Corpus resources and tools have come to play an increasingly important role both in Translation Studies research and in translation practices. In Translation Studies, corpora have provided a basis for empirical descriptive research. Corpus-based studies usually involve the comparison of two (sub) corpora, in which translated texts are compared with either their source texts (parallel corpus) or with another (sub)corpus constructed according to similar design criteria (comparable corpus), either in the same or in another language. These corpora are used to investigate regularities of translated texts, regularities of translators and regularities of languages. Regularities of translation may consist either of universal features which are hypothesised to be distinctive of translated texts as opposed to non-translated texts, or of translation norms and strategies which characterise texts translated under specific social and historical circumstances. Regularities of translators are individual linguistic habits manifested through consistently different (unconscious) patterns of choices, independently of the source texts. Parallel corpora are used, together with bilingual or multilingual comparable corpora, to compare and contrast regularities of languages. Bilingual and multilingual corpora have also found application in translator education. Parallel corpora offer learners a repository of translators’ strategies and choices, while comparable corpora provide them with a mapping of the words and structures employed by different linguistic communities for building discourse. Approaches to corpus use in translator education include both training in computer-assisted translation and experiential approaches to corpus use. Practical applications of corpora in translation comprise resources and tools for assisted and automatic translation, including translation memories and corpus-based statistical machine translation. This chapter provides an overview of how corpora are used for descriptive and applied purposes, and discusses research findings and implications for translation trainees, professionals and users.
10.1 Introduction

Corpus linguistics, as the computer-assisted study of language in use, has its foundations in the technological advances which have made it possible to collect and analyse large quantities of text in electronic format. Its contribution to Translation Studies has been significant, since corpora have provided a basis for descriptive research and allowed for the empirical testing of theoretical hypotheses. The first applications of corpus linguistics to translation research date back to the 1980s, and corpus-based Translation Studies (CTS or CBTS) have grown since then into an established subfield. In 1998 Laviosa entitled her introduction to a special issue of the journal Meta ‘The Corpus-based Approach: A New Paradigm in Translation Studies’, and a number of monographs and collected volumes have followed in its wake. Similarly, articles and papers on translation and corpora abound in both corpus linguistics and Translation Studies journals, conferences and volumes (Laviosa 2002; Olohan 2004; Tengku Mahadi et al. 2010; Zanettin 2012).

Corpora have also had a decisive impact in translation practice, as concerns the work of translation learners, professionals and users. In translator education, applications include both experiential approaches to corpus use and training in computer-assisted translation. Most professional translators today rely to a large extent on computer-assisted methodologies to carry out their work, and translation memories, a specific type of dynamic parallel corpora, are a standard tool of the trade. Finally, practical applications of corpora in translation also include automatic machine translation systems which millions of people use in their everyday life, and which rely largely on corpus-based statistical machine translation techniques.

This chapter provides an overview of how corpora are used in translation research for descriptive and applied purposes, and discusses research findings and implications for translation and language learners, as well as for translation professionals and users.

10.2 Corpora in translation research

In the 1980s a descriptive approach to the study of translation established itself as a major conceptual framework in the field that has since become known as ‘Translation Studies’. This approach, which has been variously labelled with the terms ‘empirical’, ‘target-oriented’, ‘poly-system’, ‘manipulation school’ and ‘systemic’ (Hermans 1999: 7–16)