Corpus linguistics and language ideologies

Rachelle Vessey
Newcastle University, United Kingdom

Although "language ideology” has been a topic of research for linguists since the 1970s (e.g. Silverstein, 1976), studies have been primarily oriented toward the field of linguistic anthropology. More recently, Milani and Johnson (2008) have highlighted how critical discourse analysis can be usefully applied to studies of language ideology. This paper outlines how corpus linguistics – and more specifically the corpus-assisted discourse studies (CADS) approach (e.g. Baker et al., 2008) – can provide valuable dimensions. More specifically, considerations of frequency, concordances, keywords, and dispersion can provide greater insight into both explicit and implicit language ideologies. While corpus methods and tools could never replace ethnographic approaches, they may usefully be combined with them. Crucially, such a combination of methods also allows researchers to examine cross-linguistic and cross-cultural data; this is particularly relevant in language ideology research, which is premised on the need to compare "the cultural and historical specificity of construals of language" (Woolard, 1998: 4).

Drawing on a large comparative study of French and English Canadian newspapers, this paper illustrates how corpus linguistics can contribute to the study of language ideologies in three ways. First, corpus linguistics can overview linguistic patterns and quantify discursive representations so that explicit representations of languages can be identified. Second, since the CADS approach is premised on comparative studies, it also enables researchers to identify "unsaid” or inexplicit language ideologies (Blommaert and Verschueren, 1998: 191; Woolard, 1998: 6). Finally, corpus linguistics can help to identify ideological processes such as indexicalisation and essentialisation, which contribute to the production and reproduction of language ideologies. To conclude, while the paper emphasises the value of corpus methods, it also notes that socially-oriented corpus linguistic research could also benefit from the rich literature and ethnographic data of language ideology, as some researchers (e.g. Subtirelu, 2013) have begun to demonstrate.

Keywords: corpus-assisted discourse studies, Canada, language ideology.