The Lancaster Corpus of Mandarin Chinese

(LCMC)

by

Tony McEnery

Richard Xiao

Lancaster University

Preface

The Lancaster Corpus of Mandarin Chinese (LCMC) addresses an increasing need within the research community for a publicly available balanced corpus of Mandarin Chinese. *LCMC* has been constructed as part of a research project undertaken by the Linguistics Department, Lancaster University. The corpus is designed as a Chinese match of the *Freiburg-LOB Corpus of British English (FLOB)*, and, as such, will provide a valuable resource for contrastive studies between English and Chinese as well as a sound basis for monolingual investigations of Chinese. The LCMC corpus is distributed by the European Language Resources Association (Cat. No ELRA-W0039).

We are obliged to the UK **Economic and Social Research Council** for funding our project (see Grant Ref. RES-000-220135). Without their help, this corpus would not have been built. We would also like to thank the presses, libraries and websites, as listed in the bibliographic document of this corpus, for providing the required texts, and Miss Xin Huang, for proofreading the scanned electronic texts.

The *LCMC* corpus, together with a spoken Chinese corpus and two comparable English corpora, is used on our new ESRC-funded project *Contrast English and Chinese* (Grant Ref. RES-000-23-0553).

Tony and Richard

February 2004

Contents