

『日本語日常会話コーパス』の設計と研究の可能性

小磯 花絵¹

概要：言語生活の本質を解明するには、日常の言語生活を反映したコーパス設計が不可欠である。これまで種々の会話コーパスが構築・公開されてきたが、その多くは特定の場面や話者層に偏っており、日常生活の中で私たちがどのような言語行動をとっているかを調査することは難しいという問題があった。そこで国語研究所では、さまざまなタイプの日常会話 200 時間をバランス良く収録した『日本語日常会話コーパス』(*Corpus of Everyday Japanese Conversation, CEJC*) の構築を進めている。CEJC は、(1) 日常場面の中で当事者たち自身の動機や目的によって自然に生じる会話を対象とすること、(2) 多様な場面の会話をバランスよく集めること、(3) 音声だけでなく映像まで含めて収録・公開し会話行動を総体的に解明するための研究環境を提供することを目指しており、その一部にあたる 50 時間の会話を 2018 年 12 月に試験公開したところである。本発表では、CEJC の設計やデータの特徴について概説した上で、CEJC を用いることでどのような研究の可能性が開けるかを具体的な研究事例とともに示す。

¹ 国立国語研究所