The Effect of the Semantic Depth of SpanishVerbs on Processing Demands of Wh-questions

Ashlee Norris
-Orlando Alba Linguistics Award Winner

Wh-movement in Spanish is an intricate and much studied subject. Much work has been done on the underlying syntax that governs wh-movement, the following subject-verb inversion, and the various restrictions of wh-movement (adjunct island constraint, complex NP constraint, subject island constraint, etc.). This paper will look at another factor that determines the grammaticality of a wh-question: the semantic qualities of verbs introducing embedded clauses within wh-questions in which the matrix verb subcategorizes to a gap located in the embedded clause. We will see how the semantic weight of these verbs affects working memory and processing constraints, thereby affecting the acceptability of the wh-utterance.

The acceptability or grammaticality of a wh-utterance is governed, among other things, by the understandability of the filler-gap relationship. In the sentence

1. *¿Qué compró Juan?*

Qué represents the direct object of the verb *comprar*, which, in an unmarked declarative sentence, would be located after the verb to which it corresponds. In this construction *Qué* is the filler and its gap is found after *comprar*, in the unmarked direct object position. If in the course of the sentence that relationship is blurred or brought into question, the acceptability of the sentence decreases. This is not to say that the filler and gap cannot be separated by intervening words, so long as they do not violate such restrictions as the adjunct island constraint. Under proper stipulations the filler and gap may be located in different clauses, as in the sentence

2. *¿Qué dijo Juan que Maria compró?*

Here, the gap corresponding to the filler *Qué* is found after *compró*, *Qué* being in the matrix clause, the gap being in the embedded clause. This separation does not affect grammaticality.

However, although syntax is supposedly semantically blind (Jackendoff, 9), certain verbs will result in ungrammaticality in exactly the same syntactic context.

3. **¿Qué susurró Juan que María compró?*

Clearly the verb *susurrar* is closely related to the verb *decir* in (2). Why, then, the difference in acceptability? Both verbs report speech, and both are semantically plausible within the given context. The only difference between the verbs, then, is the semantic weight that each verb carries.

Breadth, Depth, and the Law of Inverse Proportionality

Let us first look at the verbs *decir* and *susurrar* individually. *Decir* reports a speech act, but out of context it is difficult to ascribe any additional meaning to this verb. The verb alone contains no information as to the emotion, importance, frequency, or volume of the utterance. While it may initially be assumed to report something that was uttered verbally, it may easily refer so something someone "said" in an email, for example. In fact, the verb *decir* is also flexible in terms of its corresponding agent, which is not necessarily animate, as in this sentence:

4. *La Biblia dice que Dios nos ama.* Because of its lack of semantic information, the verb *decir* can be applied to nearly any event during which information is transferred, whether verbally or in print. In Peirceian terms, the

"interpretable information associated with the sign", or in other words its "*quality* or its *signification* […] constitutes the sign's depth" (Roberston,1998, p.2). Thus, the verb *decir* is shallow; it contains very little interpretable information, but is very broad in its possible application.

In the verb *susurrar* we find quite the opposite. Much semantic information can be derived or inferred from the verb itself, without any aid from contextual clues. The verb alone contains information on the volume of the utterance, the agent (only animate agents can whisper), and the medium (one can whisper only vocally, not, for example, via email). On a more instinctual and inferred level, the verb *susurrar* seems to carry the idea of intimacy or secrecy, thus the verb conveys to some degree the emotional involvement of the participants of the whispered conversation. Therefore, the verb *susurrar* is semantically deep, but significantly narrower in its application than *decir*. Another way of expressing this relationship is to say that *decir* is more marked and *susurrar* less marked (Robertson,1998, p.23).

Robertson (1998) likens this inverse relationship between words like *decir* and *susurrar* to the Universal Law of Inverse Proportionality, defining the law in these terms: "Any division of a fixed quantity requires that the sum of its divided parts, however they are divided, be equal to the whole." Perhaps a definition more immediately applicable to the subject at hand would be that of C.S. Peirce, "one cannot devote a thing to a particular use without making it less available for other applications" (Robertson,1998, p.7). In terms of language, this means that as a word is given more semantic meaning or depth, it is necessarily restricted in its frequency, application, and breadth, and vice versa; as a word is given more applications or breadth, it will lose specificity and depth. For example, the English verb 'do' is so shallow that one could ask any individual engaged in any activity, "What are you doing?" On the other hand, only certain activities would permit a question like, "What are you building?" Here we see an illustration of a principle of markedness: the more marked option can be replaced with the less marked. Thus one may speak of the past or future in terms of the present, of plural in terms of singular, etc., and one may always replace *susurrar* with the less marked counterpart *decir.*

Wh-questions and processing

So far we have examined the differences in semantic breadth and depth in the verbs *decir* and *susurrar*. In this section we will discuss why those differences affect the grammaticality of wh-questions. Let us start by examining one way of interrupting the filler-gap relationship, thus decreasing the acceptability of wh-questions.

      5. a. *¿Qué dijo tu mamá de eso?*
        b. **¿Qué tu mamá dijo de eso?*

Goodall (2008) explains the reason for the difference in grammaticality between (5a) and (5b):
> One could reasonably speculate that they result from differences in ease of processing, since this factor is known to affect acceptability. Wh-dependencies place special demands on the processor, in that the filler must be held in working memory until the subcategorizing verb is processed, at which point a gap can be posited and linked to the filler (e.g. Gibson 2000). Importantly, processing this dependency and processing the reference of the intervening subject appear to make use of the same limited pool of resources (Warren and Gibson 2002) (p. 2-3).

Of the immense capacity of the human brain, only a limited amount of energy is devoted to understanding speech at any given time. When these resources are strained (i.e. holding a filler in

working memory while trying to process the reference of an intervening subject) the sentence is ungrammatical.

Also of note is the fact that the type of intervening subjects bears sway on the degree of grammaticality of the question. While outside of the Caribbean a question such as

6. *¿Qué tú dijiste?

is ungrammatical, it is more grammatical, more acceptable, and more easily understood than a sentence such as

7. *¿Qué el medico dijo?

This is because, as Goodall (2004) reports, "the more D-linked [discourse linked or lexical] an element is, the heavier its processing load" (4). Thus we find the continuum of acceptability of intervening subjects: 2p pronoun>3p pronoun>lexical (Goodall, 2008, p. 2). As the amount of information contained within the intervening subject increases, the demand on processing increases, thus decreasing grammaticality.

A similar phenomenon may be seen if we return to the initial examples of embedded clauses within wh-questions where the matrix contains the filler and the embedded clause contains the gap. (Note that for the remainder of this paper we will focus only on these types of wh-questions. We will also restrict the discussion and all data to questions beginning with *qué* which will demonstrate the most relevant filler-gap relationship for the purposes of this paper.)

8. a. *¿Qué dijo Juan que María compró?*

   b. *¿Qué susurró Juan que María compró?*

In (8b), the hearer must first hold the filler *Qué* in working memory, then process all the information contained in the semantically deep verb, *susurrar,* (animate agent, volume, transmitted vocally, reflecting a degree of intimacy or secrecy) before arriving at the subcategorizing verb *compró* and positing the gap. This results in a processing overload and ungrammaticality. This is why Mark Davies's Corpus del Español shows many instances of such questions using the verb *decir* while reporting no instances with *susurrar*.

To further illustrate this, consider the verbs of volition *querer, esperar,* and *desear*. Instinctively one may surmise that, of the three, *querer* is the least marked, the most shallow. This is confirmed by the *Real Academica Española*, which reports ten possible definitions, the assumption being that more definitions reflects greater breadth and, according to the Law of Inverse Proportionality, less depth. Using this measurement we see that *esperar* is less marked than *desear*, with six and three definitions, respectively. One would then expect to find a decreasing number of instances in wh-questions as we continue on the breadth-depth continuum: *querer>esperar>desear*. The corpus confirms this expectation, as seen in the following table.

Table 1. Verbs of Volition

Verb	Instances in wh-questions	Number of definitions
Querer	26	10
Esperar	5	6
Desear	1	3

Further evidence can be seen in the verbs of cognition *creer, pensar, suponer, presumir,* and *sospechar*, which increase in markedness in that order, again based on the number of definitions provided by the *RAE*. Note, however, that instinctual interpretation of markedness

would seem to correlate the data found by *RAE*. *Creer* and *pensar are* instinctually more common, *suponer* seems to hint semantically at the level of desire to know or admit something, *presumir* hints at level of evidence for the belief as well as attitude of the agent. *Sospechar* contains all of the above semantic depth, in addition to the negative connotation that one or more parties involved is displeased with the situation, or wishes the suspicion to be proven false—one is suspected of a crime, not of a heroic act. (Consider also this scenario: a wife walks in unexpectedly to her surprise birthday party. After the initial surprise her husband asks, "Did you suspect anything?" Here, the husband hopes that she did not; her knowledge would ruin his plan. Even in the best of situations, the verb *sospechar* carries a negative aspect.) The following table reflects the same results found with verbs of volition.

Table 2. Verbs of Cognition

Verb	Instances in wh-questions	Number of Definitions
Creer	30	7
Pensar	19	3
Suponer	1	5
Presumir	1	3
Sospechar	0	2

As expected, increased markedness leads to fewer definitions and fewer instances in wh-questions.

　　Here we see that the continuum *querer>esperar>desear* and the continuum *creer>pensar>suponer>presumir>sospechar>* follow closely the pattern found by Goodall, 2p pronoun>3p pronoun>lexical. As semantic depth is added, the grammaticality of the question decreases because demands on processing increase, straining the filler gap relationship.

　　At this point the observant reader will most likely have the following objection: could the decreasing frequency of semantically deep verbs in wh-questions be accounted for by the Law of Inverse proportions? In other words, as the verbs increase in semantic depth, one would assume that they become more restricted in breadth or application and thus show fewer instances of them overall, not just in wh-questions. Why, then, the appeal to processing constraints to explain the decrease in frequency?

　　If the Law of Inverse Proportions alone were enough to account for the data of tables (1) and (2), one would expect that roughly the same percentage of each verb's total instances found in the corpus would be in wh-questions. If processing constraints actually affect the acceptability of semantically deep verbs that introduce an embedded clause within wh-questions, one would expect this percentage to decrease. The following table shows the percentage of instances of each verb contained in wh-questions.

Table 3. Wh-question percentages

Verb	Total Instances	Wh-questions	Percentage
Volition			
Querer	58378	26	.0445

Esperar	17249	5	.029
Desear	6733	1	.0149
Cognition			
Creer	35327	30	.0849
Pensar	24705	19	.0769
Suponer	7825	1	.0128
Presumir	930	1	.108

Here the verb *presumir* is assumed to be an anomaly—the overall frequency has decreased to such a degree that even one occurrence in a wh-question results in a high percentage. Elsewhere we see, exactly as expected, that, as the markedness or depth of the verb increases, both the overall frequency of the verb (due to the Law of Inverse Proportionality) and the percentage of that verb's uses found in wh-questions (due to processing demands) decrease.

Conclusion

The difference between *decir* and *susurrar* may initially seem minimal, but when one considers all of the information contained in *susurrar* that is not conveyed by *decir*, its potential effect on processing demands becomes more apparent. All data recovered from the *Real Academia Española* and the Corpus del Español seem to support this hypothesis. Thus we may conclude, upon good evidence, that the semantic depth of verbs that introduce embedded clauses in wh-questions inversely affects the grammaticality of the question.

References

Davies, M. (2002). Corpus del Español (100 million words, 1200-1900s). Available online at
       http://www.corpusdelespanol.org
Goodall, G. (2008). Experimenting with wh-movement in Spanish. *Interactions in Romance:
       Selected Proceedings of the 38th Linguistic Symposium on Romance Languages.* 2008.
       Amsterdam.
Goodall, Grant. "On the Syntax and Processing of *Wh*-questions in Spanish." *Proceedings of the
       23rd West Coast Conference on Formal Linguistics,* 2004, Davis, California, 101-114.
Jackendoff, Ray. "Alternative Minimalist Visions of Language." *Proceedings from the Annual
       Meeting of the Chicago Linguistic Society* 2/2005:189-226. Web
Real Academia Española, *Diccionario de la lengua Española.* www.rae.es 19 April, 2010
Robertson, John S. "The Breadth and Depth of Markedness." Draft revised March 4, 1998. Web