

エクセルで1文1行コーパス作成をする際の覚書

①文区切り

- ・執筆者番号のあとにハイフンを加え、文番号を順次施す。タイトル文は、文番号 00 をふり、本文は文番号 01 から開始する。
- ・原文に「。」があるところまでを1文とすることを原則とする。
ただし、述語文が「?」「!」で終了している場合や、文末形式がある場合（連体節等の埋め込み節である場合や、活用ミス等の誤用であると2名の分析者が判断した場合を除く）は、「。」がなくても一文と認定して「。」を付加する
- ・タイトル文は、末尾に「。」は付加しない。

②文境界の認定

- ・「あるいは／それとも／すなわち」といった接続詞でつながれている場合、原文が接続詞の前の文と後ろの文を「、」で区切っている場合は、原文を重視して全体で一文と認定。

例 JP117-01 世の中にインターネットが普及し、インターネットでニュースを見られるようになった現代社会において、従来からの情報源であった新聞や雑誌はこれからも必要とされるだろうか、あるいは不要となるのだろうか。

TM052-02 今日何か大事なことが起きたのか、世界で何があったのか、それとも生活に関する情報とか、すべてのインフォメーションはマウスをクリック押すだけで、あっという間に目に映るのだ。

TM052-23 だからこそ、私はインターネットの便利さを利用しつつ、新聞や雑誌も読む必要がある、即ち、新聞や雑誌は必要だと思っている。

- ・「だが／けれども」といった接続詞でつながれている場合、原文が接続詞の前の文と後ろの文を「、」で区切っている場合、前文に文末形式があれば誤用と判断し、二文に分ける。

例 TM052-14 新聞と雑誌は本を読んでいると感じさせる、。

TM052-15 だが、インターネットニュースはパソコンを見るだけのよう
な気がする。

③モダリティのコーディング

- ・文が引用句で囲まれている場合、すなわち直接引用であることが明示されている場合は、その文はモダリティの分析の対象から外す。