

Comparing the use of Idiomatic Expressions in European Portuguese and European Spanish



An exploratory research

Ana Sofia Santos, Daniel Monteiro, Leonor González

Faculty of Social Sciences and Humanities – NOVA University Lisbon
Graduation in Language Sciences

1 Introduction

The present work presents an exploratory research on the comparative analysis of the use of 10 idiomatic expressions in European Portuguese and Spanish to **see which of the two languages most uses idiomatic expressions as an expressive-communicative resource**, in which discursive contexts both languages use them, as well as the frequency of use and the degree of variation of the given expressions.

2 What is an Idiomatic Expression?

An idiomatic expression (IE) is a crystallized and indecomposable phraseological unit, whose meaning is not literal. For example, the Portuguese IE **“feito ao bife”** (literally: *made to the steak*) means **“to be in trouble”**. This meaning is not created by the combination of the lexical units that compose it, but by means of **cultural sense**.

The IE is indecomposable in the sense that it requires a **specific number and type of lexical units, without which the idiomatic sense disappears**. Although IEs are generally rigid, some allow different rigidity levels. This indicates that a (un)predictable number of forms are possible for the same idiomatic meaning.

3 Corpus and Methodology

To analyze both languages, obtaining representative data for each language, and considering the significantly erratic usage of IEs as a communication resource, we needed a reference corpus for each language.

For Portuguese, the **CRPC** (*Corpus de Referência do Português Contemporâneo*¹) was key. This corpus contains **308 million words**. For Spanish, we used the **CREA** (*Corpus de Referência del Español Actual*²). This contains **350 million words**.

Both these corpora allow the researchers to analyze the typological dispersion of the IEs in the existent textual categories of the corpus (e.g., literature, press, legal texts, etc.)

For this exploratory study, we restricted the list of IEs to 10 per language, selected randomly from already existent and referenced lists. Our methodology for extracting the selected IEs was using **regular expressions** and **Corpus Query Language** (CQL) for Portuguese, and **native speaker intuition** on possible variations for Spanish, due to the inexistence of a Spanish reference corpora that allowed CQL as a method of research.

We defined the query expressions for each corpus, performed the queries and analyzed the data.

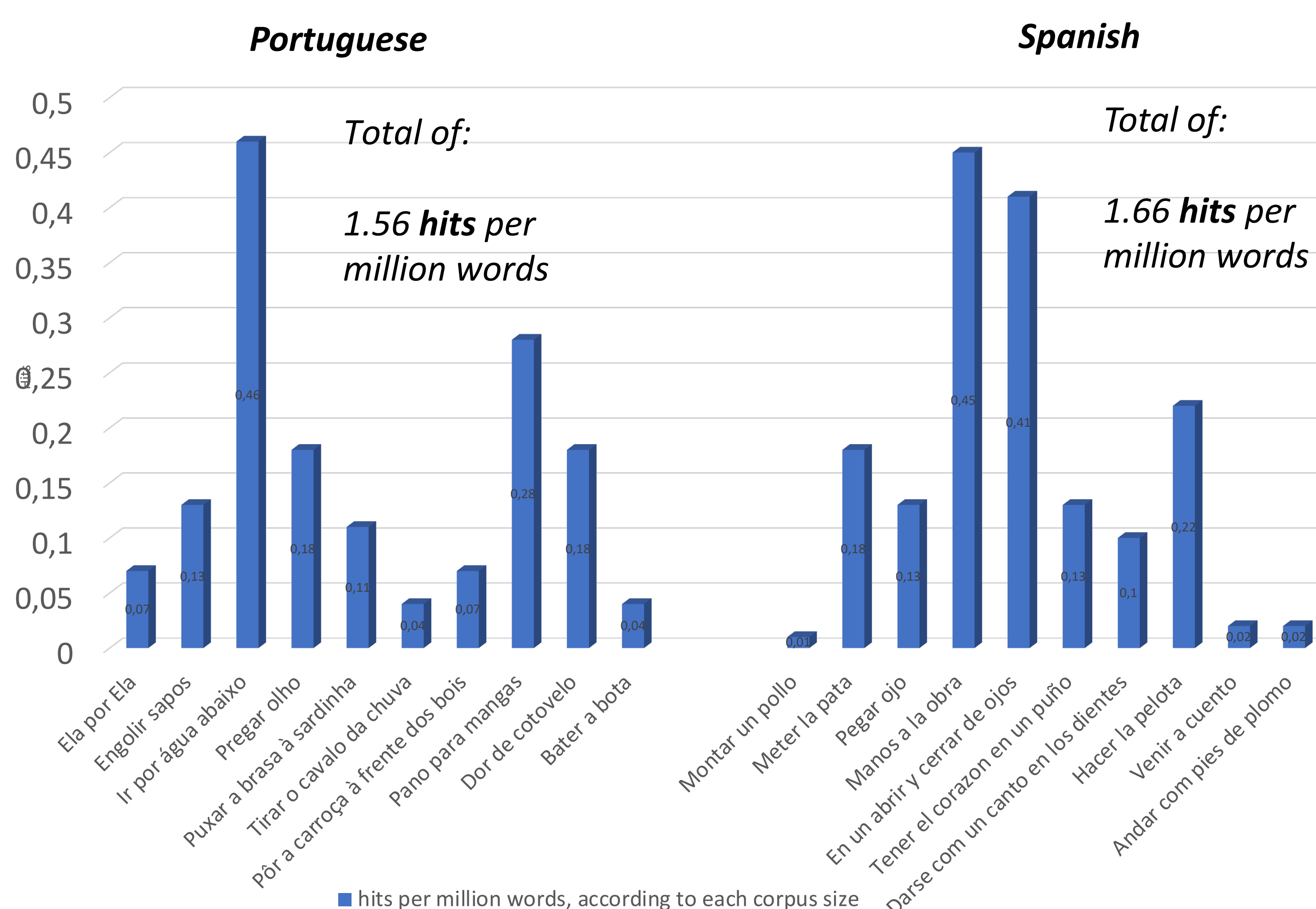
CQL expression: `{{puxar} | {trazer}} (_QNT*)? _DA* {brasa}_PREP* (_POSS*)? {sardinha}`
→ *puxar/trazer (tanto) a brasa à (minha/tua/sua/...) sardinha(s)*

1 <https://clul.ulisboa.pt/en/projeto/crpc-reference-corpus-contemporary-portuguese>

2 <https://corpus.rae.es/creanet.html>

4 Data

Portuguese IE	Spanish IE
“ela por ela” (=same amount)	“montar un pollo” (=start a scandal)
“engolir sapos” (=eat crow)	“meter la pata” (=speaking foolishly)
“ir por água abaixo” (=go down the drain)	“pegar ojo” (=to sleep)
“pregar olho” (=to sleep)	“en un abrir y cerrar de ojos” (=in an instant)
“puxar a brasa à sardinha” (=to feather one’s own nest)	“manos a la obra” (=hands at work!)
“tirar o cavalo da chuva” (=don’t hold your breath)	“venir a cuento” (=take the chance)
“pôr a carroça à frente dos bois” (=putting the cart before the horse)	“hacer la pelota” (=to speak in a way with certain interest in mind)
“pano para mangas” (=have legs)	“darse com un canto en los dientes” (=to conform with something)
“dor de cotovelo” (=only sour grapes)	“tener el corazón en un puño” (=to be anxious or afflicted)
“bater a bota” (=to die)	“andar com pies de plomo” (=to walk carefully)



5 Final Remarks

Our analysis makes it possible to conclude that the Spanish language has more frequent use of the IEs analyzed, although the difference between both languages is quite small. None of the expressions in question reached one occurrence per million words, which allows us to conclude that their frequency of use is not prominent in any of the languages.

Regarding textual dispersion, this appears to be the case in Spanish, although there are more occurrences in fiction. In Portuguese, the IEs appear more frequently in newspapers and politics. However, since the text genres and types are not exactly coincident in both corpora, further research on these data is in order.

Some IEs demonstrate great flexibility in verbal forms, with few of them not allowing variation, even in other elements. That is, from the list of IEs observed, there are few expressions with a high level of rigidity.

Thank you!