

## BTSJに基づいて文字化作業をする際のQ&A まとめ

### 《文字化について》

**Q1:** 聞き取れないところがあり、文字化作業がなかなか進まない。

**A1:** 始めは聞き取れなかった部分が、何度も聞いていくうちに突然はつきりと聞こえる場合もある。もし、聞き取れない部分があってもそこに固執することなく、できるところから文字化をしていき、後でもう一度戻ればよいと思う。

**Q2:** 文字化は最初からPCで入力するのか？

**A2:** 作業者によって、楽な方法をとればよいと思う。なれてくると、PCで入力するのも楽になる。

### 《発話文認定について》

**Q3:** 発話文の認定の仕方は基本的にどのようなにすればいいのか？

**A3:** 構造的に「文」を成しているかどうか。ただし、実際の音声聞きながら判断する。「話者交替」や「間」なども重要である。例えば、構造としては倒置文と捉えられるものでも、音声を聞いてみて長い間があった場合は、2つの発話文であると認定する。

**Q4:** 1人が長く話しているとき、どこで1発話文とすべきかどうか場合に迷う。

**A4:** 構造的に終わっているかどうかポイント。しかし、間があいている場合は切ることもある。

**Q5:** 間があいた部分で発話文を切るかどうか？間をどのように判断するか？

**A5:** 他の部分と比べて相対的に長いかどうかや、イントネーションを考慮する。

**Q6:** 中途終了型発話なのか、次の発話に続けるべきか迷う。

**A6:** 第一には、構造的な面で判断する。第二の判断材料はイントネーションで、発話者が終わっているように話している場合は、そこで1発話文を区切り、中途終了型発話とみなす。

**Q7:** あいづちと笑いの扱い方について(1 発話文とするか否か)

**A7:** その会話におけるその話者の基本状態と照らして判断する。その会話における間の取り方や声の大きさなどの基本状態から見て、当該のあいづちや笑いに突出した特徴があるかどうかを判断する。

**Q8:** BTSJでは、実質的な発話文とあいづち的な発話文を区別できるようになっているか？

**A8:** Basicな部分としては、実質的であってもあいづち的であっても、「発話文」として同様に扱い、区別する記号などは特にない。研究目的によってその区別が必要な場合は、発話内容セルの右側のコーディングセルを用いて、「S(substantial)」、「B(backchannel)」などの記号で区別することが考えられる。

### 《沈黙の捉え方について》

**Q9:** どのくらいの沈黙が続いたら、間として捉えるべきであるか？

**A9:** 「話のテンポの流れの中で、少し間が感じられている時は、《少し間》と記すが、原則として1秒以上の間は、沈黙としその秒数を《沈黙 1.5 秒(秒数)》のように記す」と定義されているように、その会話でのテンポの流れから沈黙に感じられるところは間、沈黙として捉える。

### 《あいづちに関すること》

**Q10:** 1発話文と認めるか否かの判断は？

**A10:** 基本的にそのあいづちの前後に間がある場合や、声が大きかったり特徴があったり会話の中で意味を持っていたりする場合は1発話文としてとらえる。相手の発話内に入れてもよいあいづちは、かなり小声のものであり、相手の発話に完全に重なっているものである。

**Q11:** 非常に小さなあいづちなどでも表記する必要があるのか？

**A11:** 基本的に、聞こえたものはすべて書き起こす。

### 《記号及び表記について》

**Q12:** 「？」や「??」の区別がわからない。

**A12:** 疑問や確認の意図で発せられる発話には「？」をつける。語尾が軽く上がり、話者が相手に同調を求める意図が強く感じられる発話は「??」をつける。

**Q13:** 〈笑い〉と〈笑いながら〉の区別が難しい。

**A13:** 要はその場で何が起きているかが分かるようになっていればよいので、自分の判断で〈 〉の中の書き方を決めればよい。そのデータの中で書き方が一貫していればよい。

**Q14:** 「JBM(Japanese Base Male)」など、会話者の記号の付け方をどうすればよいのか？

**A14:** 自分の実験の条件統制しだいで、自分で判断すればよい。たとえば、親疎を操作している実験なら I(Intimate)と U(Unacquainted)という記号をつけたり、話す言語を操作するなら、J(Japanese)と C(Chinese)という記号をつけたりするなど。会話の流れを重視するなら(読みにくくなるなら)仮名を用いてもよい。目的はプライバシーを保護することである。

**Q15:** 会話の中で、何のことが分からないことが出てくる場合は？

**A15:** できるかぎり調べる。本人に聞けるなら、確認をとる。

**Q16:** 「…」をつける基準は？

**A16:** 音声的ないいよどみ。発話の最後が小さくなったり、遅くなっていたりするもの。文の構造には関係なくつける(例えば、中途終了だからつける、などということではない。

**Q17:** 文脈情報はどこまで細かく入れるべきか？

**A17:** [鼻をすする音][紙に書く音][語尾は小さい声で]など、なるべくその会話が連想できる必要な情報は明記しておく。

**Q18:** 発話の重複とラッチングの区別が難しい。

**A18:** ラッチングは、改行される発話と発話の間が、当該

の会話の平均的な間の長さより相対的に短い、まったくしないことを示すためにつけるものであり、発話の重複は、同時発話されたことを示すためにつけるものである。基本的に、聞こえたものを書く。

### 《その他》

**Q19:** 会話の内容で、人格を著しく傷つけるような発話が現れた場合、どうするか？

**A19:** 現在当研究室に蓄積されているデータには、幸いそのような場合が生じたことがないが、その場合は伏せる。伏せた上、文脈情報〔 〕にそのまま記載できない内容である旨を記しておく。

**Q20:** 留学生や学部学生などに向けて、自然会話の文字化体験をさせ、その特徴に気づかせたい。その場合、BTSJが適しているか(複雑すぎないか)？

**A20:** 少しの会話でもよいから、文字化させてみることに意義はあると思う。BTSJでは、あいづち、オーバーラップなど、自然会話に特有な現象をチェックするシステムであるので、体験した学生はより明示的にその特徴を認識しやすくなるだろう。

**Q21:** コーディング項目の数に制限はあるのか？

**A21:** 表計算ソフト上の操作上は、数の制限はない。ただし現実的には、1回の分析にそんなに多くの観点から分析項目を立てても收拾がつかなくなるので、整理して、3~5項目くらいにするのがよいのではないかな。

**Q22:** フォローアップアンケートを行う場合、学習者の母語に翻訳して使っているか？

**A22:** 宇佐美研究室には韓国語・中国語・英語の翻訳版がある。上級学習者にアンケートをとるときは日本語だけを使う場合もあるが、中級学習者にフォローアップアンケートを行うとき、翻訳版も提示する場合もある。