

GLI OMONIMI ITALIANI PIÙ FREQUENTI:
UN' ANALISI *CORPUS*-BASED

FEDERICA CASADEI

Studi Italiani di Linguistica Teorica e Applicata, anno L, 2021, numero 2

ABSTRACT

The purpose of this research was to obtain frequency information about Italian high-frequency homonyms using corpus-derived data. A list of 6.415 homoforms was compiled, corresponding to 15.669 homonyms (e.g. the homoform *anche* cor-responds to the two homonyms *anche* adv./conj. and *anche* plur. of the noun *anca* 'hip'). For each homoform, the relative frequencies of the corresponding homonyms were obtained by manually checking the occurrences in the *Corpus di riferimento dell'italiano scritto* (CORIS). It emerges that most homonyms are strongly unbalanced, since for 85% of them the most common one accounts for 90% or more of the total occurrences of the homoform. Two sets of high-frequency homonyms are provided in the paper: a list of the 45 most frequent homonyms, whether or not they share the same part of speech (e.g. *comune* adj. and *comune* noun); and a list of the 50 most frequent homonyms belonging to the same part of speech (e.g. *calcio* 'kick' and *calcio* 'calcium'). Some pedagogical and research implications of these findings are briefly discussed.