「日本語話し言葉コーパス(CSJ)転記データの分析」

学籍番号：2J19F508-7

氏名：川崎咲希

　日本語話し言葉コーパス(CSJ)の転記データの分析結果と考察した内容を纏める。

**使用データ：**

　分析に使用したデータは以下のとおりである。

1. A01F0067/学会講演(女)
2. S00F0014/模擬講演(女)
3. A01M0020/学会講演(男)
4. S00M0053/模擬講演(男)

**【結果】**

　平均発話長および講演に占める発話の割合について、女性と男性のデータを表１に示す。

表１　転記分析結果



まず平均発話長について比較すると、模擬講演よりも学会講演が長くなっている。男性、女性の両方において同様の結果である。また、講演に占める発話の割合は学会講演よりも模擬講演の方が大きくなっており、男女ともに同様の結果となっている。

そこで、平均発話長について*t*検定を行ったところ、女性のデータにおいて有意差がみられた（*t*=5.701,*p*<.05）。男性のデータにおいては有意差がみられなかった（*t*=1.894, *n.s.*）。

発話長の標準偏差は模擬講演よりも学会講演が大きかった。

**【考察】**

　転記分析から得られた上記の結果より、学会講演は、講演時間に占める発話の割合が模擬講演よりも低いにも関わらず、発話長は長いか、同等であることがいえる。この結果に基づく学会講演の特徴として、ひとつの文の長さおよびポーズ長が長いことが挙げられる。この原因として学会講演は模擬講演に比べ自発性が低く説明的な内容であることが影響していると考えた。またベルが鳴った際に発話が途切れることは、ポーズ長が長くなることのほか、発話長にばらつきを生じる要因とも思われた。一方、模擬講演の特徴としてひとつの文の長さおよびポーズ長が短いことが挙げられる。模擬講演の内容には個人の主観が多く含まれており学会講演と比較すると日常会話に近いものとなっている。そのためひとつの文は極端に長くならず、短く速いテンポで発話されることが考えられる。これは発話長のばらつきが学会講演に比べ小さくなった要因でもあると思われた。