先行研究におけるトランスクリプト・システムの概観と「基本的な文字化の原則 - Basic Transcription System for Japanese (BTSJ) の意義[1]

木山幸子・関崎博紀・木林理恵・施信余・金庚芬

1. はじめに

話し言葉の分析には、その音声を文字に起こした資料、すなわちトランスクリプトという道具が必要になる。トランスクリプトは、研究目的や分析の観点に応じて、それぞれの研究者にとって興味ある事象を効率よく浮かび上がらせるように構成することが不可欠といえる。それゆえ、同じ音声資料を文字化するにも、その研究目的・分野によって、トランスクリプト・システムの様相は大きく異なる。たとえば、1994 年 7 月に行われた社会言語学研究会シンポジウム「談話資料の観察・記述」では、複数の領域の研究者が同じ音声資料を文字化しているが、ここからも、表記のし方、発話の区切り方、またその配置のし方についてなど、非常に多様であることが一見して分かる。

様々なトランスクリプト・システムが存在する中で、本稿では、「言語社会心理学的アプローチ」[2]に適するように開発された、「基本的な文字化の原則 - Basic Transcription System for Japanese (BTSJ)(宇佐美 1997、改訂版:宇佐美 2003)」の位置付けを明らかにし、どのような点に「BTSJ」の特徴と意義があるかを示したい。

そこで、まず 2 節において、Edwards(1993、*Talking Data* 1 章)をとり上げ、文字化の方法を設計する上での原則や留意点、文字化の対象となるいくつかのカテゴリーにおける文字化方法の選択肢などを確認する。続いて 3 節で、*Talking Data* (Edwards & Lampert 1993)の 2 章～6 章より、各研究目的に応じて開発されたトランスクリプト・システムを概観する。さらに、4 節では日本語のトランスクリプト・システムの概観を踏まえた上で、5 節において、会話分析への言語社会心理学的アプローチに適するように、「基本的な文字化の原則(BTSJ)」が設計されていることを確認する。

2. 談話を文字化するための原則－Edwards(1993) 'Principles and Contrasting Systems of Discourse Transcription' のまとめ -

文字化の方法の基礎的原則を確認するために、Edwards(1993)をとり上げる。ここには、トランスクリプトにとって重要なことが詳細にわたって議論されているが、本稿でも、これを順次追っていくことにする。Edwards の論稿の進め方に沿い、まず、2.1 でトランスクリプト・デザインについての原則を、続く 2.2 で、文字化のシステムの比較とそれらが示唆することについて確認する。

2.1 トランスクリプト・デザインについての原則 (Edwards :2 節)

Edwards は、トランスクリプトを設計する際の 3 つの原則を挙げている。それは、カテゴリー・デザイン、読みやすさ、コンピュータ操作のしやすさである。以下、各原則について、Edwards(1993)の述べるところを概括していく。

①カテゴリー・デザイン

文字化作業をする際には、例えば、特定の長さを持ったポーズを「長い」「短い」と分けるといったように、ある情報をいくつかのカテゴリーに括ることになる。カテゴリー・デザインというのは、そうした文字化の対象となる情報を括るカテゴリーをどのように設定するかということである。Edwards

---

(1) 本稿は、東京外国語大学の宇佐美まゆみ教授の指導のもと、筆者たちによる討論を経てまとめたものである。

(2) 「言語社会心理学的アプローチ」については、その関心事や具体的な方法論の詳細が宇佐美(1999、2001)などに述べられている。

によると、そうしたカテゴリーを設計する場合に重要となる 3 つの特質は、体系的に識別できるものであること(systematically discriminable) 、網羅的であること(exhaustive)、体系的に対照させることができること(systematically contrastive)である。各カテゴリーが「体系的に識別できる」というのは、あらゆる場合において、データがどのカテゴリーに該当するか(例えば、長いポーズ、短いポーズなど)について明確な基準を設けていなければならないということを指す。「網羅的である」とは、データ中のどのような事例にも、それが適合するカテゴリーが用意されてなければならないということを指す。「体系的に対照させることができる」というのは、各カテゴリーが、関心のある事象を最もよく抽出するような形で対照されうるものでなければならないということを指す。

②読みやすさ

研究の効率を考慮するとき、読みやすさは非常に重要な役割を果たす。Edwards は、読みやすいトランスクリプトを設計するには、本や新聞、時刻表や広告などのような、ふだん読み書きしているとおりの慣行にならった表記をすることが重要だと述べている。そうすることで、研究者が必要とする情報を最も迅速に抽出することができるのだという。情報を効率的に把握するという観点から、太字や下線などの視覚的な卓立(visual prominence)を利用したり、文字化した情報の空間的配置のし方(spatial arrangement)を工夫したりするなどの方法も提示されている。

他にも、様々な種類の事象や情報が互いに密接に関連するものであるほど、関連の薄いものよりも空間的に近くに配置させること、その反対に、質的に異なる事象や情報（発話された語と研究者のコメント、コードとカテゴリーなど）を明確に区別できるように記述することが挙げられている。さらに、時間的に先にある事象を、時間的に後に来る事象よりもページの中での早い位置（下より上、右より左）に設けること、発話を解釈するために論理的に必要不可欠な情報を、それと関連する発話よりもページの中の早い位置に設けること、特定の現象を記号化する際には、通読するときに意味の復元がしやすいよう、直接解釈可能な省略形か象徴的な図形表示方式を用いること、意味が簡単に復元できる範囲でコード化されたものの区別のし方をできるかぎり少ない符号で表し、不要なまぎらわしい混乱を最小限に抑えることなどが紹介されている。

③コンピュータ処理のしやすさ

Edwards がコンピュータでの分析にとって最も重要であると指摘するのは、符号化の組織性(systematicity)と予測性（predictability）だ。データにコンピュータ処理を施す際にこの 2 点がおろそかにされると、研究結果にも影響するような検索エラーが生じかねないからである。検索エラーには、データの中の関連する事例を見逃してしまう「不十分な選択(underselection)」と、関連のある事例に混じって関連のない事例まで検索してしまう「過剰選択(overselection)」の 2 つが挙げられている。前者は、'because'の意で発せられた'cause'を発音の通りに表記した場合に別のものとして認識されてしまうといった類のことだ。これを防ぐためには、変異形の隣に予測される標準形が記される。後者は、発話のラインに加えてコメントが書かれているラインからも対象を検索してしまうといったことであるが、対処法、検索の際にそれらが性質の違うものとして扱われるよう体系的にマークしておくことが述べられている。

2.2 文字化のシステムの比較とそれらが示唆すること（Edwards ：3 節）

次に、Edwards は、空間的配置、記述の種類・レベルという観点から様々なシステムを比較し、それらのシステムの類似点と相違点が示唆することを考察している。以下、その 2 点を追っていく。

①空間的配置

トランスクリプトの空間的配置としては、別々の話者によるターンをどう配置するかということと、発話に関連のある文脈的コメント、身ぶり、韻律、コーディングなどの情報をどのように表し配置する

かという 2 点が問題となっている。

別々の話者によるターンの配置には、垂直(Vertical)、列(Column)、分割(Partiture)の 3 つの方法が紹介されている。

・**垂直(Vertical)**：話者たちが相互作用の過程に平等に従事し、平等に影響を及ぼしていることを印象付ける。

例)

A: Did you just get [back]?
B: [Yes], or rather 2 hours ago. It was a great film.
A: Really?

・**列(Column)**：成人対幼児の会話に見られるような、相互行為者同士における非対称性を浮き立たせるのに効果的である。時間は垂直の次元で表わされ、2 人の話者による同時発話は、[ ] に入れたり、水平に並べたりすることにより表わされる。

例)

Speaker A	Speaker B
Did you just get [back]?	[Yes], or rather 2 hours ago.
	It was a great film.
Really?	

・**分割(Partiture)**：多くの同時発話や様々な行為が関連する相互行為の広がりを、高度に効果的に把握するものである。これは、話者たちのコミュニケーション上の地位が同等であることを暗示したり、ターン・テイキングの様相を強調したりしつつ、一貫して時間の流れを示しているというように、垂直形式にも列形式にも通じる特徴を持つ。会話に従事する同等の立場の話者間において、ターンのタイミングや連鎖を効果的に浮き立たせたり、オーバーラップを見つけやすくするという長所がある一方、トランスクリプトを修正する際に、特別な作業が必要になるという短所もある[3]。

例)

A: Did you just get [back]? Really?
B: [Yes], or rather 2 hours ago. It was a great film.

トランスクリプトの空間的配置に関わるもう1点は、文脈的コメント、非言語行動、韻律、コーディングの配置をどのようにするかということである。これには 4 つの選択肢が紹介されている (各例は、Edwards(1993)に紹介されていたものの二次引用である)。

・**ランニング・テキスト(Running Text)**

：非言語行動や文脈的事象を、発話に挿入する形で時系列に表示する(以下の例では、非言語行動が二重かっこ「(( ))」内に示されている)。

⑶ この問題を緩和するための、特別なコンピュータ・プログラムも存在する。

例) Atkinson & Heritage(1984, p. xiii)

Tom: I used to ((cough)) smoke a lot
Bob: ((sniff))He thinks he's tough
Ann: ((snorts))

・ **UPC(Utterance-Plus-Clarification)**

:非言語行動や文脈的事象を、発話を明確にするための情報として扱い、個々の発話とはラインを分けて表示する(以下の例では「*」で始まるラインが発話内用を表し、「%」で始まるラインが非言語行動や文脈的事象を表している)。

例) Mac Whinney & Snow (1985)

*MOT: what did you see?
%sit: Mother and Allison sitting on the chai; Allison wearing
half zippered jacket; fingers in her mouth
*MOT: what did you see over there?
%gpx: <aft>points to monitor
*ALI: Mommy.
%gpx: looking at monitor with fingers in her mouth

・**点在(Interspersed)**

:韻律情報などの特定の事象やその一部を詳述したラインをどこに表示するかということに関して、その記述が、発話とは区別できる形で正確に符号化されているのであれば、それを発話内容が記述されているラインに挿入する (以下の例では、「^」「/」などが韻律的情報を表している)。

例) Svartcik and quirk (1980), text S1. 3

1 3 7212280 1 1 A 11 and at ^h\/ome #.
1 3 7212290 1 1 A 11 she`s not a ^b\it the way shi is at c/ollege#

・**SPS(Segment-Plus-Specification)**

:韻律情報などの特定の事象やその一部を詳述したラインをどこに表示するかということに関して、その記述が複雑になりすぎる場合には、各情報を複数の層で記述していく。

例) Fillmore, Peters, Oestman, Larsen, O'Connor & Parker (1982)

S: He wen' (=went) out./
m: he *went out
s : (nla PN.3.M) (v *V PT)

これらの 4 つのうち、ランニング・テキスト形式が、時間に関するすべての事象を同じラインで扱っていると述べられている。ある一時の発話の構造の認知(オーバーラップ、連鎖、リズムなどの特質)を容易にするものであり、「読みやすさ」を重視する場合には適しているということである。

②記述のレベルとタイプ

さらに、Edwards は、トランスクリプトに記述する事柄のカテゴリーのレベルとタイプについて紹介している。その記述のカテゴリーとして、言葉の表し方(Word Forms)、分析の単位(Units of Analysis)、韻律(Prosody)、ターン・テイキング(Turn-Taking)、非言語的事象・非言語行動(Nonverbal Events or Actions)を挙げ、さらにそれらのカテゴリーに沿って情報を記述する際の選

択肢(例えば、標準的な正書法で表記するのか、IPA を使うのかといったこと)に触れている。以下、各カテゴリーで言及されていることを見ていく。

・言葉の表わし方(Word Forms)

「標準的な正書法は多くの目的に適うものであるが、特定の発音や方言が重要性を持っていたり、書き言葉にはない形式を話し言葉から符号化したりする場合においては、補てんが必要だ」(4)と指摘されている(p.20)。補てんの例としては、修正された正書法を用いること、IPAを使用することという 2 つの方法が挙げられる。前者は、'cause', 'cuz'の横に'because'と記すような方法を指す。この方法は特殊な言語学的訓練を要さないという長所がある一方、書記素と音声の一致があいまいであるという短所もある。また、英語の非母語話者である研究者にとってはこれが難しいという可能性なども指摘されている。後者は、前者よりも正確な方法といえるが、特別な訓練を要するし、使用する際にも時間がかかる。これらのうちどちらを採用するかは、各研究目的、データの使用者に求められる発音記述の正確さの程度によるとされている。

・分析の単位(Units of Analysis)

研究目的によっては、改行や明示的マーカー(スラッシュなど)を用いることにより、談話テキストがいくつかの種類の単位に分割されることがある。例えば、イントネーションにもとづく単位('tone units', 'intonation units'など)や、間によって区切られる単位('production units'など)、もしくは両方を組み合わせた単位などが挙げられている。

・韻律(Prosody)

トランスクリプトを作成するときにとり上げられる韻律的な側面としては、ポーズ(Pause)、伸ばすことと縮めること(Lengthening and Shortening)、プロミネンス(Prominence)、イントネーション(Intonation)が挙げられ、それらに関する注意点が述べられている。

まず、ポーズについては、その長さをどのように捉えるか、どの基準によってそれをとり上げるか、また、ターンの間に生じたポーズをどのように表示するか、といった問題に触れている。これらをどのように文字化するかは、研究目的に応じてどういった情報が必要とされるのかによって異なる。例えば、相互行為者同士の認知にできるだけ近い形でポーズを捉えようとするときには、話者の発話速度にあわせて沈黙の長さを捉える、当該の場所に置かれるポーズとして期待されたものより長いか短いかという基準を使う、などの方法が述べられている。

伸ばすことと縮めること(Lengthening and Shortening)というのは、音節の長さの変化を指している。これは、規範的な長さを逸脱していた場合や、相互行為において意義があると考えられる場合にのみ、とり上げられる。

プロミネンス(Prominence)については、大半のトランスクリプトでは特異なものについてのみ記述されているが、研究目的によっては、すべてのプロミネンスを記述する場合もあるということだ。

イントネーション(Intonation)は、研究目的に応じて、楽譜を記すような文字化の方法をとってイントネーションを細かくとり上げるものから、発話中のプロミネンスのあるところから発話の末尾に向けてトーンがどのように変化しているかを捉えるものや、発話の末尾の音節におけるイントネーションの変化のみを記述するものなどが紹介されている。

・ターン・テイキング(Turn-Taking)

ターン・テイキングを描き出す際には、ターンの間で、あるいは、ターンを越えて見られる話者間のリズム上の同調に焦点をあてるシステムや、発話が完結しているか否か、つまり自分で終えたかまたはもう1人の話者によって発話が遮られたかを強調するシステムなどが紹介されている。また、

---

(4) これより以下、本稿における日本語訳の部分は、すべて関崎によるものである。

ラッチング(2 人の話者によるターンのあいだに期待される間(ポーズ)が短い現象)や、ターンのあいだに期待されていながら、欠落している間(ポーズ)なども、記述の対象となっている。

・非言語的事象・非言語行動 (Nonverbal Events or Actions)

たいていの会話のトランスクリプトでは、非言語行動については、大まかに触れているだけであるが、研究目的によって、目の動きや、うなずいているときの手の位置など、より詳細な記述を必要とすることもあると述べられている。

3. *Talking Data* に紹介されているトランスクリプト・システム

本節では、各分野における研究者が、目的に応じてどのようなトランスクリプトを作成しているかを概観していく。その際、2 節でとり上げた Edwards (1993)の紹介を踏まえながら、本稿では各トランスクリプトにおける表記のし方、音声的情報の記述、周辺言語情報の記述、トランスクリプトの空間的配置といった 4 つの側面から紹介し、各システムの特徴を簡単に示していく。

3.1 Chafe(1993)

Chafe(1993)は、「情報の流れ(information flow)」、とりわけ、情報の活性状態を捉えるのに適したトランスクリプト・システムを紹介している。ここでは、発話をイントネーション・ユニットやアクセント・ユニットといった単位に区切る作業が行われる。以下、トランスクリプト・システムと、これを用いた分析を簡単に見ることにする。

まず、発話の表記は、英語の標準正書法(standard orthography)によっている。ただし、uh や gonna のように、広く使用されている口語体の表現であればそのまま記述される。

そして、イントネーション・ユニットとアクセント・ユニットの核心と境界を作り出すのに有効であるという理由により、ポーズの長さ、通常よりも長い音節、句末ピッチ曲線のあり方、そしてプロミネンス(一次アクセントと二次アクセント)が扱われている。ポーズの表記については、長い、普通だ、短い、といった知覚的な区別が発見されるまでは、持続時間を絶対的に表示するという考えで、0.2 秒以上の聞き取れるポーズは、(. . .)で、0.5 秒以上のポーズは、かっこ内にその時間長を記入して表すようになっている。句末のピッチ曲線については、それが、落ちるか落ちないかを問題としており、落ちないピッチをコンマ(,)で、文末を示す落ちるピッチをピリオド(.)で表記している。プロミネンスは、通常のアクセントよりも大きな振幅や高いピッチを持つ語を、一次アクセントを持つ語、二次アクセントを持つ語というように同定している。一次アクセントは二次アクセントよりも、二次アクセントは無アクセントよりも、振幅が大きくピッチが高いものとなっている。

分析の単位には、イントネーション・ユニットと呼ばれる、句末ピッチ曲線によって区切られる分節を設定している。Chafe は、これを機能面から見て、実在的(substantive)と調整的(regulatory)とに、また形式的な面から見て、始まったものの完結されていない(fragmentary)ものと、完結したものとを区別しており、トランスクリプト上にそれぞれの頭文字(S)(R)(F)を記入するようにしている。また、上に一次アクセントを持つ語の後に、それぞれの一次アクセントの境界を示す作業がなされる。その境界は、統語論的な基盤によって決定されるものであり、(|)の記号を使って示される。その境界に区切られた単位がアクセント・ユニットと呼ばれ、イントネーション・ユニットの下位単位となっている。

以上のように作成したトランスクリプトを用いて、Chafe は、ある情報が、人間の意識の中で活性的(active)、半活性的(semi-active)、非活性的(inactive)のいずれの状態にあるかを捉えようとしている。活性的な情報とは、ある瞬間において話し手の意識の焦点となっている少量の情報、非活性的な情報とは、長年の記憶ともいわれる利用可能な大量の情報、そして、半活性的な情報とは、十分に活性的ではないが、非活性的な情報よりは利用しやすいものであると説明されている。そ

の中で、非活性的な状態から活性的な状態に移った情報を、Chafe は「新」情報と呼ぶ。

Chafe は、分析の結果、そうした情報の活性化はアクセント・ユニットごとに起こることを見出している。さらに、アクセント・ユニットにも実質的、調整的という区分を設け、それと情報の活性化との関連を探っており、情報の活性的、半活性的、非活性的という区分は、実質的なアクセント単位にのみ適用されるといっている。このように、アクセント・ユニットは情報の活性化を探る上での 1 つの基本的な単位となる。一方、イントネーション・ユニットは複数のアクセント・ユニットからなるものだが、そこに含まれる新情報、つまり活性化された情報を担うアクセント・ユニットは 1 つだけになる。つまり、イントネーション・ユニットは 1 つの新情報と、0 個以上の、半活性状態にある情報から構成されるものであるということだ。

したがって、ここでのイントネーション・ユニットは、新情報を見つける際の 1 つの指標となっている。また、これに含まれている半活性的な情報は、「新情報を理解するときに必要となる文脈をもたらすもの(p.41)」とされている。イントネーション・ユニットは、新情報を捉える際にも、またそれを理解する上でも有効に働くものと捉えられている。

Chafe は、情報が活性的であるか、半活性的であるか、もしくは非活性的であるのかを捉えことに関心を持っており、イントネーション・ユニットやアクセント・ユニットといった単位を設けていた。ここで紹介したトランスクリプトは、その関心に見合うように作成されるものであることから、韻律的情報を詳細に記述しており、発話の表記に関しては、発音の正確な再現を追求しないし、周辺言語情報の記述も特にしない。また、本文中では特に触れられていなかったが、データの空間的配置の方法を見てみると、各ライン内はランニング・テキスト形式で記述されていることが分かる。その理由は、分析の便宜のためにイントネーション・ユニットごとに改行するようにしているから、発話を記述した各ラインの間に韻律的情報などを記述したラインを設けたりすると、質的に異なる情報をもつラインを混在させることになり、分析の妨げにもなりかねない、ということだと考えられる。

3.2 「ディスコース・トランスクリプション(Discourse Transcription)」

次に、Bois et al.(1993)の紹介する「ディスコース・トランスクリプション(Discourse Transcription)」というシステムを見る。ディスコース・トランスクリプションでは、基本とされる記号と、個別の研究目的や興味による必要に応じて、選択的に使用することができる記号の 2 種類が設けられている。

まず表記は、読みやすさのために、通常は標準正書法による表記を奨励しているが、必要とされる表記の正確さの程度などによっては、音素による表記法も認められている。

改行の単位は、基本的にイントネーション・ユニットに求めている。イントネーション・ユニットは、それぞれの末尾のコンマ(,)やピリオド(.)で区切られる。他にイントネーション・ユニットに関わる記号としては、それらが最後まで言い切られずに途中で終った場合や、当該のイントネーション・ユニットが、他のイントネーション・ユニットに埋め込まれたものである場合や、1 人の話者によるイントネーション・ユニットが始まってから終わるまでの間に、別の話者による完結したイントネーション・ユニットが挿入され、先行するイントネーション・ユニットを便宜的に分割して表示する場合などを表すものが用意されている。

また、音声的情報として、ポーズは、話し言葉におけるポーズの位置とそのタイミングは、話し手の談話産出プロセスと、進行中の会話の相互行為への方向づけについての重要な情報を伝達しているという観点に立ち、会話中に見られる全てのポーズを抽出するようになっている。そして、抽出したポーズは、その持続時間に応じて、「短い」(0.2 秒以下)、「中間」(0.3 秒～0.6 秒)、「長い」(0.7 秒以上)、という 3 段階に区別するようになっている。ここでは、ポーズの持続時間を、発話のテンポとの相対的なものとして主観的に判断するのではなく、その物理的な継続時間に応じて表示することを選択している。また、2 人の話者によるターンの間にポーズが見られない「ラッチング」

や、2 人の話者によるターンが重複していることも記述する。さらに、ディスコース・トランスクリプションでは、分節(segment)を伸ばして発音することはイントネーション・ユニットを区切る際の判断材料の 1 つとして認識されており、（ = ）という記号で表される。プロミネンスについては、一次アクセント(primary accent)、二次アクセント(secondary accent)をとり上げている。一次アクセントは、「イントネーションによる行為(intonational action)」がとられたことをマークするものとして認識されている(5)。続いて、イントネーションは、機能、音声と、2 つの側面から捉えられている。まず、機能的な側面としては、終了(Final)、継続(Continuing)、アピール(Appeal)の 3 種類を文脈から判断して区別するようにしており、これらが移行の際の継続性(transitional continuity)を示すものだとして説明されている(6)。一方、音声的な側面から見たイントネーションの 1 つとしては、ピッチを扱っている。イントネーションのどの側面に注目し、それをどのように分類するか、ということは研究目的次第であるとしながら、ここでは、ピッチに、下降(Fall)、上昇(Rise)、平板(Level)、という 3 つの分類法が紹介されている。

なお、ブースター(booster)についても、選択的にマークできるように記号が用意されているし(7)、話し手の態度や発話生成のプロセスのうち言葉には表れない側面を表出するのに非常に重要な役割を果たすという理由から、声の大きさや、発話のテンポ・リズム、声の質などの表し方も用意されている。

周辺言語情報については、相手に進行中の言語的相互行為における微妙なキューを与えるものとして、咳払いや、あくび、鼻を鳴らす音などの音声的ノイズと、声門閉鎖音、吸気音、呼気音、笑い、などの記号が用意されている。このうち笑いに関しては、その持続時間の表記の仕方が示されていたり、笑いの種類(例えば、「鼻で笑った」など)を区別できるようになっていたりする(8)。

このように、ディスコース・トランスクリプションでは、話者同士の相互作用をよりよく捉えられるような形で、種々の記号が用意されている。上に見たイントネーション・ユニットに関するものからだけでも、相互作用上で起こりうるいくつかの状況を想定していると分かる。また、そうした観点に立っているため、アクセント・ユニットの境界を表す記号、発話の再開を示す記号、言い直し(false start)、コードスイッチが起きたことを表す記号、周辺語(marginal words)を表す記号、吸気や呼気、単語や笑いなど、単純なできごとの持続時間を示す記号と、ポーズとヘッジ表現の連鎖などの持続時間を表示する方法なども、必要に応じて選択的に使用できるよう記号が用意されている。このように、相互行為上で想定される現象を広く認めている中でも、韻律的情報に関しては、特に詳細な区別がなされており、そこに相互作用を捉える上での重要さを認めていることが伺える。

⑸ ちなみに、一次アクセントにあたる言葉のところで生じるピッチの変動の形は、一般にトーン（tone）と呼ばれているが、ディスコース・トランスクリプションでは、これがイントネーションの最も示差的な意味を担っているものと判断して、下降(Fall)、上昇(Rise)、下降上昇(Fall-Rise)、上昇下降(Rise-Fall)、平板(Level)、の 5 種を区別している。

⑹ 移行の際の継続性とは、同一話者によるイントネーション・ユニットが 2 つ並んでいた場合に、先に発されたイントネーション・ユニットが、次のイントネーション・ユニットに継続するのか、継続しないでそこで終わるのか、ということに関する性質である。それは、イントネーション・ユニットの末尾におけるイントネーションによって示されるようである。

⑺ ブースターとは、「非常に大まかにいえば、期待されたピッチよりも高いものである(p.58)」というように説明されている。

⑻ 他にも、ディスコース・トランスクリプションでは、発話中に引用文を表すための記号や、発話自体とは別に、会話の状況を説明するときなどのための記号なども用意されている。

最後に、トランスクリプトの空間的配置の方法については、基本的には文字列がランニング・テキスト形式によって表され、文字化作業者のコメントなどもライン内に記述されるようになっている。ただし、あまり注釈が長くなる場合には、SPS(Segment Plus Specification)形式によって、それらを行間に記載することもある。そして、各ラインは垂直(vertical)形式で配置されている。

3.3 「会話のやりとりの文字化 (Transcribing Conversational Exchange)」

この節では、Gumperz & Berenz(1993)から、彼らが採用しているトランスクリプト・システムをとり上げる。Gumperz らは、会話の中で使われるコミュニカティブなサインが会話の参与者の理解に働きかけるときの影響力を捉えようとしており、それに応じたトランスクリプトは、情報を伝達したり理解したりするために使われる言語的、非言語的サインが、話し手や聞き手に知覚されたとおりに描かれるよう設計されている。

表記法は、読みやすさを考慮し、会話の正書法(conversational orthography)と呼ばれるものを中心としつつ、なお音声を正確に表現したいときのために、一般的な口語体のつづりも適宜併用するという選択肢が提示されている。口語体のつづりを使用する場合には、どの語彙を口語体で表記したかを明記するように述べられている。

分析の単位は、「インフォメーショナル・フレーズ(informational phrase)」と呼ばれる独自の単位である。これは、「リズムの面から見て区切られるものであり、また、韻律面から決定される発話の一部分でもあり、さらに、1 つのイントネーション曲線をもつもの(p.95)」と述べられているように、発話に表れる諸々の特徴から総合的に判断して認定される単位だということである。

音声的情報について、まずポーズは、その長さを3段階に区別している。長さの段階に応じて、表記されるドット(.)の数が1〜3個になるが、1秒以上のポーズについては、それが、会話の中で自然と形成されるやりとりのリズムからの逸脱を知る手がかりとなるため、そのおおよその時間長を記述するようになっている。また、ポーズに関連して、2 人の話者によるターンの間に予期されるポーズが見られないラッチングや、2 人の話者の発話が重複した場合なども記述するようになっている。イントネーションは、そこに、「インフォメーショナル・フレーズ」の境界を示したり、ターンを保持するか譲渡するかを示したり、どれほどの確信を伴った発話であるかを示したりするといった機能を認め、それらをマークする。プロミネンスとしては、アクセントや声の大きさ、音節の長さなどをマークしている。アクセントの記述については、アクセントが付くことが予測できないものについては、そこにアクセントがあることを示すために、当該の単語に記号をつけるようにしている。声の大きさについては、声の大きさのみの記述ができるように、(ff)や(pp)などの記号が用意されている。このようにするのは、「叫んでいる」「ささやいている」などと任意の言葉で記述する場合のように、声の大きさ以外の韻律的要素に対する作業者の判断が入らないようにするためだと思われる。

また、会話の円滑さや修正(repair)を捉える際に重要になるものとして、韻律的特徴の連続である律動性(rhythmicity)もマークされている。さらに、「インフォメーショナル・フレーズ」のピッチ域やリズム、テンポなども、フレーズ全体の韻律的特質を表すものとしてマークされる。これらを表示する絶対的な数値は存在しないため、ここでは、当該の会話中における「普通の」高さや速さにとっての相対的な評価が表示される。

周辺言語情報の記述に関しては、特に記号を設けるのではなく、発話と同一ライン中に角かっこ([ ])に入れ、[nod]や[handing paper to doctor]などと記述するようになっている。

Gumperz らは、会話の中で使われるコミュニカティブなサインが会話の参与者の理解に働きかけるときの影響力を捉えようとしている。そのため、諸々の現象を、絶対的基準によって表すのではなく、話し手や聞き手に知覚されたように描くようになっている。この点が、上に見てきた 2 つの文字化システムと比べても特徴的といえる。なお、トランスクリプトの空間的配置については、特に

発話の背景的情報を提供する必要がある場合には、それを発話のラインとは別のラインに記述するUPC (Utterance-Plus-Clarification)形式が採用されているが、基本的には、各ライン中はランニング･テキスト形式によって書かれている。各ラインの配置は、会話の参与者同士によるやりとりを捉えやすい垂直(vertical)形式が採られている。

3.4 「HIAT」

ここでは、Ehlich(1993)から、「HIAT」というトランスクリプト･システムをとり上げる。「HIATは、コミュニカティブなデータを文字化するためのシステムである。(p.143)」とあり、教室内でのコミュニケーションや裁判でのコミュニケーション、子供の会話や討論など、様々な形態のコミュニケーションがHIATによって文字化されている。さらに、イントネーションや非言語行動を詳述できるバージョンとしてHIAT2も用意されている[9]。

Ehlich(1993)を見ると、HIATは、a)簡略性(simplicity)と妥当性(validity)[10]、b)読みやすさと修正のしやすさ[11]、c)文字化する人と読み手とがする訓練が最小限であること、という 3 点を念頭において作成されたものということである。

表記法には、標準正書法とIPA表記の双方の問題点を考慮した結果[12]、文字通りの表記(literary transcription)と呼ばれる表記法が採用されている。これは、基本的には標準正書法を踏襲しつつ、発音の再現が容易なように、適宜、発音された通りにその言語における一般的な文字を用いて文字化するといった方法である。

次に、HIAT で扱っている音声的情報には、ポーズ、イントネーション、トーン、抑揚の表現などがある。ポーズは、話し言葉においては発話の計画やターンの調整と関係するものとしての重要性を認め、その位置と絶対的な持続時間を表示するようになっている。また、イントネーション、トーン、抑揚についても記述されることになっているが、特定の場合に限られている。強勢や抑揚(長音化)は、標準的なパターンから逸脱していると捉えられるものについてのみ表記するようになっているし、トーンも声調言語を記述する場合やトーンの種類が意味の違いをもたらす場合にのみ、表記するようになっている。

⑼ イントネーションを詳しく記述したい場合には、発話の上に 5 段階のラインを設け、その動きが記述できるようになっている。また、非言語コミュニケーション(nonverbal communication)や身体の動き(action)を詳述したい場合には、それぞれを記述するためのラインを設けられるようになっている。

⑽ それぞれについての説明はとくになされていないが、それに先行する問題提起より、「簡略性」とは、もっとも本質的な情報を分かりやすく保存しつつ、読み手に過度の負担を強いたり分析を妨げたりしかねない多量の情報は記載を避ける、ということを指していると考えられる。また、「妥当性」とは、簡略性を実現するために、使用する記号は意味の復元が容易であるに設定するということを指していると考えられる。

⑾ 「修正のしやすさ」とは、分析の過程でデータに対する理解が進むにしたがってデータを修正していくことが生じるが、その際に修正を加えやすいように作られなければならない、ということを指しているのであろう。

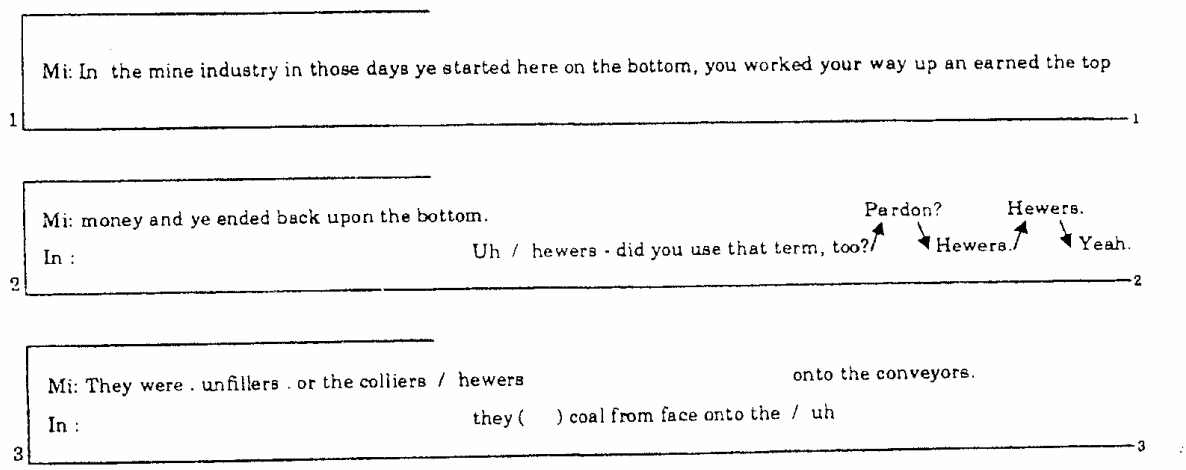
⑿ 標準正書法の問題点は、「後の分析において重要になってくる音声情報が膨大に失われる」(p.125)と指摘されているが、その解消を IPA に求めると、「情報が多くなり、それと引き換えに、使いやすさが失われる」(p.125)という問題が浮上する。また、IPA については、「書き起こす側にも、それを読む側にも特殊な音声学的訓練が必要になる」(p.125)という、HIAT 作成の理念に反する特徴も指摘されている。

笑いやくしゃみなどの現象や、犬のほえる声やドアが閉まる音などの描写的なコメントについても、これらが談話の文脈と関連すると考えられる場合についてのみ、記述される。

さらに、トランスクリプトの空間的配置については、HIAT では、会話における諸事象の「同時性」、「等時性」を表すために、スコア形式(score)と呼ばれる独特の書式が採用されている。下に、作成の際の原則を引用し、スコア形式の例を紹介する（表 1）。

1. 複数話者の発話を書いたり、イントネーションや非言語行動などの注意書きを書き込んだりするための、十分な縦幅をもったスペースを用意する。
2. 話者が変わった場合には、下の列に移ってその発話を文字化する。
3. 元の話者が発話する場合には、その話者の発話を文字化するラインに戻って文字化する。
4. ページの端まで続ける。
5. 1 から 4 を繰り返す。

表 1 HIAT が採用するスコア形式 (p.133)



上の例のように、各スコアは、左端の垂直線と、そこから伸びる上下 2 本の横線によって描かれる。下の線を引くのは、読みやすさのためであり、上の線を引くのは、羅列されたそれぞれのスコアが、どこから始まったものであるかを明確にするためである。

以上、HIAT のシステムを見たが、表記法には、正書法の持つ問題点と IPA 表記が持つ問題点を考慮した結果、折衷的なものが採用されていた。ここでは、トランスクリプトの役割として、特別な訓練をせずに文字化し利用できることや、読みやすさを備えていることに加え、必要に応じて発音が再現できるということをも重視しているからである。また、HIAT の基本バージョンでは、非言語行動を記述する方法は特に用意されていないが、これらの情報を詳述するために HIAT2 が用意されており、身ぶりにかかわった身体部位を表す 26 の記号によって、左右どちらの身体部位によるものかをも区別することができる。他にも、身ぶりの持続時間の長短を、簡略性との兼ね合いを考慮したうえで記述することができるようになっている。さらに、トランスクリプトの空間的配置は、各ターン内については特に言明されていないが、ポーズや割り込みの記号はランニング・テキスト形式

によって文字化されている。抑揚や発話のテンポなどを表示するには、発話を文字化したラインの上に付け加えるようになっており、SPS(Segment-Plus-Specification)形式が採用される部分も見受けられる。それらのターンは、分割(partiture)形式によって配置され、Edwards(1993)で指摘された、修正に特別な労力を要するという問題点も、独自のプログラムによって解消されている。加えて、HIAT 独自のスコア形式により、コミュニケーションが展開していく過程を捉えやすいように工夫が施されている。

3.5 「子どもの言語研究のためのトランスクリプトとコーディング (Transcription and Cording for Child Language Research)」

Bloom(1993)は、子どもの言語を研究するにあたって、ビデオを用いた分析法を提示している。ここで Bloom は、トランスクリプトを作成する際には、多くのデータを偏りなく文字化することが望ましいが、人間は本来的に興味のある情報を捉えようとする傾向があり、また物理的な限界があることも現実であり、文字化という作業にあたっては、どうしても作業者の興味・関心の対象となる情報が中心に記述される傾向が出てくるということを指摘する。ビデオには特に膨大な量の情報が含まれるため、その傾向がより強くなる。この問題を解消するために、Bloom は、1 つの資料に対して、複数の作業者が各々割り当てられた別々の項目(covariables)を文字化・コーディングし、後からそれらを合成するという手法を考案した。例えば、ある作業者は子どもの発話を文字化し、ある作業者は子どもの非言語行動をマークし、ある作業者は母親の動きをマークするといったことである。作業者は研究目的を知らされていないため、当該の項目に関してより公平な作業が可能になる。また、作業者を複数用意することで、より多くの項目、広範にわたる情報を抽出することができる。

この際、作業は、コンピュータとビデオ、タイムコーダーとモニターを使ってなされる。その際、コーダーが抽出した項目はタイムコーダーによって時間が記録されているため、後から、各項目を生起した時間順に並べることができる。そして、作業者が別個に抽出した項目を合成してトランスクリプトを作成するときには、目的に応じて各項目を自由に組み合わせることも可能である。組み合わせた後も、それぞれに生起した時間が記録されているため、それらの項目を時間順にべることができる。このようにして作成されたトランススクリプトは、やりとりの展開を検証するにも、描写的な分析をするにも利用できる。また、データはコンピュータによって容易に処理できるため、定量的な分析にも適する。

Bloom が開発した手法に特徴的なことは、1 つには、始めから目的に合わせてトランスクリプトを作成するのではなく、まず、複数の作業者に別個の項目を割り当ててそれらを抽出し、データを分析する段階になってから、研究目的に応じて項目を自由に組み合わせ、トランスクリプトを構成することができるという点である。こうすることにより、1つのデータを複数の研究目的に対応させることができるし、それぞれが観察したい項目を偏りなく抽出することもできる。また、このシステムでは、一連の作業がコンピュータを用いて行なわれ、データもコンピュータ処理が容易に施せるようになっている点も特徴の 1 つといえる。

4. 日本語のトランスクリプト・システムの概観

3 節では、*Talking Data* に紹介されているトランスクリプト・システムについての詳細な議論を追ってきたが、日本語においては、どのようなトランスクリプト・システムが用いられているのか、研究の関心・アプローチに応じてどのようなシステムが用いられているかを概観することにする。

日本語におけるトランスクリプトから、本稿では、沢木(1984)の紹介するトランスクリプト、水川(2001)の紹介するエスノメソドロジー的会話分析におけるトランスクリプト、工学的研究への応用も視野に入れたコーパス構築で使用されているトランスクリプト、佐々木(1996)による談話研究で使

用されているトランスクリプト、沖(2001)によるトランスクリプトをとり上げ、それぞれの特徴を、他のシステムと比較しながら順に見ていくことにする。

4.1 沢木(1984)の紹介する、方言研究におけるトランスクリプト

沢木(1984)は、「その方言の話者でない人にも録音が理解できるようにすることと録音された談話を後々の分析ができる形にすること」を目的とし、方言録音資料を作成する具体的な方法[13]を述べている。

まず、文字化の表記は、音声記号としてのカナが用いられる。これは、その方言の音韻体系を決定した音素的文字化と音声的文字化の中間的方法ということだ。この場合のカナは音標文字としてのものであるから、共通語における正書法は無視される。すなわち、格助詞の「ヲ」や「ハ」は「オ」「ワ」と表記されるし、長音を表わす「ウ」なども用いられない。国立国語研究所『方言談話資料(6)』では、表記法について次のように説明されている。

> 文字化は原則として表音的カタカナ表記によることとした。これは、利用者の便宜、文字化作業の能率などを考慮してのことである。ただし、対象とする方言の性格によって、カナ表記では特殊な字母を多数必要とし、かえって煩雑になると判断される場合は、国際音声字母による表記も可とした。なお、それぞれのカナで表わす具体的音声の範囲・内容については、各担当者が「解説」の中で説明することとした。(pp. 8-9)

また、言いさしやあいづちなどについても、全く知らない方言では、どれに意味があり、どれが大して意味のない言い方なのか区別がつかないという理由から、それらも一通り文字化するとされている。このような表記法に加え、方言の共通語訳が文字化した原文の理解の助けとして付せられ、さらに、訳で説明できない部分が注釈として補われる。

沢木の紹介するトランスクリプトは、音声の文字化、その方言の音韻体系についての解説、共通語訳、注釈付け、録音に関する情報の記録で構成される。このうち方言研究に向けられたトランスクリプトとして特徴的な点は、発音の精密な記述を試みていることだ。これは、その方言独特の珍しい調音をトランスクリプトに留めるためである。この点に主な関心があるため、その他の音声的情報、周辺言語情報、発話の区切り方、各種情報をトランスクリプト上にどのように構成し配置するか、といったことには言及していない。

4.2 水川(2001)による、エスノメソドロジー的会話分析におけるトランスクリプト

水川(2001)は、エスノメソドロジーに基づいた「会話分析(conversation analysis)」のためのトランスクリプトの一案を提示している。エスノメソドロジーは、社会学的観点から「実際に行われた日常的な相互行為における会話を具体的、詳細に分析する」手法ということだが、ここでは、録音・録画資料をともに扱い、音声に加えて動作の分析もなされる。トランスクリプトも、それに応じて、音声、視線、身体動作の 3 種類で構成されることになる。

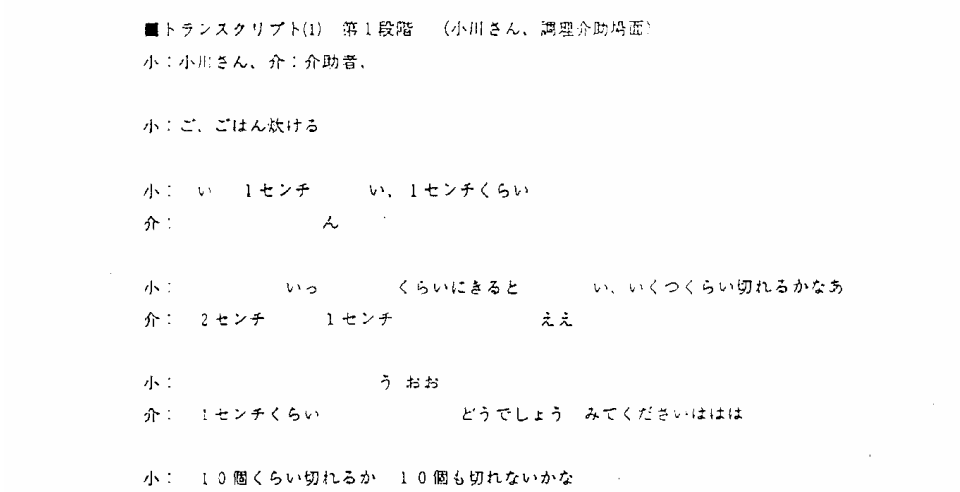
この音声トランスクリプトでは、例に示したように(表2)、基本的に2人の話者のターンが並行して配置され、それらがひとまとまりとなっている。水川は、これを「オーケストラ形式」と呼ぶ。また、オーバーラップやラッチングなど、欧米のエスノメソドロジー研究において標準的方法とされている記号を用いて示される。さらに視線・身体動作も、それらが相互行為に関連しているかという基準に照らして記述される。

---

(13) このトランスクリプト・システムは、国立国語研究所『方言談話資料』(1)〜(10)などで用いられているものである。

音声トランスクリプトは、分割形式によって構成されている。これは、Edwards(1993)が、分割形式について「相互行為の広がりを、高度に効果的に把握するものである」と位置づけているように（本稿 p.3、l.21）、話者同士の相互作用をより効果的に把握できるようにするためだと考えられる。表記は、漢字仮名交じりでなされているが、この点については特に詳細には論じられていない。

表 2. 水川が提唱するオーケストラ形式(p.79


■トランスクリプト(1)　第１段階　（小川さん、調理介助場面）
小：小川さん、介：介助者、

小：ご、ごはん炊ける

小：　い　　１センチ　　い、１センチくらい
介：　　ん

小：　　いっ　　くらいにきると　　い、いくつくらい切れるかなあ
介：　２センチ　　１センチ　　ええ

小：　　う おお
介：　１センチくらい　　どうでしょう　みてくださいははは

小：　１０個くらい切れるか　１０個も切れないかな

4.3 工学的研究への応用をも視野に入れたコーパス構築におけるトランスクリプト

日本語において、すでに、データの公開を目指したいくつかの話し言葉のコーパス構築が進んでいる。本稿では、それらの中から、言語のみならず工学的研究への応用も視野に入れたものを 2 つ紹介し、特徴を述べる。

①「日本語話し言葉コーパス（Corpus of Spontaneous Japanese: CSJ）」（前川他 2000、小磯他 2001）

「日本語話し言葉コーパス（以下 CSJ と記す）」は、国立国語研究所、郵政省通信総合研究所、東京工業大学の 3 機関のプロジェクトとして、音声認識などの工学的研究の要請および言語学的研究の要請に同時に応えるものとすることを目的としている。データは、学会講演、模擬講演、一般の講演や大学の講義という 3 種類から成る。

まず、文字化作業にあたっては、「発音形」と「基本形」という 2 種類の表記がされることになっている。これは、音響モデルと言語モデルを構築する必要性に基づいたものだ。「発音形」は、実際に発音された音にできる限り忠実に、カタカナを利用して表記される。「基本形」は漢字仮名交じりで表記されるが、表記のゆれが存在しないように、用語リストを作成して表記の際の原則を定めている。両表記の対応がとれるよう、文節に相当する単位で改行すると述べられている。

発話の区切り方は、「転記基本単位」として次のように定めている。CSJ では、転記基本単位の認定基準として物理的な指標を採用した。原則として、200 ミリ秒以上のポーズ（言語音の途切れに相当。息・咳・笑い声などの非言語音もポーズに含む）に挟まれた範囲を転記基本単位とする。ポーズの長さが 200 ミリ秒未満の場合には、そこで単位は分割せず、ポーズも 1 単位内に含める。ただし言語的な文末形式(述語の終止形や終助詞など)が存在している場合には、ポーズが 50 ミリ秒以上 200 ミリ秒未満であってもその末形式の後で転記基本単位を分割する。（小磯他 2001）

さらに、談話において生じる、発話以外のさまざまな現象を体系的に表現するため、トランスクリ

プトにタグが用いられ、フィラーや笑い、咳などを表す記号が、基本形と発音形の両方に付与されている。

②「日本語地図課題対話コーパス」(堀内他 2000、市川他 2000)

「日本語地図課題対話コーパス」は、千葉大学および外部組織から参加した研究者たちにより、より自然で音声品質の高い対話収録を目指して構築が進められているタグ付きコーパス[14]である。同コーパスは、言語学、心理学、文化人類学などの基礎研究的側面、音声処理、インタフェイス設計、言語処理などの工学的側面のいずれの研究関心にも応えることを目的とするということだ。データは、2 人一組の実験参加者が地図課題を共同で達成する過程を収録したものである。

まず、表記は、転記者によるばらつきが出る可能性のある方法は採らないため、漢字は用いられず、ひらがな表記となっている。ひらがなであれば、IPAやローマ字の分かち書きなどのような高度な専門知識を必要としないし、読みやすく、単語の検索が容易にできるからだ。漢字は、読みが一意に定まらず、何の漢字を用いるかも作業者の主観に依存し、また送り仮名も一定しないため、当コーパスでは採用していない。また、句読点も、話し言葉においては、書き言葉の持つ統語的構造が必ずしも明確に存在せず、どの位置に示すべきかの客観的基準がないため、使用していない。長音・促音[15]についても、同様の理由で使用されない。

発話の区切りについては、ポーズという物理現象を規準に発話単位(Utterance Unit: UU)を設け、この単位ごとに改行している。これを、「自己発話内の 400 ミリ秒以上の無音区間によって区切られた音声的連続」と定義している。また、音韻情報、談話情報、非言語情報などを包括的に表現し、各研究者の関心によって必要になる情報をも追加できるよう、共通交換フォーマット[16]を用いて、発話や各種の情報が配置される。

これらのコーパスのトランスクリプトは、いずれも大量のデータを処理することや、言語学的研究のみならず工学的研究にも利用するという目的があることから、文字化作業にあたっては、複数の作業者によってずれが生じないよう、絶対的な基準が設けられている点に特徴がある。したがって、表記についても、揺れが存在しないよう綿密な工夫を施しているし、発話の区切り方も、物理的な時間長を基準としている。またここで紹介したコーパスでは、例えばエスノメソドロジーにおけるトランスクリプトで見たように、相互行為の様相を視覚的にトランスクリプト上に表現するようには便宜を図っていない。

4.4 佐々木(1996)によるトランスクリプト

佐々木(1996)は、自然会話における話し方のスタイルの特徴を実証的に明らかにすることを目的にした研究の前提として、発話の単位やトランスクリプトについて言及している。

まず、会話のスタイルを分析する単位として、「発話文」を設けている。この認定基準は次のように定められる。

(1)ポーズによって区切られている。
(2)意味のまとまりを示す。
(3)構成部分が一定の音韻・構文上の決まりに従って結びついている。

---

(14) この対話コーパスの収録デザインは、エディンバラ大学における地図課題コーパスを踏襲したものであるということだ。

(15) 促音については、相当する部分が 100[ms]以上無音である場合、その時間長を表記している。

(16) この共通交換フォーマットは、TEI(Text Encoding Initiative)というアメリカの研究グループによって発表された、SGML というシステムにならったものである。

(4)発話者の交代がある[17]。

この発話文を単位としたトランスクリプト・システムに、12 項目にわたる文字化の際の原則を設けている。表記は基本的に漢字仮名交じりとすること、改行は同一話者の 1 発話文ごとにすること、あいづちも独立した発話と扱うこと、複数の話者による発話の重なりは位置をそろえて書きはじめること、音声的情報は顕著なものに限り示すこと、さらに、周辺言語的情報や状況説明は注釈として＜　＞内に示すことなどが記されている。

4.5 沖(2001)によるトランスクリプト

沖(2001)は、関連する先行研究を概観したうえで、談話のしくみを解明するという言語学的関心のもと、談話の最小単位の発見と検討を目的とした文字化のし方について述べている。

沖は、談話の最小単位は、川上蓁の提唱する、談話を形成する最小単位である「句」とするべきだと主張している。これは、文法論として説かれるものではなく音調単位であり、「その途中に切れ目を感じさせない、長さ不定の言葉」と定義される。この単位を採用する理由を、次のように述べている。

> 「句」という単位は、あらかじめ定まった長さをいうものではない。話し手の考えること表現したいことに対応して、短くも長くもまとまりをつけることでき（ママ）、聞き手は、その音調を手掛かりに、話し手の言いたいことを「句」の意味のまとまりで捉え、「句」が次々に短く繰り出されるのを手掛かりに、きめ細かく相手の思考を追うことができるのである。

そして、この「句」という談話の最小単位は、内容のまとまりをともなった音調形式であり、これによって様々な現象の説明が可能になると述べる。

トランスクリプトには、「楽譜形式」が採用されている。これは、実時間に添って話者の交替が明示され、沈黙やオーバーラップも視覚的に示すことができるよう構成されたものだ。表記については、具体的な言及はされていないが、トランスクリプトを見るとカタカナが用いられている。音韻的情報は、イントネーションや卓立が、発話の文字化されている行に示されており、状況の最低限の説明も、発話の下の行に記されている。

5. 「言語社会心理学的アプローチ」に適する「BTSJ―基本的な文字化の原則」

4 節で日本語のトランスクリプト・システムを紹介し、話し言葉を対象とする研究には、様々な観点や関心に応じて、どのような情報を浮き立たせ、どの程度まで精密に表わすかが変わってくることを確認した。しかし、日本語におけるトランスクリプトにおいては、3 節で見た *Talking Data* に紹介されている論文のようには、トランスクリプト・システム作成に関する原則が体系的に吟味され、目的に見合ったシステムの方針についての詳細な議論が十分にされているといえない部分がある。さらに、話し言葉を分析対象とする研究目的・分野は、4 節で紹介したものには留まらない。ここで、「言語社会心理学」という観点を提示したい。

「言語社会心理学」は、会話を分析対象としているが、主関心は、会話の構造そのものや単なる言語的特徴ではなく、会話を通して見る「社会の中に生きる人間の心理であり行動（宇佐美 2001）」である。そしてもう 1 つこのアプローチに特徴的なことは、会話の定量的分析を行い、結果の信頼性を高め一般化に迫ろうとしている点だ。この目的のもと、条件を統制して会話データを

（17） ただし、これらすべてが発話文であるための必要条件ではなく、(4)が起きないケースもあるということだ。

収集し、会話の中の現象をコーディングし、定量的処理を施すという手順が採られる（その枠組みは宇佐美1999に詳述されている）。このようなアプローチに適するトランスクリプト・システムの必要性により、「基本的な文字化の原則（BTSJ）」（宇佐美、1997：2003）が考案されるに至った[18]。

ここで、BTSJ はどのようなシステムであるかという特徴を、分析の単位、表記法、音声的情報、周辺言語情報、空間的配置という5つの観点から具体的に示す。そのうえで、最後に、BTSJの全体的な意義について述べたい。

・分析の単位

日本語では、スピーチレベルの分析など、文単位でコーディングする必要があるものがある。このため、BTSJ では、「発話文」を分析の単位としている。「発話文」とは、会話という相互作用の中において「発話された文」を意味し、基本的に「文」を成していると捉えられるもののことである。しかし、実際の会話の中で「文」を認定するためには、「話者交代」や「間」も考慮に入れられる。「発話文」は以下のように定義される。

> 「発話文」の定義は、会話という相互作用の中における「文」とする。そして、以下のように認定する。基本的に、「文」を成していると捉えられるものを「1発話文」とする。しかし、自然会話では、いわゆる「1語文」や、述部が省略されているもの、あるいは、最後まで言い切られない「中途終了型発話」など、構造的に「文」が完結していない発話もある。そのような場合は、話者交替や間などを考慮した上で「1発話文」であるか否かを判断する。つまり、「発話文」の認定には、「話者交替」、「間」という2つの要素が重要になる。

・表記法

沢木によるトランスクリプトでは、その方言の独特の調音をトランスクリプトに残すため、音標文字としてのカタカナが用いられていた。また、工学的研究への応用も視野に入れたコーパス構築にあたっては、あくまでも絶対的な基準を設けるという方針のもと、表記のゆれが生じる漢字仮名交じりは避けられ、カナ表記が採用されたものがあった。

BTSJでは、これらとは異なり、読みやすさを重視し、漢字仮名交じり表記を採用する。また、読点についても、読みやすさのために日本語表記の慣例にしたがって打つことになっている。ただし、なるべく話された通りに忠実にトランスクリプトに転記するために、日本語表記慣例にそぐわなくとも、間があった場合には打っている[19]。例えば、1 語を発する中で間が認められれば、「明日」を「あー、した‘明日’」のように表記する。さらに、データの共有化をも目指しているBTSJでは、なるべく統一された表記法であることが望ましい。そのため、長音や数字の表わし方などについての基本的な原則が提示されている。

・音声的情報

たとえば、3 節で紹介した Chafe のシステムでは、イントネーションやアクセントなどの韻律的情報を分析の中心的単位としているため、これらを詳細に記述していた。それに対して、ＢＴＳＪでは、上述したとおり、「発話文」という単位を用いている。発話文は文を基本とした単位ではあるが、可能な限りの情報を考慮しながら多角的な分析ができるように、発話の状況をできるだけ伝えるための基本的情報として、イントネーション、ポーズ、言いよどみを表すことにしている。

イントネーションは、上昇、平板、下降の 3 種類をマークしている。ポーズは、1 秒以下のものを「/少し間/」と記すようになっている。1 秒以上のものは沈黙として扱い、その時間長を「/沈黙 2 秒/」などと秒単位で記す。というのも、BTSJ では、話者同士がやりとりを交わしている中での「1 秒」という時間を、相互作用上相当の有標的（marked）な比重を占めるものと認めるからである。また、話者交替された 2 人の話者による発話と発話のあいだ、および 1 人の話者による複数の発話文のあいだについて、当該の会話の平均的な間の長さより相対的に短いか、まったくないラッチング、発

(18) BTSJ は、基本的には、会話分析への言語社会心理学的アプローチの関心事を捉えるのに最も適するものである。しかし、他のアプローチを採る研究にも利用できるシステムであることを目指しており、研究目的に応じて、原則や記号を追加して用いることをも推奨している。

(19) 句点のほうは、発話文ごとに打つが、コンピュータ処理上の役割も担っているため、必ずしも慣例的でない用い方となっている。

話の重複などもマークするようになっている。言いよどみの感じられるところには、「…」が付けられる。他に、研究者のメモとして分析の際に注釈的に活用できるように、声の大きさ、高さ、調子、発話の速さなどは、[大きい声で][ささやくように]と[ ]に入れて記述するようになっている。

・周辺言語情報

周辺言語情報については、笑い、あいづちを記述する様式が定められている。その他は、特記の必要のあるものは[ ]の中に自由記述するようにし、研究者のための注釈として役立てられる。なお、音声の文字化作業を中心に扱う BTSJ では、身ぶりやジェスチャーの表し方は定められていないが、もし特定の動作についての記述が必要な場合には、各研究者が、当該の動作を表すものとして独自に記号を定めることも、推奨されている。笑いについても、基本的には「<笑い>」というように一様に表記するが、個別の研究目的によっては、笑いが共起している発話に下線を引いたり、笑いの種類を区別したりすることなどもできる。

・空間的配置

BTSJ では、発話文や会話における各種情報の配置の仕方は、定量的分析を念頭に置いて設計されている。

まず、表計算ソフトを用いて、ライン番号、発話文番号、話者、発話文終了、発話内容という 5 つの列を設けるのを基本とする。また、発話内容と、音声的情報・周辺言語情報を同一のセルの中に記述するというランニング・テキスト形式が採られる。

改行は、基本的に話者交替によってなされるが、話者交替がなくとも、1 人の話者が複数の発話文を連続して発する場合は、1 発話文が終了したところで改行するというルールを設けている。このとき、各ラインの末尾には、発話文が終わって改行されているか否かの記号が必ず付せられ（終わっていれば「。」、終わっていなければ「,,」）、さらに発話文終了の項にも「＊」が付せられ、二重にチェックされる。これは、人間同士の相互作用としての会話を扱い、それを膨大に蓄積していこうとする際にはどうしても生じてくる間違いを、未然に防ぐために設けられたチェック機構である。ラインに付せられる番号も、通し番号とともに、発話文が終了したか否かを示す番号の 2 種類のものを各ラインに付していく。このようにして改行されるラインは、会話の参与者が相互作用に平等に従事し、その流れに平等に影響を及ぼしていることが把握しやすい垂直形式によって配置される。

以上、分析の単位、表記法、音声的情報、周辺言語情報、空間的配置と、5 つの観点から基本的なルールを紹介してきたが、BTSJ は、各研究者が独自の目的に応じてこのシステムを使用することを奨励している。会話の中で焦点を当てて分析したい項目があれば、各研究者が独自にコーディングの列を設け、コンピュータ処理できるようにも設計されているのである。そして、トランスクリプトを起こし分析していく過程において、「発話文」という単位の認定、改行のし方、さらにコーディングの段階で、客観性が保たれていることを判断するために、2 人の作業者間の判定の一致率(Cohen's Kappa)が一定の水準を満たしているかどうかを確認するという手順を踏んでいる。こうすることにより、各作業の信頼性が保証される。

すなわち、BTSJ は、日本語のトランスクリプトとしての汎用性を目指したシステムであるとともに、

会話における複雑な現象をコーディングし定量的処理を施していくのに有効なシステムであるといえる。

6. おわりに

ここまで、*Talking Data* で紹介されているトランスクリプト・システムの原則と、種々の研究目的・アプローチに応じたトランスクリプト・システムを概観し、そのうえで「言語社会心理学」的アプローチに適する BTSJ の特徴と意義を示した。今後、言語社会心理学の観点による複数の言語の対照研究が促進され、ひいては言語を越えた普遍性を追求するためには、日本語以外の言語においても、「基本的な文字化の原則」を確立する必要がある。そのために、現在すでに、韓国語版（BTSK）、中国語版（BTSC）の構築が進められている。

くりかえし述べたように、各目的によって適するシステムは当然異なるわけだが、各々が目指そうとする目的以外のトランスクリプト・システムにも目を向け、それらの間の類似点と相違点、それぞれの長所と短所を明らかにすることは重要と思われる。そうすることにより、ある現象を重視するときに必要な条件、不必要な条件といったことが見えてき、各システムをいっそう洗練させていくことができると考えるからである。さらに、異なる研究目的・領域間の位置づけを明確にすることは、トランスクリプト・システムの向上においてだけでなく、会話を分析対象とする諸領域の全体的な発展にもつながるであろう。

付記

本稿をまとめるにあたって、*Talking Data*(Edwards & Lampert 1993)第 1 部について、各章を要約した上での議論が持たれた。以下に担当者の名前を記す。

1 章　木山幸子*、2 章　金庚芬**、3 章　施信余*、黄瓊芸*、4 章　木林理恵**、5 章　関崎博紀*、6 章　謝オン**

*　東京外国語大学大学院博士前期課程

**　東京外国語大学大学院博士後期課程

参考文献

市川熹・堀内靖雄・土屋俊(2000)「日本語地図課題対話コーパス」『音声研究』Vol. 4, No. 2、4-15.

宇佐美まゆみ（1997）「基本的な文字化の原則（Basic Transcription System for Japanese: BTSJ）の開発について」『日本人の談話行動のスクリプト・ストラテジーの研究とマルチメディア教材の試作』文部省科学研究費基盤研究（C）研究成果報告書、12-26.

宇佐美まゆみ(1999)「談話の定量的分析 - 言語社会心理学的アプローチ - 」『日本語学』Vol. 18, No.11、明治書院、40-56.

宇佐美まゆみ(2001)「言語社会心理学」『言語 2001.2.別冊：言語の 20 世紀 101 人』Vol. 30, No. 3、大修館書店、231-234.

宇佐美まゆみ（2003）「改訂版：基本的な文字化の原則（Basic Transcription System for Japanese: BTSJ）」『多文化共生社会における異文化コミュニケーション教育のための基礎的研究』（科学研究費補助金基盤研究（C）2：研究代表者　宇佐美まゆみ）研究成果報告書、4-25.

Edwards, J. A. (1993) Principles and contrasting systems of discourse transcription. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 3-31.

沖裕子(2001)「談話の最小単位と文字化の方法」『人文科学論集　文化コミュニケーション学科編』(信州大学)、55-72

Gumperz, J. J. & Berenz, N. (1993) Transcribing conversational exchanges. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 91-121.

国立国語研究所(1982)『方言談話資料(6)-鳥取・愛媛・宮崎・沖縄-』

小磯花絵・土屋菜穂子・間淵洋子・斉藤美紀・籠宮隆之・菊池英明・前川喜久雄(2001)「「日本語話し言葉コーパス」における書き起こしの方法とその基準について」『日本語科学』9、43-58.

佐々木倫子(1996)「日米対照:女性の座談ー発話文の数量的分析を中心にー」『国立国語研究所研究報告集 17』、239-272.

沢木幹栄(1984)「方言録音資料の作成法」、飯豊毅一・日野資純・佐藤亮一編『講座方言学 2ー方言研究法ー』国書刊行会、338-347.

社会言語学研究会シンポジウム「談話資料の観察・記述」『社会言語学研究会予稿集』(1994)

Chafe, W. L. (1993) Prosodic and functional units of language. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 33-43.

Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993) Outline of discourse transcription. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 45-89.

Bloom, L. (1993) Transcription and coding for child language research: The parts are more than the whole. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 149-166.

堀内靖雄・中野有紀子・小磯花絵・石崎雅人・鈴木浩之・岡田美智男・仲真紀子・土屋俊・市川熹(1999)「日本語地図課題対話コーパスの設計と特徴」『人工知能学会誌』Vol. 14, No. 2、261-272.

前川喜久雄・籠宮隆之・小磯花絵・小椋秀樹・菊池英明(2000)「日本語話し言葉コーパスの設計」『音声研究』Vol. 4, No. 2、51-61.

水川喜文(2001)「会話分析による録画記録の利用法〜トランスクリプト作成の方法論〜」『北星女子短期大学紀要』Vol. 37、77-84.

Ehlich, K. (1993) HIAT: A transcription system for discourse data. Edwards & Lampert eds. *Talking Data*. Lawrence Erlbaum Associates, Inc. USA. 123-148.