(関崎他 2004)。ゆえに、本稿で言う「典型的な言語形式」とは、ある程度操作的に定義されたものである。