

A Corpus-Based Study of Demonstrative Phrases in English and Arabic

Mai Zaki¹

Abstract

This paper attempts a corpus-based study of demonstrative phrases in English and Arabic with the aim of comparing the syntactic, semantic and pragmatic behaviour of demonstratives in both languages. The corpora used for this study are (a) part of the British component of the International Corpus of English (ICE-GB), and (b) part of the Contemporary Corpus of Arabic (CCA) compiled at the University of Leeds and freely available for download. This study starts from the theoretical hypothesis put forward by Diessel (to appear) that the main function of demonstratives is to maintain joint attention in discourse. In order to evaluate this hypothesis, the paper focuses on demonstratives as determiners through the analysis of two features. First, it discusses the relation between the demonstrative determiners and their antecedents, in terms of classifying the types of use as direct coreferential, indirect coreferential, or anaphoric. In the cases of indirect coreference, a further classification is done on the basis of the semantic relation holding between the head noun in the demonstrative phrase and the antecedent (eg. hypernymy, synonymy, etc.). Second, the paper highlights the effect of genre on the use and function of demonstrative phrases. The English and Arabic corpora are both divided into two sub-categories: fiction (short stories) and non-fiction (essays on natural and social sciences). Preliminary results show, for example, that the vast majority of demonstrative determiners in non-fictional texts are directly coreferential with their antecedents. Examples of indirect coreference and non-nominal antecedents present interesting cases for the assessment of the amount of inference needed in the process of reference assignment. The analysis of those features in English and Arabic are also compared to similar analyses done for other languages. The results made available through this study could also serve as background information for the development of reference resolution systems for English and Arabic.

References

- Ammari, A. (2004). *al-Jumlah al-`Arabiyah : dirasah lisaniyah*. Rabat: `Abd al-`Aziz al-`Ammari.
- Buckley, R. P. (2004). *Modern Literary Arabic: A Reference Grammar*. Beirut: Librairie du Liban Publishers.
- Cantarino, V. (1974–75). *Syntax of Modern Arabic Prose*. Vols. I-III. Bloomington: Bloomington University Press.
- Diessel, H. (1999). *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins.
- Diessel, H. (forthcoming). *Demonstratives, joint attention, and the evolution of grammar*. *Cognitive Linguistics*.

¹ *e-mail*: mz106@mdx.ac.uk

- Klinge, A. (2006). The unitary procedural semantics of the, this and that. *Acta Linguistica Hafniensia*. pp. 54–77.
- Lakoff, R. (1974). Remarks on this and that. In *Proceedings of the Chicago Linguistics Society, Chicago 10*, pp. 345–56.
- Moutaouakil, A. (1986). *Dirasat fi na.hw al-lughah al-`Arabiyah al-wa.zifi*. al-Dar al-Bay.da: Dar al-Thaqafah.
- Vieira, R., Salmon-Alt, S. et al. (2002). Coreference and anaphoric relations of demonstrative noun phrases in multilingual corpus. Paper given at the 4th Discourse Anaphors and Anaphor Resolution Colloquium. Lisbon, 2002.
- Wilson, D. (1992) Reference and relevance. *UCL Working Papers in Linguistics 4*: 165–91.