

On topic expressions in Spanish: Contrasting corpus and questionnaire data in the analysis of near-synonymy

Anton Granvik
HANKEN School of Economics

This paper addresses the issue of prepositional synonymy from the perspective of what I call topic expressions in Spanish. Topic is defined as the meaning expressed by prepositions such as *about* in a phrase like *talk about politics*, or *on* in the title of this paper. As such, this notion of topic is related to, but more restricted than the one used in discourse studies, where it is defined broadly as ‘what the sentence is about’ (Zubizarreta 1999: 4218; Halliday 1967: 212).

There are several topic expressions in Spanish, including the prepositions *de* and *sobre* as well as a series of compound prepositions (Lindstromberg 2010) such as *acerca de*, *en torno a*, *(con) respecto a/de*, *en cuanto a*. In this paper I will focus on the first four, which form a family of largely interchangeable expressions. They are frequently used as verbal and nominal complements to express the topic meaning, as shown in examples (1) and (2):

- (1) Siempre que me encuentro con mi hermana acabamos envueltos en una discusión de/sobre/acerca de/en torno a cómo hay que educar a los niños, si con amor o con dureza.
‘Every time I get together with my sister we end up arguing about how to raise your children, either with love or with discipline.’
- (2) Mi psicóloga me ha advertido de/sobre/acerca de/en torno a el peligro existente en negar las cosas que tan claramente forman parte de ti, aunque sean negativas y poco agradables.
‘My therapist has warned me about the danger of denying things that clearly are a part of you, even though they are negative and unpleasant.’

The aim of this paper is to find out what the relations are between these four expressions, that is, to what extent they are interchangeable (i.e. synonymous) and what differences exist in their usage.

Following Vanhatalo’s (2003: 362) conclusion that surveys provide significant additional information on word semantics that is inaccessible to corpus analysis, I set out approaching the situation from two angles. Firstly, I extracted 100 topic uses of each expression from the 100 million word Corpus del Español (Davies 2002-) and analyzed them according to both formal (genre, communicative situation, syntactic context, syntactic construction, etc.) and semantic criteria (word class, animate/inanimate agent/subject, semantic context, etc.). Secondly, I created a survey in order to complement the corpus data. The survey consisted of two slightly different questionnaires, in which 74 native speakers of Spanish were asked to complete given sentences by using the topic expression they found most appropriate for the given context. Examples (3) to (6) show some of the sentences included in the questionnaires.

- (3) En mi familia no hablamos _____ política en la mesa.
‘In my family we don’t talk _____ politics at the table.’
- (4) Me gustaría contarte un poco _____ lo que voy a hacer este verano.
‘I would like to tell you a bit _____ what I’m planning to do this summer.’
- (5) Necesitamos más información _____ el fenómeno antes de poder presentar nuestras conclusiones.
‘We need more information _____ the phenomenon before we can reach any conclusions.’
- (6) ¿Qué piensas _____ la explicación de Guardiola _____ la derrota ante el Chelsea?
‘What do you think _____ Guardiola’s explanation _____ the loss against Chelsea?’

Both data sets were analyzed both qualitatively and quantitatively. The results show that despite their similarity, each expression shows signs of constructional and contextual specialization. In the corpus data, *sobre*, *de* and *acerca de* are the most frequent, *en torno a* being relatively rare. While *acerca de*, *en torno a* and *sobre* are more frequent in combination with nouns than verbs, *de* is mostly used with verbs, especially verbs of communication (*hablar* ‘talk’) and cognition (*acordarse* ‘remember’).

The survey data builds on the corpus data. *Sobre* stands out as the default topic expression in Spanish, despite being highly polysemous, while *en torno a* is very rarely used by the informants, indicating its marginal status. *Sobre* and *acerca de* are often used in the same contexts and seem to constitute a subclass of their own, *acerca de* being the more explicit, marked alternative for *sobre*. Furthermore, there is an inverse relationship between the syntactic context and the choice of topic expression: the more marked the syntactic construction (for example, a topicalized interrogative structure), the less marked the topic expression.