

韓国手話コーパス

ウォン・ソンオク、ホ・イル、ホン・スンウン、イ・ヒュンファ
(韓国・韓国福祉大学校)

2015年に開始された韓国手話コーパスプロジェクトは、大韓民国でろう者によって使用されている韓国手話の言語学的なコーパスを作成することを目的としている。このために、ソウル地域から、60名の生え抜きのろう者と生え抜きに近いろう者がペアで招かれた。ろう者は13のタスクをこなすように求められた。それらには、自由会話や多彩な話題についての議論、絵物語や映画の断片の再話などが含まれている。このほとんどが視覚的な刺激に基づくコーパス開発に加えて、すべてのタスクをテストするテストフェーズと、編集、時間の計測が最終版にむけて実施された。この自然主義的で誘導・コントロールされたコーパスの長さはおよそ3時間で、これは、完了した収録がおよそ90時間の手話言語のデータとなることを意味する。データの3分の2は、能力を有する韓国手話通訳者によって韓国語に翻訳された。12時間弱の手話言語のデータはELANを用いてタグ付けされた。ELANは、オランダにあるマックスプランク心理言語学研究所によって開発された専門的ツールで、映像・音声資源に対して複雑なタグ付けを行うことができる。手話言語のデータの転写は、大韓民国においては初めての試みであるとみることができ、そのため、転写フェーズは、韓国手話の運用能力はあるが、基礎的な言語学や転写方法に付き教育を受けたことのないタグ付け実施者のトレーニングを含んでいった。時間に同調したELANによるタグ付けが意味するのは、転写の焦点がID注釈(Johnston 2008)を用いて一貫したやり方で手話記号の単位を同定する、韓国手話の基礎的な転写であることを意味している。多数のタグ付け実施者がトークン化を行うことによる注釈に基づく転写を実施するためには、タグ付け規則を作成することが必須である。この規則では、一貫して体系的な標準的タグ付けを行うためにどのように注釈を用いるかを示した。

韓国手話コーパスプロジェクトの第2段階として、われわれは手以外を用いる記号などの詳細な転写を行うことを計画している。それとは別に、転写中にELANをサポートするための、技術的な解決を開発する努力が必要であるだろう。