Word frequency distribution per corpus language and media type, following Zipf's Law