SEMANTIC AMBIGUITY: THE ROLE OF NUMBER OF MEANINGS AND RELATEDNESS OF MEANINGS IN WORD PROCESSING

Juan Haro Rodríguez

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

Semantic ambiguity: The role of number of meanings and relatedness of meanings in word processing

Juan Haro Rodríguez

Doctoral Thesis

Universitat Rovira i Virgili
2018

Juan Haro Rodríguez

SEMANTIC AMBIGUITY: THE ROLE OF NUMBER OF MEANINGS AND RELATEDNESS OF MEANINGS IN WORD PROCESSING

DOCTORAL THESIS

Supervised by Dr. Pilar Ferré Romeu

Departament de Psicologia

UNIVERSITAT ROVIRA i VIRGILI

Tarragona

2018

UNIVERSITAT
ROVIRA i VIRGILI

Departament de Psicologia
Facultat de Ciències de l'Educació i Psicologia

Carretera de Valls, s/n
Campus Sant Pere de Sescelades
43007 Tarragona
Tel: 977 55 80 75
Tel: 977 55 80 88

I STATE that the present study, entitled “Semantic ambiguity: The role of number of meanings and relatedness of meanings in word recognition”, presented by Juan Haro Rodríguez for the award of the degree of Doctor, has been carried out under my supervision at the Department of Psychology of this university.

Tarragona, 25 d'abril de 2018

Doctoral Thesis Supervisor

Dra. Pilar Ferré Romeu

Acknowledgments

A Pilar Ferré. Por tu atención, comprensión y cercanía. Gracias por enseñarme tanto.

A José Eugenio García-Albea y a Rosa Sánchez-Casas.

A mis compañeros del Grupo de Investigación en Psicolingüística. A Toni.

A todos los alumnos y voluntarios que han participado en los experimentos.

A mis padres y a mi hermano. A mis abuelos y abuelas.

A Javi.

Y a Laura.

This doctoral thesis has been developed under a PhD grant received from the Universitat Rovira i Virgili (2015PMF-PIPF-16).

Table of contents

1. Introduction

1.1. A brief definition of semantic ambiguity

Words are strings of letters that have *a meaning*. However, some words have *more than one meaning*. These words are called ambiguous words, and they are present in all languages. For example, *bark* is an English ambiguous word, as it refers to both “the sharp cry of a dog, fox, or similar animal” and “the outside covering of the woody stems of plants”. Of course, *bark* is not a rare linguistic exception; there are a lot of ambiguous words: *bat, cell, cricket, thread, jam, prune, pitch,* etc. It is estimated that approximately 44% of the English words are ambiguous (Britton, 1978) and, importantly, the abundance of ambiguous words seems to be larger if we examine only the most commonly used words in English. In fact, Britton (1978) found that 85% of a sample of high frequency English words had more than one meaning. Thus, ambiguous words comprise a large proportion of the vocabulary of any language and, hence, they are very common in the everyday use of language. Given all this, it is not unreasonable to say that ambiguity is a key feature of language. For this reason, understanding how ambiguous words are processed and represented is necessary to provide a full explanation of human language.

On the other hand, semantic ambiguity is not a homogeneous phenomenon, since there are important qualitative and quantitative differences between ambiguous words. One of the most relevant is the relatedness between their meanings. The best-known type of ambiguity is represented by ambiguous words with unrelated meanings. This type of ambiguity is called *homonymy*. For example, *muñeca* is a Spanish homonymous word, as it has two completely unrelated meanings: “parte del cuerpo humano en donde se articula la mano con el antebrazo" and "figura de persona, hecha generalmente de plástico, trapo o goma, que sirve de juguete o de adorno". But there are also ambiguous words with related meanings. For instance, the ambiguous word *bueno* means "de valor positivo" as well as "gustoso, apetecible, agradable, divertido". This type of ambiguity is called *polysemy*, and the related meanings of a polysemous word are known as *senses*. It should be noted that polysemy is a more common phenomenon than homonymy. That is, there are more words with multiple related senses than with multiple unrelated meanings. For instance, Rodd, Gaskell, and Marslen-Wilson (2002) analyzed 4930 words of the Wordsmyth dictionary (Parks, Ray, & Bland, 1998), and found that while 84% of them had more than one related sense, there were only 7.4% with more than one unrelated meaning. In addition, polysemy can be further categorized into *metonymy* and *metaphor*. Metonymy is a type of polysemy in which the senses of a word are directly connected or are related to the same concept. For example, *agenda* could be considered a metonym, since its two meanings "libro, cuaderno o dispositivo electrónico en que se apunta, para no olvidarlo, aquello que se ha de hacer" and "relación ordenada de asuntos, compromisos o quehaceres de una persona en un período" are directly connected to the same concept. More examples of metonyms can be found, for instance, in words referring to the name of an animal and the meat of that animal (e.g., *pollo, cerdo*). On the other hand, metaphor defines a type of ambiguity where a word refers to both a literal and a symbolic meaning that are interrelated. Usually, a metaphorical word is created when the primary (and literal) meaning of a given word is eventually applied to refer to another concept with which it is symbolically related. For example, the word *gema* is a metaphorical word, given that its original meaning "nombre genérico de las piedras preciosas, principalmente de las denominadas orientales" (probably from the latin word *gemma*) was eventually used to define "persona o cosa tan valiosa como dicha piedra debido a su belleza o valor".

Finally, ambiguous words also vary in the frequency of their meanings and how equiprobable these meanings are. The frequency of a meaning indicates the extent to which it is familiar to speakers. In this way, ambiguous words can be classified into balanced (i.e., those with meanings of similar frequency) or unbalanced (i.e., those with meanings of distinct frequency) depending on the difference in frequency of their meanings. Unbalanced ambiguous words, which are more common than balanced ones, have a dominant meaning (the most frequent or familiar) and a subordinate meaning or meanings (the less frequent ones). For example, the meaning "de fuego" of the word *llama* is much more frequent than the meaning "animal". Thus, we could say that "de fuego" is the dominant meaning of *llama*, "animal" the subordinate one, and that *llama* is an unbalanced ambiguous word. On the contrary, both meanings of the word *heroína* (i.e., "droga" and "héroe de género femenino") are equally frequent, so that *heroína* could be considered a balanced ambiguous word.

1.2. How to measure semantic ambiguity

After this brief introduction to semantic ambiguity, I am going to describe how semantic ambiguity has been characterized in experimental research. Indeed, the first challenge that any researcher who wants to study semantic ambiguity should face is how to properly define what an ambiguous word is. It may seem a simple task, since an ambiguous word is *just a word having more than one meaning*. However, it is not clear which is the proper criterion to determine *how many meanings a word has*. Probably the most straightforward and accessible option is to look at the dictionary and count the number of entries for a given word (e.g., Jastrzembski, 1981; Rodd et al., 2002). Thus, assuming that each dictionary entry refers to a word meaning, those words with a single dictionary entry would be unambiguous and those with more than one entry would be ambiguous. For example, the word *pipa* could be classified as ambiguous since it has two entries in the Spanish Language Dictionary published by the Real Academia Española (RAE) (2014), one for the meaning “pipa de fumar” and the other for the meaning “semilla”. Instead, the word *jabón* could be classified as unambiguous because it has just a single entry (“producto de limpieza o higiene”). A further challenge is how to distinguish between ambiguous words with related meanings (i.e., polysemes) and those with unrelated ones (i.e., homonyms). Again, one way to do this is to look up the entries and senses of a word in the dictionary (e.g., Rodd et al., 2002). Usually, senses that are grouped within the same dictionary entry refer to related meanings of a word, whereas separate entries refer to unrelated meanings. For example, the polysemous word *diario* has a single entry in the RAE dictionary that lists all its related meanings, one for each sense. On the other hand, the homonymous word *bonito* has two dictionary entries. In the first entry we can find the meaning “pez” whereas in the second one there are listed several senses referring to the meaning “lindo”.

This so-called *dictionary approach* is an useful method because of the exhaustiveness of dictionaries in listing meanings. However, there are some concerns about its use. Probably, the most significant limitation is that not all the meanings that appear in the dictionary are known by the speakers. For instance, the word *aguja* has 32 dictionary definitions in the RAE dictionary, but many of them are unknown by the majority of speakers; for example, “pastel de hojaldre largo y estrecho relleno de carne picada o de dulce”. This is mainly because dictionaries include many definitions that are no longer used (or that are jargon), and that often do not incorporate new word meanings. Regarding the difference between homonymous and polysemous words based on the distinction between dictionary meanings and senses, the main concern is that neither all the words that have several dictionary entries are homonyms nor all those that have multiple senses within the same single entry are polysemes. Just to mention one example, all the unrelated meanings of the homonym *banco* are listed in the RAE dictionary as senses within the same single entry.

Thus, one could argue that the number of dictionary definitions is not a proper index of the number of meanings that are represented in the speakers' mind, and that the distinction between homonyms and polysemes cannot be made on the basis of the number of dictionary meanings or senses. To address this problem, an alternative possibility is to employ subjective approaches. Indeed, in order to classify words as ambiguous or unambiguous one could directly ask participants to provide the number of meanings of a word. This can be done in different ways. For example, by asking them to write down all the meanings they know of a word or to write the first word that comes to mind when they read a word (i.e., a lexical associate; e.g., "humo" o "semilla" for *pipa*, or "burbuja" for *jabón*) (e.g., Nelson, McEvoy, Walling, & Wheeler, 1980; Twilley, Dixon, Taylor, & Clark, 1994). Then, these responses are clustered according to the meaning to which they refer, and the number of meanings of the word is calculated. Another, simpler approach is to ask participants to indicate whether the word has one (1) or more than one meaning (2) (e.g., Kellas, Ferraro, & Simpson, 1988; Pexman, Hino, & Lupker, 2004). As such, average ratings close to 1 would indicate that the word is unambiguous, and ratings close to 2 would indicate that it is ambiguous. Similarly, in order to distinguish homonyms and polysemes, one might consider asking participants to indicate how related the meanings of a word are. To do so, some researchers provide participants with two sentences that include the same ambiguous word, but where each one refers to a different meaning. Then, participants have to indicate the degree of relatedness between the meaning of the word in each sentence (e.g., to indicate how related the two meanings of the word "muñeca" are, using a 1 to 7 Likert scale: "la niña jugaba con la muñeca" y "la niña se rompió la muñeca"). Another approach consists of asking participants to think about all the meanings of a word, and then to indicate how related they are by using a Likert scale from 1 (totally unrelated) to 7 (highly related).

It should be noted, however, that subjective approaches to estimate number of meanings and the relatedness between them are not devoid of problems. With respect to the number of meanings, the definitions or associates provided by the participants have to be later classified by judges. This is a task prone to errors and subjective biases (especially when it is necessary to determine whether a given definition or associate refers to a particular meaning), which is especially difficult to do with ambiguous words having related meanings. On the other hand, asking participants to indicate whether a word has one or more meanings often overlooks the differences in number of meanings between ambiguous words, so that, for example, a word with two meanings and another with four meanings would end up with a similar rating. Regarding the estimation of the degree of semantic relatedness between meanings, the subjective approach has also some limitations. For example, the relatedness ratings will be biased if the sentences used in the task refer to some meanings of the word, but not to others, or if participants do not know all the meanings of the word.

In view of the above, it is not surprising that the correct definition and assessment of semantic ambiguity has been a recurring issue in the more than 50 years of study of this phenomenon. Moreover, this concern is of great importance because it may have had a significant impact in the experimental results obtained in semantic ambiguity research. In brief, and before addressing this issue in more detail later (see next section), there are conflicting results between those studies that employed a dictionary approach (e.g., Rodd et al., 2002) and those that used subjective approaches (e.g., Pexman et al., 2004). With this in mind, the first objective of the present thesis is to provide a large database of Spanish ambiguous and unambiguous words, which will include several subjective and objective semantic ambiguity variables to properly categorize and characterize ambiguous words. This will allow us to examine whether the approach employed to categorize ambiguous and unambiguous words can influence the experimental results. It will also provide a large number of stimuli from which to select the materials for the rest of the experiments of this thesis.

1.3. Research on semantic ambiguity

In this section I will describe how semantic ambiguity has been examined in Psycholinguistic research. To do so, first I will present the main questions that have guided this field of research since its very beginning, and then I will depict how these questions have been addressed. The first of these questions focuses on the processing of ambiguous words: What happens when someone reads a string of letters that refers to more than one meaning? And the second one concerns their mental representation: How is represented in the mind the one-to-many mapping between the orthographic representation and semantic representations of an ambiguous word? It should be noted that addressing these questions has contributed not only to understanding how ambiguous words are processed and represented, but also to a deeper comprehension of human language. For instance, semantic ambiguity research has helped to elucidate how semantics and orthography interact during word recognition (e.g., Balota, Ferraro, & Connor, 1991; Hino, Lupker, & Pexman, 2002) or how a previous context affects meaning activation during reading (e.g., Swinney, 1991; Van Petten & Kutas, 1987).

The above questions have been addressed in several ways. On the one hand, some studies have examined how ambiguous words are processed within a context. For instance, by using semantic priming paradigms some authors have analyzed the time-course activation of ambiguous words that differ in their meaning relatedness (e.g. Kleipousniotou, 2002; Klepousniotou, Pike, Steinhauer, & Gracco, 2012; MacGregor, Bouwsema, & Klepousniotou, 2015). The usual procedure employed in these studies consists in comparing the recognition of polysemes (e.g., *aguja*) and homonyms (e.g., *banco*) when they are preceded by a related prime (e.g., *jeringa-aguja*, *dinero-banco)* with respect to when an unrelated word serves as a prime (e.g., *pelota-aguja, viento-banco*). Overall, the results of these studies have showed that polysemes exhibit a larger priming effect than homonyms, suggesting that polyseme meanings are faster retrieved than homonym meanings.

In a distinct line of research, several studies have explored the role of context and frequency of meanings in the retrieval of ambiguous word meanings (e. g., Kambe, Rayner, & Duffy, 2001; Martin, Vu, Kellas, & Metcalf, 1999; Onifer & Swinney, 1981; Simpson, 1981). In general, the evidence from these studies reveals that both variables influence ambiguous word processing. That is, if the context is congruent with one of the meanings of the ambiguous word, such meaning is retrieved easier (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982). In addition, when the context is biased towards the interpretation of the most frequent (dominant) meaning, the less frequent (subordinate) meaning is inhibited. Instead, a context biased towards the subordinate meaning would also activate the dominant meaning (Swaab, Brown, & Hagoort, 2003). For example, if participants were presented with a sentence containing the ambiguous word *pension* which was biased towards its dominant meaning (i.e., “ayuda económica”; e.g, “José tenía una pensión abundante”), and immediately after they were asked to decide if a string of letters corresponds to a word or not, they would recognize faster a word related with the dominant meaning of *pensión* (e.g., *renta*) than a control word (e.g., *mesa*). In contrast, a word related with the subordinate meaning of *pensión* (i.e., “establecimiento de hostelería”; e.g., *hotel*) would be recognized more slowly than the control word. Conversely, if the sentence was biased towards the subordinate meaning of *pensión* (e.g., “José tenía una pensión en el centro”), both the word related with the dominant meaning (i.e., *renta*) and that related with the subordinate meaning (i.e., *hotel*) would show a facilitation with respect to the control word.

On the other hand, many other studies have examined how ambiguous words are processed in isolation, that is, without context. These studies have mainly employed tasks that do not require meaning access to respond (e.g., lexical decision task) and, to a lesser extent, tasks where meaning access is required (e.g., semantic categorization tasks and sense judgement tasks). The results of the later reveal that ambiguous words are usually responded to more slowly than unambiguous words (see Eddington & Tokowicz, 2015, for a review). For instance, participants require more time to decide if *banco* is the name of an animal in comparison to an unambiguous word such as *mesa*. The explanation for this finding is that these tasks usually require a specific meaning of the ambiguous word to be activated. Consequently, this may cause a competition between the multiple meanings of the ambiguous words, leading to slower responses for ambiguous words in comparison to unambiguous words. However, most of the research on ambiguous word processing in isolation comes from studies that have employed the lexical decision task (hereafter LDT), a task in which participants are presented with strings of letters and asked to indicate whether the string is a word or not. The evidence obtained from LDT studies has largely contributed to explain ambiguous word processing, as well as to the development of word recognition models. For this reason, and because the

LDT is the task mainly used in this thesis, in what follows I will focus on the studies that examined ambiguous word processing using LDT.

1.3.1. Research on ambiguous word processing using LDT

The first LDT study that examined ambiguous word recognition dates from 1970, and was conducted by Rubenstein, Garfield and Millikan. However, their main objective was not to examine the recognition of ambiguous words *per se*, but to "test this view that the recognition of words involves consulting the internal lexicon" (page 487). If such a view were true, the authors expected experimental differences between various types of English words, for example, between words differing in their number of meanings. Participants of Rubenstein et al. (1970) were presented with 180 English words, 120 of which were unambiguous and 60 were ambiguous. In addition, 165 nonwords were also included. The results showed that ambiguous words were recognized faster than unambiguous words. Moreover, the number of meanings of ambiguous words affected their recognition, as ambiguous words with more than two meanings were recognized faster than those with only two meanings.

The study by Rubenstein et al. (1970) was a pioneer in the field of ambiguous word recognition, and paved the way for a fruitful line of research. In addition, it was the first proof of the existence of the so-called *ambiguity advantage*, the experimental finding that ambiguous words are recognized faster than unambiguous words in LDT. Rubenstein et al. suggested that such an effect might be due to the fact that ambiguous words are represented by multiple lexical entries, one for each of their meanings. Thus, under the assumption that word recognition implies consulting the internal lexicon, the more lexical entries a word has, the faster it will be recognized.

The effect found by Rubenstein et al. (1970) was later reproduced by Rubenstein, Lewis and Rubenstein (1971). However, in 1973, Clark showed that the ambiguity advantage found in these studies might be the result of an experimental artifact. Of note, this has been an ongoing feature of semantic ambiguity research: Each piece of evidence supporting the ambiguity advantage has been eventually challenged by subsequent studies. Namely, Clark (1973) demonstrated that Rubenstein et al. (1970, 1971) did a mistake in treating words as a fixed factor, and consequently their results could not be extrapolated to a new sample of words. Clark reanalized the data from Rubenstein et al.'s studies, but treating words as a random-effects factor, and found that the ambiguity effect was no longer significant. This result was also replicated by Forster and Bednall (1976), who found no differences between ambiguous and unambiguous words in a LDT when words were treated as a random-effects factor.

Considering the weaknesses of Rubenstein et al.'s studies, one might wonder whether the ambiguity advantage they found was a genuine effect or rather the result of an experimental confounding. To address this question, Jastrzembski and Stanners (1975) conducted a further LDT study, in which they tried to overcome the limitations of the work of Rubenstein et al. Following Clark's (1973) suggestion, they treated words as a random-effects factor. In addition, they suggested that the proper way to study the effect of ambiguity would be to compare ambiguous and unambiguous words differing in many meanings. Consequently, Jastrzembksi and Stanners compared words with a high number of meanings and words with a low number of meanings. Of note, Jastrzembksi and Stanners employed a dictionary approach to categorize their stimuli, so the number of words meanings was defined as the number of dictionary meanings. The results showed that words with a high number of meanings were recognized faster than words with a low number of meanings. In light of this finding, the authors argued that the lack of a significant ambiguity effect in Rubenstein et al.'s studies (when treating words as a random-effects factor) might have been produced by a difference not large enough between ambiguous and unambiguous words in number of meanings.

As mentioned above, studies conducted during the 1970s examined semantic ambiguity as a way of testing lexical access models. It was not until the 1980s when researchers began to analyze how the language system gets through the problem of processing a string of letters linked to multiple meanings, and how those meanings are represented in memory. To our knowledge, the first of these studies was that of Jastrzembski (1981). In their own words, “The effect for words with multiple meanings is important because a glance through an unabridged dictionary reveals that relatively few words have only a single meaning. Consequently, it would be expected that the mental lexicon would have properties to reflect the fact that a particular letter string can be associated with several, often quite unrelated, meanings. Understanding the means by which the multiple meanings of words are maintained and appropriately retrieved will considerably further our understanding of the mental lexicon and of the ability to comprehend language communication” (pages 278 and 279). The main aim of Jastrzembski (1981) was to analyze in depth the ambiguity advantage in order to find those variables that could modulate or cancel it. The results of the study showed a) that the ambiguity advantage was independent of frequency and, furthermore, was twice as large as that of frequency (Experiment 1); b) that it could not be attributed to differences between ambiguous and unambiguous words in terms of orthographic structure, in particular those related to the positional bigram frequency, and that c) that recency did not interact with the number of meanings, so that the effect cannot be due to ambiguous words being found more recently.

Jastrzembski (1981) made an important effort to study how ambiguous words are processed, and it set a strong precedent for the robustness of the ambiguity advantage. However, three years later, the study was challenged by Gernsbacher

(1984). She suggested that the ambiguity advantage found by Jastrzembski might be due to the fact that ambiguous words in that study were more familiar to participants than unambiguous ones. Indeed, given that ambiguous words have more than one meaning, speakers are likely to find them in more contexts than unambiguous words. Therefore, the experiential familiarity of ambiguous words would be greater than that of unambiguous words. To test this possibility, Gernsbacher (1984) presented words that varied orthogonally in familiarity and number of meanings, showing that familiarity facilitated RTs, but number of meanings did not. In addition, Gernsbacher raised serious concerns about the approach employed by Jastrzembski to measure the number of meanings of a word, which consisted of counting its number of entries in the dictionary (i.e., dictionary approach), as she argued that it is not a psychologically valid measure of number of meanings.

The two concerns raised by Gernsbacher were further addressed by Kellas et al. (1988) and Millis and Button (1989). In both studies, familiarity was matched between ambiguous and unambiguous words. In addition, they employed a subjective approach to measure word meanings, which consisted of asking participants to indicate the number of meanings a word has (Kellas et al., 1988) or to write down its meanings (Millis & Button, 1989). The results of both studies showed that subjectively-defined ambiguous words were recognized faster than subjectively-defined unambiguous words. Importantly, the same finding was obtained in many studies that employed subjective approaches to categorize ambiguous and unambiguous words during the 1990s and early 2000s (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Ferraro & Hansen, 2002; Hino & Lupker, 1996; Pexman & Lupker, 1999; Pexman et al., 2004; Piercey & Joordens, 2000).

Thus, almost all the evidence gathered from the late 1980s to early 2000s went in the same direction by supporting the ambiguity advantage, and all this evidence suggested that the effect was caused by the number of meanings of ambiguous words. However, in 2002, a study conducted by Rodd et al. presented strong evidence against this view. These authors examined the stimuli employed in some selected previous studies (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Millis & Button, 1989) and found that the ambiguous and unambiguous words of these studies did not differ in their number of meanings, but instead in their number of related senses. It should be noted that Rodd et al. employed a dictionary approach for this review, so they considered each dictionary entry to be a meaning and each dictionary sense to be a sense. Based on this, they suggested that the ambiguity advantage observed in previous studies might be due to the number of word senses rather than to the number of word meanings. To test this hypothesis, they orthogonally manipulated the number of senses and meanings of a set of words, and presented them in different LDT experiments. The results showed that, surprisingly, words with more than one meaning were recognized slower than words with one meaning. That is, they found an *ambiguity disadvantage* or *homonymy disadvantage*. By contrast, they observed that words with many related senses were recognized faster than words with few related senses, resulting in a *sense advantage* or *polysemy advantage*. To explain these results, Rodd et al. suggested that unrelated meanings would compete during word processing, whereas related senses would contribute together to word recognition (a more detailed explanation of this account will be provided in the next section [1.4. Models of ambiguous word recognition]).

The findings of Rodd et al. have been undoubtedly one of the greatest challenges to the ambiguity advantage, as they suggested that such a facilitation not only would not exist, but there would even be a disadvantage for ambiguous words with multiple meanings. Furthermore, they have broadened the study of semantic ambiguity by pointing out the relevance of the relatedness between ambiguous word meanings in word processing. This issue has motivated almost all the studies on ambiguous word recognition conducted since then to date. The aim of most of these studies has been to address to which extent the relatedness between meanings influences the processing of ambiguous words. In line with the results of Rodd et al., several studies have found a disadvantage for homonyms along with an advantage for polysemses (Armstrong & Plaut, 2008, 2011; Beretta, Fiorentino, & Poeppel, 2005; Tamminen, Cleland, Quinlan, & Gaskell, 2006). However, several other studies have found a similar facilitation for polysemes and homonyms with respect to unambiguous words (e. g., Hino, Kusunose, & Lupker, 2010; Hino, Pexman, & Lupker, 2006; Lin & Ahrens, 2010; Pexman et al., 2004). This line of evidence, which indicates that the relatedness between ambiguous word meanings has no effect on word recognition, is also congruent with previous studies in suggesting that multiple meanings facilitate word processing (e.g., Borowsky & Masson, 1996; Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Kellas et al., 1988; Millis & Button, 1989; Rubenstein et al., 1970).

Taken the above into account, it is not clear whether number of meanings facilitates or inhibits word recognition, and which is the role of relatedness of meanings in such process. Addressing these issues constitutes the main objective of the present thesis. My working hypothesis is that the conflicting evidence found in the literature may be partly due to methodological differences among studies, mainly regarding the approaches they employed to categorize ambiguous words. As summarized in the Section 1.2, distinct approaches have been used to measure number of meanings (hereafter, NOM), but the extent to which this has influenced experimental results is unknown. However, there seems to be a connection between the approach used and the results obtained in the studies conducted over the last two decades: While all the studies that have used a dictionary approach have observed a disadvantage for words with multiple meanings (Armstrong & Plaut, 2008,2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006), those that have employed subjective methods have found an ambiguity advantage (Hino et al., 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004). Thus, it is important to determine whether these conflicting results can be explained by differences among studies in the approach used. This objective will be addressed in this thesis in a series of experiments in which the categorization approach will be manipulated. In addition, I will also examine whether the relatedness between the meanings of ambiguous words (hereafter, ROM) influences word recognition. To assess whether ROM affects word processing, a LDT study comparing polysemes and homonyms will be conducted. In this experiment, neurophysiological correlates (Event Related Potentials [ERP]) will also be recorded.

1.4. Models of ambiguous word recognition

In the previous section I presented the evidence obtained from the studies that examined ambiguous word processing in isolation. These findings have led to the proposal of several models of ambiguous word processing, most of them based on or integrated within existing architectures (e.g., serial search, logogen, interactive activation, or PDP models). As new models have appeared since the pioneering work of Rubenstein et al. (1970), and some of the existing ones have been updated or disappeared, the explanations found in the literature are quite diverse. For instance, the first studies on ambiguous word recognition were aimed at testing the predictions of lexical access models based on serial search (e.g., Forster & Bednall, 1976; Rubenstein et al., 1970). These models suggested that the system performs a search in the mental lexicon to find the lexical entry corresponding to the word that has been presented as a stimulus. Word recognition would therefore take place once the system retrieves such entry. Under these models, facilitation for ambiguous words occurs because they would have a lexical entry for each meaning, allowing the system to find one of the multiple entries of these words faster than the single entry of an unambiguous word.

The explanation raised by Rubenstein et al. (1970, and also by Rubenstein et al., 1971) after observing for the first time the ambiguity advantage was based on those models. They proposed that the search process consists of different sub-processes. In the first sub-process, the system segments the presented stimulus into smaller units, that is, into letters. These units will then be used to select a subset of lexical entries. For example, after segmenting and identifying the letters "c-a-" of the stimulus *casa*, a subset of entries starting with those letters will be selected; for example, *casa*, *calor*, *caso*, *cabalgata*, etc. All the selected entries will be marked for further processing at a later stage, in which they will be compared one by one against the new information received from the process of letter segmentation. In this way, entries that do not match the new information are removed. For example, after segmenting the letter *s* of the stimulus *casa*, the entries *calor* and *cabalgata* will be deleted from the selected subset, leaving only *casa* and *caso*. It is important to note that the comparison of entries was assumed to be made at random, so that all entries have the same preference within this stage. This comparison process is performed in a loop until the entry that best fits the response criterion established by the participant is selected. Thus, assuming that ambiguous words have a lexical entry for each meaning, and that the comparison process is carried out in random order, the probability of selecting an ambiguous word entry will be greater than that of an unambiguous one (i. e., represented by a single entry).

On the other hand, the account provided by Jastrzembksi and Stanners (1981) to the ambiguity advantage was based on the logogen model (Morton, 1979). This model assumes that a logogen exists for every word a speaker has in his or her memory. Each logogen accumulates evidence from both the visual information provided by the stimulus (e. g. letters) and from contextual information. When the accumulated evidence reaches a given threshold, word recognition occurs. For example, during the presentation of the stimulus *casa*, all the logogens which are similar to the stimulus (e.g., *caso*, *calor*, *cabalgata*, besides *casa*) would accumulate evidence. While the structure and processing of such a model is significantly different from that of serial search models, the explanation for ambiguity advantage is quite similar: Ambiguous words would be represented by as many logogens as meanings; thus, it would be more likely that one of the logogens of an ambiguous word reaches the threshold of recognition faster than the logogen of an unambiguous word.

The 1990s was the decade of important progresses in the scientific understanding of how ambiguous words are represented and processed. These progresses were directly associated with a) the rise of the interactive model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982), b) the hypothesis that word meaning is available before word recognition (e.g., Balota et al., 1991), and c) the hypothesis that ambiguous words were not represented by multiple lexical entries, but by multiple semantic representations. In what follows, we will address these three points.

The interactive activation model is a model consisting of different levels of word processing and representation. Each of these levels includes representation units of varying degrees of abstraction: letter features, letters, and words. These levels are linked sequentially, in an order analogue to the degree of abstraction they handle. Furthermore, each level and the next one communicate bidirectionally. This communication is achieved through activation spreading from each level to the next one (feedforward) or to the previous one (feedback). This activation indicates the degree of correlation between the stimulus presented as input and the units of representation contained in each level (e.g., words or letters). Consequently, the amount of activation at one level would influence the amount of activation of every level to which it is connected. For example, during the presentation of the stimulus *asa*, all the lexical representations containing those letters (e.g., *casa* and *pasa*) would be activated. This activation, in turn, would influence the letter level, activating representations of some letters not present in the stimulus (e.g., *c* [*casa*], and *p* [*pasa*]). Finally, a word is recognized when its lexical unit of representation reaches a given activation threshold.

Based on the interactive activation model, Balota et al. (1991) suggested an explanation for the ambiguity advantage (hereafter, *semantic feedback hypothesis*). The authors' argument was that there was sufficient evidence in the literature indicating that the meaning of a word influences its recognition (e.g., *concreteness effect*, see James, 1975), and thus some meaning information would be available before a word is recognized. This was a significant difference from previous models, which assumed that word meaning was accessible only once the word was recognized (e. g., Forster & Bednall, 1976). However, the original interactive activation model did not include a semantic level of representations, making it impossible to explain the ambiguity advantage. To this end, Balota et al. suggested the addition of such a level, which would be bidirectionally connected to the word level. Thus, the greater the amount of semantic information associated with a word, the larger its influence on the word and letter levels. Consequently, since ambiguous words were assumed to have more than one semantic representation (one for each of their meanings), their semantic-to-orthographic feedback would be greater than that of unambiguous words, and thus they would reach the recognition threshold faster.

The semantic feedback hypothesis has obtained large support (e. g., Hino & Lupker, 1996; Hino et al., 2006, 2010; Pexman et al., 2004; Pexman & Lupker, 1999). In addition, it has also been able to account for the interactions observed between ambiguity and other variables, such as word frequency (Hino & Lupker, 1996; Pexman et al., 2004) and difficulty of nonwords in the LDT (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Piercey & Joordens, 2000). According to the semantic feedback hypothesis, high-frequency ambiguous words do not show an ambiguity advantage because their letter-to-word links are very strong. This would allow forward activation between these levels to be abundant and sufficient to reach the recognition threshold. Thus, there would be few time for semantic feedback to significantly influence the activation of inferior levels. Low-frequency words, instead, would have weaker letter-to-word links, so that semantic feedback would have a greater influence on inferior levels and thus on reaching the recognition threshold. A similar explanation might be provided for the interaction between ambiguity and nonword difficulty. The ambiguity advantage increases depending on the difficulty of nonwords because when nonwords are more wordlike the system would establish a higher recognition threshold. The purpose would be to reduce the number of errors in the task by requiring more evidence (i.e., a higher threshold) to distinguish between a word and a nonword. Hence, by increasing the recognition threshold, semantic feedback would have more time to influence the activation of inferior levels.

Apart from the accounts based on the interactive activation model, some PDP (Parallel Distributed Processing) models have also been able to accommodate the ambiguity advantage. For instance, Joordens and Besner (1994) tried to simulate such a facilitation effect by using Masson's (1991) distributed memory model. This model includes two word processing modules, one representing the orthography of the word and the other representing its meaning. The model was trained with stimuli representing ambiguous words (formed by an orthographic activation pattern and two semantic ones) and stimuli representing unambiguous words (formed by an orthographic and a semantic activation patterns). After the learning phase, during which the network tried to learn the matching pattern, the network was tested to see if it could retrieve one of the semantic activation patterns of an ambiguous word, and how long it would take. Interestingly, the network constantly retrieved a semantic pattern including information from both meanings. Furthermore, this blend semantic pattern was retrieved faster than the semantic pattern of an unambiguous word. Borowsky and Masson (1996) suggested that this blend semantic pattern would provide a high degree of familiarity to execute a correct response in LDT, and since it is reached quickly, it would explain the advantage for ambiguous words in LDT. To test this hypothesis, they measured the degree of familiarity during word recognition using a feature of Hopfield networks called energy. Network energy was measured as the sum of the activation of orthographic and semantic levels during word processing. This measure indicates the distance from the current network state to a learned activation pattern. In addition, word recognition would occur within the network when the energy measure reaches a given level. The model simulations showed that, after presenting an ambiguous word, a blend semantic pattern containing information from both meanings of the ambiguous word was quickly reached. As this blend pattern was similar to the two learned semantic patterns of the ambiguous words, these words elicited a higher level of semantic activation than that produced by unambiguous words. Thus, ambiguous words reached the energy criterion for word recognition faster than unambiguous words.

A somewhat different model to those proposed during the 1990s is that of Kawamoto, Farrar and Kello (1994). Its main difference with respect to the two mentioned above models is that the ambiguity effect was not assumed to be caused by a greater semantic activation for ambiguous words, but by a greater orthographic activation. This model is based on PDP principles, and consists of two processing modules, one for orthography and one for meaning. As in Borowsky & Masson (1996)'s model, ambiguous words were represented by an orthographic pattern and more than one semantic pattern, whereas unambiguous words were represented by an orthographic and a semantic pattern. After the training phase, the authors evaluated network performance by presenting only the orthographic pattern of a word, and observed that ambiguous words reached the orthographic criterion for lexical decision faster than unambiguous words. After examining the network to account for such a finding, they noted that the network had strengthened the connection weights of the orthographic units of ambiguous words. This would eventually help to overcome the inconsistency between the orthography and semantics of these words, thus facilitating ambiguous word recognition.

To sum up, there are three main models that account for the ambiguity advantage: the model of Kawamoto et al. (1994), the model of Borowsky & Masson (1996) and the interactive activation model represented by the semantic feedback hypothesis (Balota et al., 1991; Hino & Lupker, 1996). The evidence accumulated during the last 30 years has provided strong support for these models. However, there are some significant differences between them. On the one hand, the differences between the model of Kawamoto et al. and the other two models are that the former suggests that ambiguous words do not benefit from a greater amount of semantic activation, but from creating strong links at the orthographic level. On the other hand, the models that suggest that ambiguous words benefit from triggering a large amount of semantic activation (Balota et al., 1991; Borowsky & Masson, 1996; Hino & Lupker, 1996) also differ in two main aspects. The first is that the Borowsky and

Masson model assumes that the link between the orthographic and semantic level is unidirectional, so that the activation at the semantic level has no influence on lower levels. Instead, the semantic feedback hypothesis (Balota et al., 1991; Hino & Lupker, 1996) suggests that the link between the orthographic and semantic levels is bidirectional, thus semantic activation may influence orthographic processing. This leads to the second difference. According to the semantic feedback hypothesis, the semantic activation triggered by ambiguous words influences lower levels, causing the orthographic activation threshold required for word recognition to be reached quickly and resulting in a faster recognition for these words. In contrast, in the Borowsky and Masson's model, the advantage occurs because ambiguous words would increase global activation at the semantic level with no effect on orthographic processing.

All the models described above provide an explanation for the effect of multiple meanings on word recognition. There is a last model that was developed, instead, to explain the effect of the relatedness between ambiguous word meanings, that is, the difference in processing between polysemes and homonyms. This is the model of Rodd, Gaskell and Marslen-Wilson (2004), developed to explain the results of Rodd et al. (2002) and which can also account for the findings of subsequent studies (Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Tamminen et al., 2006). This model provides an explanation for both the homonymy disadvantage and the polysemy advantage. In this model, each meaning is represented as a basin within a semantic network. This basin is attractive, that is, it attracts the activation of the network towards itself. The network starts from a state of random activation, and word recognition takes place when the network accesses one attractor basin. Under this model, polysemes and homonyms differ in the structure and location of their semantic basins. Related senses are represented by neighboring attraction basins, forming a wide and shallow single basin; in contrast, unrelated meanings are represented by attraction basins located in distant regions of the semantic network. Thus, the semantic network should access faster the wide basin of a word with multiple senses than the narrow basin of a word with few senses, and this would explain the polysemy advantage. On the other hand, a blend of all the unrelated meanings of a homonym would be activated in the early stages of the semantic network settling. Then, the semantic network should escape from this blend state and move into one of the basins representing a meaning of the homonym. Assuming that moving away from the blend state involves a high processing cost, and that meanings compete during word processing, the model predicts a disadvantage in recognizing words with multiple meanings compared to words with one meaning.

The final objective of the present thesis is to test the predictions of the above described models. First, I will try to determine the cause of the NOM effect. On the one hand, I will examine whether ambiguous words benefit from creating strong orthographic links (e.g., Kawamoto et al., 1994) or from triggering a large amount of semantic activation (Borowsky & Masson, 1996; Hino & Lupker, 1996). To this end, ambiguous and unambiguous words will be compared in relation to some ERP components associated with orthographic processing (i.e., N200) and semantic processing (i.e., N400) during a LDT. On the other hand, I will contrast the models that suggest that ambiguous words trigger a large amount of semantic activation, to determine whether such semantic activation boost orthographic processing (Balota et al., 1991; Hino & Lupker, 1996) or not (Borowsky & Masson, 1996). This will be assessed with a task that are assumed to tap orthographic processing (i.e., a two-alternative forced-choice task). Finally, and following the tenets of the Rodd et al. (2004)'s model, I will examine whether ROM affects word recognition in a LDT experiment comparing polysemes and homonyms, where both behavioral and EEG data will be recorded.

1.5. Aims and organization of the thesis

So far, the introduction covered three relevant aspects of ambiguous word processing: 1) the different approaches used to define semantic ambiguity for experimental research purposes; 2) the experimental evidence obtained in semantic ambiguity research, and 3) how models of word recognition can account for such evidence. Along the lines, I have presented the main objectives of the present thesis, which are strongly linked to the above three aspects. Each of these objectives will be assessed in one or more than one of the studies that are included in the Experimental section. In what follows, I briefly present the objectives of the thesis and the studies in which each objective will be assessed:

1. To provide different objective and subjective measures to categorize Spanish ambiguous and unambiguous words (Study 1)
2. To assess whether the approach used to categorize ambiguous and unambiguous words has any influence on ambiguous word recognition (Study 2)
3. To examine the processing of ambiguous words which differ in their meaning relatedness by recording both behavioural and neurophysiological data (Study 3)
4. To test the predictions of some of the word recognition models that account for the ambiguity effects (Studies 3 and 4)

2. Experimental section

The following studies are included in this section:

1. Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics, 38,* 457-475. doi: 10.1017/S0142716416000266
2. Haro, J., & Ferré, P. (2018). Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition? *Journal of Psycholinguistic Research, 47*, 679-698. doi: 10.1007/s10936-017-9554-3
3. Haro, J., Demestre, J., Boada, R., & Ferré, P. (2017). ERP and behavioral effects of semantic ambiguity in a lexical decision task. *Journal of Neurolinguistics, 44*, 190-202. doi: 10.1016/j.jneuroling.2017.06.001
4. Haro, J., Comesaña, M, & Ferré, P. (submitted). Is there an orthographic boost for ambiguous words during their processing?

2.1. Study 1: Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics, 38,* 457-475.

2.1.1. Description of the study

There are different methods to categorize ambiguous and unambiguous words. Briefly, there is an approach based on the dictionary definitions, and another based on the definitions or ratings given by participants. Each of these methods has its advantages and disadvantages. In addition, the approach employed may influence experimental results. For these reasons, it is important to have different measures at hand to classify and characterize ambiguous words.

Although there are numerous databases of ambiguous English words (Azuma, 1996; Durkin & Manning, 1989; Ferraro & Kellas, 1990; Gawlick-Grendell & Woltz, 1994; Gee & Harris, 2010; Gorfein, Viviani, & Leddo, 1982; Griffin, 1999; Nelson et al., 1980; Nickerson & Cartwright, 1984; Panman, 1982; Twilley et al., 1994; Wollen, Cox, Coahran, Shea, & Kirby, 1980), there are few available normative studies in Spanish (Domínguez, Cuetos, & de Vega, 2001; Estevez, 1991; Fraga, Padrón, Perea, & Comesaña, 2017; Gómez-Veiga, Carriedo López, Rucián Gallego, & Vila Cháves, 2010). In addition, these databases have a limited number of words and provide only some ambiguity measures. Therefore, the objective of this study was to develop a database of ambiguous words in Spanish. This database should provide a significant number of ambiguous words and have a wide variety of subjective and objective measures of semantic ambiguity, both of number of meanings and of the relatedness between meanings. The development of this database was the step prior to addressing the remaining objectives of the thesis.

Applied Psycholinguistics, page 1 of 19, 2016
doi:10.1017/S0142716416000266

Semantic ambiguity norms for 530 Spanish words

JUAN HARO, PILAR FERRÉ, ROGER BOADA, and JOSEP DEMESTRE
Universitat Rovira i Virgili

Received: December 2, 2015   Accepted for publication: June 28, 2016

ADDRESS FOR CORRESPONDENCE
Juan Haro, Department of Psychology and CRAMC, University Rovira i Virgili, Crta. De Valls s/n., 43007-Tarragona. Spain. E-mail: juan.haro@urv.cat

ABSTRACT
This study presents semantic ambiguity norms for 530 Spanish words. Two subjective measures of semantic ambiguity and two subjective measures of relatedness of ambiguous word meanings were collected. In addition, two objective measures of semantic ambiguity were included. Furthermore, subjective ratings were obtained for some relevant lexicosemantic variables, such as concreteness, familiarity, emotional valence, arousal, and age of acquisition. In sum, the database overcomes some of the limitations of the published databases of Spanish ambiguous words; in particular, the scarcity of measures of ambiguity, the lack of relatedness of ambiguous word meanings measures, and the absence of a set of unambiguous words. Thus, it will be very helpful for researchers interested in exploring semantic ambiguity as well as for those using semantic ambiguous words to study language processing in clinical populations.

Language is full of words with more than one meaning (e.g., the word *bat*, which means *nocturnal flying mammal* as well as *wooden club used in baseball*), the so-called ambiguous words. It is estimated that approximately 44% of English words are ambiguous, a quantity that increases considerably for high-frequency words (Britton, 1978). Given their abundance, understanding how ambiguous words are represented and processed by the human mind is necessary for a complete explanation of human language processing and representation.

A unique property of ambiguous words is that they show a particular one-to-many mapping between their spelling and their meanings. This property makes them very interesting and useful for psycholinguistic research, especially for studying how words are processed and represented. For example, research on semantic ambiguity has contributed to elucidating how semantics and orthography interact during word recognition (e.g., Balota, Ferraro, & Connor, 1991; Hino, Lupker, & Pexman, 2002) or how a previous context affects meaning activation during reading (e.g., Swinney, 1991; Van Petten & Kutas, 1987). In addition, ambiguous words have also been employed in other fields of research. Just to name a few, they have been used to investigate implicit memory and false memories (Eich, 1984; Hutchison & Balota, 2005), depression (Hertel & El-Messidi, 2006), personality disorders (Mathews, Richards, & Eysenck, 1989), or autism (Hala, Pexman, & Glenwright, 2007).

Focusing on word recognition research, a common finding is that ambiguous words are recognized faster than unambiguous words in the lexical decision task (LDT; e.g., Hino & Lupker, 1996; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989). This processing advantage for ambiguous words in the LDT (called ambiguity advantage) has been a challenge for most of the models of word recognition, especially for classical models and parallel distributed processing (PDP) models. Classical models were unable to provide an explanation for this effect, because they assumed that the information related to the meaning(s) of a word (i.e., semantic information) should not affect word recognition. Thus, to account for the ambiguity advantage, some changes in these models were proposed (e.g., logogen models, Jastrzembski, 1981; and serial-search models, Rubenstein, Garfield, & Millikan, 1970). For instance, it was suggested that ambiguous words might be represented by more than one lexical entry, each one connected to a distinct meaning of the word. Consequently, it would be more probable to quickly select one of the lexical entries of an ambiguous word than to select the single entry of an unambiguous word, resulting in a processing advantage for ambiguous words.

In contrast, PDP models were not able to account for the ambiguity advantage. PDP models predicted a processing disadvantage for ambiguous words, given the assumption that ambiguous words have an inconsistent one-to-many mapping between their orthographic representation (e.g., *bark*) and their semantic representations (e.g., *the sharp cry of a dog* and *the outside covering of a tree*). This inconsistent mapping would produce a slower semantic coding for ambiguous words, probably due to a process of competition between their multiple semantic representations (e.g., Azuma & Van Orden, 1997). Thus, to provide an explanation for the ambiguity advantage consistent with PDP models, some authors have proposed that recognition times in the LDT do not reflect semantic coding, but orthographic coding (e.g., Hino & Lupker, 1996). Accordingly, because ambiguous words have multiple semantic representations, they would benefit from a large semantic feedback from the different meaning representations to their orthographic representation during word recognition, speeding up the orthographic coding and so then recognition and response times.

Although the above-mentioned accounts of the ambiguity advantage assume that this phenomenon is a result of words having multiple meanings, some studies conducted during the last decade strongly challenged this assumption (e.g., Armstrong & Plaut, 2011; Klepousniotou & Baum, 2007; Rodd, Gaskell, & Marslen-Wilson, 2002). Such studies examined words with multiple related meanings (i.e., polysemes; e.g., *newspaper*), words with multiple unrelated meanings (i.e., homonyms; e.g., *bat*), and unambiguous words by using the LDT. They reported a disadvantage for homonyms in comparison to unambiguous words, and an advantage for polysemes with respect to unambiguous words and to homonyms. Thus, these studies showed that the degree of relatedness between the different meanings of a word, rather than the number of meanings, facilitates word recognition. To account for this evidence, Rodd, Gaskell, and Marslen-Wilson (2004) suggested that the multiple semantic representations of a word with unrelated meanings would compete during word processing, slowing down their recognition in the LDT; in contrast, the related semantic representations of a word with related meanings would facilitate the recognition of the word. However, it should be noted that some findings are at odds with this relatedness of meanings advantage. Several studies have found an advantage of the same magnitude for polysemes and homonyms in comparison to unambiguous words (e.g., Hino, Kusunose, & Lupker, 2010; Hino, Pexman, & Lupker, 2006).

In view of these conflicting experimental findings, more research is needed for a complete understanding of how ambiguous words are processed and represented, as well as to identify the variables that influence the processes and mechanisms involved in their recognition. However, before conducting a study on semantic ambiguity, researchers have to face several important issues. Three of them are, in our view, particularly critical: (a) selecting a proper set of ambiguous words; (b) categorizing distinct types of ambiguous words, and (c) selecting a matched set of unambiguous words for comparison purposes. In order to aid researchers to address these issues, the aim of the present study is to provide a database of ambiguous and unambiguous Spanish words. In what follows, we will briefly address the above-mentioned critical issues, focusing on studies that have used homographs (i.e., words with the same spelling but more than one meaning; e.g., *fair*). Then, in the Methods section, we will explain in detail the development of the database.

Selecting an appropriate set of ambiguous words is essential for experimental research on semantic ambiguity. In order to do so, researchers have to determine whether a word is ambiguous or not. This task can be achieved by employing different approaches, either objective or subjective. For example, an objective method to categorize words as ambiguous or unambiguous consists in counting the number of entries of the words in the dictionary (e.g., Jastrzembski, 1981; Rodd et al., 2002). Using this approach (i.e., *the dictionary approach*), words with more than one entry are usually classified as ambiguous, whereas words with only one entry are classified as unambiguous. Regarding subjective approaches, a widely used method is to ask participants to generate definitions or lexical associates for a list of words. Then, the definitions or associates are grouped together according to the meaning to which they refer (Nelson, McEvoy, Walling, & Wheeler, 1980; Twilley, Dixon, Taylor, & Clark, 1994). The expected outcome is that definitions or associates for ambiguous words, but not for unambiguous words, will refer to more than one meaning. For instance, in Nelson et al. (1980) some of the associates generated for the ambiguous word *bass* were *fish*, *trout*, *drum, fiddle*, or *guitar*, which are related to their two distinct meanings: *sea perch fish* and *bass guitar*. Finally, another subjective approach is to ask participants to think about all the meanings of a string of letters and then choose a value in a 3-point scale: (0) *the word has no meaning*, (1) *the word has one meaning*, or (2) *the word has more than one meaning* (e.g., Kellas, et al., 1988; Pexman, Hino, & Lupker, 2004). With this approach, a quantification of number of meanings (NOM) is obtained: words with values close to 2 are classified as ambiguous, and words with values close to 1 are classified as unambiguous.

Each of the above-described methods has some strengths and weaknesses. Concerning objective measures, dictionaries are an exhaustive and standardized resource of word meanings. Furthermore, most of them list unrelated meanings in different entries, and related meanings under the same entry, facilitating not only the selection of ambiguous and unambiguous words but also the selection of homonyms (i.e., multiple entries) and polysemes (i.e., only one entry). Given that, dictionaries seem to be a very useful tool to easily select a set of experimental stimuli. However, a strong argument against employing dictionaries in semantic ambiguity research was pointed out by Gernsbacher (1984). She claimed that dictionary definitions should not be taken as a psychologically valid measure of meaning representation. She argued that speakers do not store in their memory all the meanings listed in a dictionary, but only a small sample of them. As an example, she showed that college professors could provide only 2 of the 30 dictionary definitions of the word *gauge*, 3 of the 15 dictionary definitions of the word *fudge*, and 1 of the 15 dictionary definitions of the word *cadet*. Apart from this criticism, another reason against using the dictionary approach is the fact that dictionaries evolve quite slowly (Lin & Ahrens, 2005). Thus, there are plenty of outdated definitions and they do not capture the emerging and gradual changes on the meaning of words.

With respect to subjective measures, asking participants to provide associates or definitions for a set of words would provide a better picture of which meanings are actually represented in the speakers' minds. The reliability of these subjective measures has been evidenced in the studies that have relied on them to select their materials. For instance, it has been demonstrated that associates of ambiguous words produce significant priming effects in LDT (i.e., a facilitation in response time to target words preceded by related [e.g., *hit-ball*] vs. unrelated [e.g., *doctor-ball*] prime words; e.g., Klepousniotou, Pike, Steinhauer, & Gracco, 2012). Nevertheless, a significant limitation of using associates is that to determine if a word is ambiguous or not as well as to count how many different meanings a particular word has, the responses of the participants have to be later examined and categorized by different judges. This task is time consuming, apart from being prone to errors and disagreements, because sometimes it is difficult to determine the meaning to which a particular response refers. For example, it is not clear if the word *land*, produced as an associate of the ambiguous word *plane*, is related to the *airplane* meaning or to the *flat surface* meaning. A similar case occurs when it has to be determined if two definitions provided for a word refer to exactly the same meaning, or instead, if they refer to distinct meanings. Consequently, this limitation may affect the ambiguous/unambiguous categorization of a word and the process of counting the number of different meanings. Finally, asking participants to assess the NOM a word has is a straightforward way to known if a word is ambiguous or not for speakers. In addition, it provides a numerical score of the word's ambiguity (i.e., a NOM rating), which facilitates the process of selecting stimuli for an experiment. However, there are two significant weaknesses of this approach. On the one hand, the criterion used by respondents to make their ratings is unknown. On the other hand, because ambiguous words can widely vary in NOM, this approach overlooks the differences between words with few meanings and words with many different meanings (Lin & Ahrens, 2005). Given the above-mentioned advantages and limitations of each approach, we considered it convenient to include several different ambiguity measures in the present database, both objective and subjective. Therefore, the database provides the number of dictionary definitions, the number of dictionary senses, NOM ratings, and an ambiguous/unambiguous categorization made by judges on the basis of word associates.

The second critical issue to face when conducting research on semantic ambiguity is how to categorize different types of ambiguous words. These words vary along several dimensions. For example, they vary in NOM (few vs. many meanings), in meaning frequency/dominance (balanced vs. unbalanced meanings), and in the degree of relatedness between their meanings (related vs. unrelated meanings). Dimensions such as NOM or meaning frequency/dominance have been extensively studied in the past (for a review, see Simpson, 1994); in contrast, relatedness of meanings (ROM) has only recently begun to gain experimental interest (for a review, see Eddington & Tokowicz, 2015). It is noteworthy that the results from this line of research have suggested that ROM is a relevant variable that determines the experimental effects obtained with ambiguous words.

In order to establish how ROM modulates word processing, a psychologically valid measure of this variable is needed. Researchers have employed different approaches to obtain it. For example, some authors ask participants to rate how related two definitions of an ambiguous word are (Azuma, 1996; Durkin & Manning, 1989). Similarly, other scientists ask participants to judge how related two sentences are, each sentence containing the same ambiguous word but varying the context in which it appears, one context being related to one meaning, and the other related to another meaning (e.g., *Do you have a goal in life? Liverpool won by a goal to nil*; Panman, 1982). These approaches provide a clear context for each word meaning, which forces participants to focus exclusively on a particular meaning when making their decisions. This facilitates the task for the participant and, at the same time, allows researchers to obtain a clean score of the relatedness between the meanings of an ambiguous word. However, the main limitation of these approaches concerns the selection of the meanings to be presented to the participant. If only a subset of all the meanings of an ambiguous word is presented, the ROM rating may be biased (e.g., in case of selecting only two meanings of a word with three meanings, when two of them are strongly related but that the third one is unrelated).

An alternative approach to obtain ROM ratings is by asking participants to choose the appropriate ROM value for an ambiguous word using a numerical scale (Hino et al., 2006). In this procedure, participants are presented with a list of ambiguous words, each one followed by a numerical scale comprising values from *unrelated meanings* to *related meanings*. Then, participants are required to think of all the meanings of each ambiguous word and to judge the relatedness of these meanings by selecting one of the values of the scale. The advantages and limitations of this approach are quite similar to those of NOM ratings. Its main advantages are that it is a straightforward approach to know if a word is polyseme or homonym, and that it provides a numerical value, which could be useful for selecting stimuli for psychological experiments. Regarding its limitations, two should be highlighted: the criterion used by participants to make a ROM decision is unknown, and it is questionable that participants consider all the meanings they know for a word before making the decision. Taking into account the strengths and weaknesses of the different ROM approaches, we decided to include two different ROM measures in the database, each one obtained through a distinct method.

Finally, as previously stated, the third critical issue when conducting research on semantic ambiguity is to choose a proper set of unambiguous words for comparison purposes. This is especially crucial for word recognition experiments, in which researchers typically compare ambiguous and unambiguous words with respect to the time needed by participants to recognize them. As pointed out above, a common finding in LDT is that ambiguous words are recognized faster than unambiguous words (e.g., Hino & Lupker, 1996; Jastrzembski, 1981; Lin & Ahrens, 2010; Millis & Button, 1989). Thus, in order to ensure that this advantage is produced by words' ambiguity, ambiguous and unambiguous words have to be matched on all the relevant variables that are known to affect word processing (e.g., word frequency, word length, familiarity, concreteness). This experimental control becomes even more important taking into account that several interactions between ambiguity and other variables have been reported. Namely, the ambiguity advantage has been observed for low-frequency ambiguous words, but not for high-frequency ambiguous words (e.g., Pexman et al., 2004); for abstract, but not for concrete words (e.g., Tokowicz & Kroll, 2007); and for neutral, but not for emotional words (Syssau & Laxén, 2012). For that reason, in the present database we included ratings for several variables, such as emotional valence, emotional arousal, concreteness, familiarity and age of acquisition.

To sum up, researchers should consider several issues before conducting semantic ambiguity experiments. To address these issues, it is very relevant to have large sets of normed stimuli. There are many databases for English ambiguous words (Azuma, 1996; Durkin & Manning, 1989; Ferraro & Kellas, 1990; Gawlick-Grendell & Woltz, 1994; Gee & Harris, 2010; Gorfein, Viviani, & Leddo, 1982; Griffin, 1999; Nelson et al., 1980; Nickerson & Cartwright, 1984; Panman, 1982; Twilley et al., 1994; Wollen, Cox, Coahran, Shea, & Kirby, 1980). However, the number of available norms in other languages is much lower. Focusing on Spanish, only three normative studies have been published to date (Domínguez, Cuetos, & de Vega, 2001; Estevez, 1991; Gómez-Veiga, Carriedo López, Rucián Gallego, & Vila Cháves, 2010). These databases are described in detail in what follows.

The norms of Dominguez et al. (2001) consist of 100 polysemous words. In this study, a sample of undergraduate students was asked to retrieve the different meanings of each word and to try to write three sentences with them. After that, the authors categorized the sentences according to the meaning to which the ambiguous word referred, and then they counted the number of distinct meanings. The database includes, for each word, the NOM, the number of responses assigned to each meaning, and the number of dictionary definitions. The database of Estevez (1991) consists of 152 homonymous words and 61 polysemous words. A sample of 104 participants was asked to write definitions for the homonyms, and 96 participants were asked to write sentences referring to the different meanings of the polysemes. All of them were undergraduate students living in the Canary Islands. The responses were categorized according to the most common meanings listed for each word in the *Spanish Language Dictionary* published by the Real Academia Española (RAE, 1984). The data set provides, for each word, the NOM, the percentage of responses for each meaning, and the number of dictionary definitions. Finally, the norms of Gomez-Veiga et al. (2010) contain 113 ambiguous words, which were evaluated by adults (ranging from 19 to 52 years old) and children (10 and 11 years old). Participants were asked to write down all the meanings they knew for each word in the same order that they retrieved them. The database includes the NOM of each word, the percentage of responses assigned to each meaning, and the familiarity of the word.

In light of the above, our opinion is that the available Spanish databases of ambiguous words lack some relevant information. Namely, (a) they only include one subjective measure of ambiguity, (b) they do not provide subjective ROM measures, and (c) none of them has a set of unambiguous words. Thus, the aim of the present study was to construct a normative database of ambiguous and unambiguous Spanish words that overcomes some of the limitations of the available Spanish databases. The present database is made up of 530 words. Four measures of ambiguity were obtained: ambiguity categorization based on lexical associates, subjective ratings on NOM, number of dictionary entries, and number of dictionary senses. In addition, we collected two subjective ROM measures that can aid researchers in selecting words to study the polysemy/homonymy distinction. Finally, subjective ratings for some relevant variables known to affect word processing were obtained in order to help researchers with the control of potentially confounding variables in ambiguity studies.

METHOD

Overview of the procedure used to construct the database

The database was developed following a series of steps. First, 641 Spanish words were selected from the *Spanish Language Dictionary* published by the RAE (2014). Second, these words were classified as ambiguous or unambiguous according to the associates generated by a group of participants. This classification was first made by four judges (the authors) and then validated by NOM ratings obtained from a different set of participants. Third, two distinct ROM measures for the ambiguous words were obtained. Fourth, ratings of concreteness, familiarity, emotional valence, arousal, and subjective age of acquisition were collected. In what follows, we will explain in detail each step.

Participants

A total of 1,213 undergraduate students from the Universitat Rovira i Virgili (mean age = 20.67 years, *SD* = 4.79; 933 females, 280 males) participated in the study in exchange for academic credits. All were highly fluent native speakers of Spanish. Each participant filled in a set of questionnaires.

Item selection

We used an automated script for randomly selecting approximately 1,000 words from an electronic version of the RAE dictionary. This set of words was selected according to some criteria; namely, they should be “potentially unambiguous” or "potentially ambiguous." Words were considered to be potentially unambiguous if they had only one dictionary entry and five or fewer dictionary senses (e.g., *aeropuerto* "airport": 1 entry and 1 sense), and words were considered to be potentially ambiguous if they had either more than one dictionary entry or more than five dictionary senses (e.g., *verso* "verse": 2 entries and 4 senses; *perfil* "profile": 1 entry and 10 senses). The number of senses for each word was computed as the total number of definitions listed in all the dictionary entries of the word. After removing words with very low frequency, the initial stimuli set consisted of 641 words: 392 potentially ambiguous words and 249 potentially unambiguous words. Words in the set had a lexical frequency ranging from 0.02 to 2,125 occurrences per million ($M = 59.59$, $SD = 175.79$; Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), their length was between 3 and 14 letters ($M = 6.46$, $SD = 2.03$), and the most common part of speech was noun (83.3%), followed by verb (7.6%) and adjective (7%). The number of dictionary entries of the words was between 1 and 6 ($M = 1.23$, $SD = 0.60$), and their number of senses ranged from 1 to 49 ($M = 8.41$, $SD = 7.01$).

Procedure

Measures of ambiguity. Following previous studies (e.g., Nelson et al., 1980; Twilley et al., 1994), we opted for collecting participants' responses in a free-association task as a subjective measure of ambiguity. All the words from the stimuli set were randomized and listed in nine questionnaire versions. Then, 236 participants (mean age = 20.4 years, $SD = 4.15$, range = 17–47 years; 46 males) were asked to complete a free-association task, in which they had to write down the first word that came to mind after reading a cue word. Based on the associates generated by the participants, four judges (the authors) categorized individually each word of the stimuli set as ambiguous or unambiguous. Words were classified as ambiguous if their associates were related to distinct meanings (e.g., for the word *mouse*, participants produced the following associates: *computer*, *cheese*, *mickey*, *cat*, *hamster*, and *keyboard*). Instead, words were classified as unambiguous if all their associates were related to the same meaning (e.g., for the word *major*, participants produced the following associates: *council*, *town*, *politician*, and *president*). Idiosyncratic responses were removed. It is important to note that only words for which full consensus was reached between the four judges were selected; consequently, 111 words that did not fulfill this criterion were removed from the stimuli set. Altogether, of the remaining 530 words, 386 were categorized as ambiguous and 144 as unambiguous.

In order to test the validity of the judges' categorization, and to provide another measure of ambiguity, we collected NOM ratings following the approach used in previous studies (e.g., Ferraro & Kellas, 1990; Pexman et al., 2004). The 530 words were randomized and listed in 10 questionnaire versions. In addition, 27 nonwords (e.g., *bresio*) were included as fillers in each version. Then, a group of 235 participants (mean age = 21.2 years, $SD = 5.53$, range = 18–57 years; 44 males) was asked to decide if a given character string had *no meaning* (coded as 0), *one meaning* (coded as 1), or *more than one meaning* (coded as 2) using a 3-point scale.

ROM measures. After classifying the words as ambiguous or unambiguous, the next step was to obtain an index of relatedness between the different meanings of each ambiguous word. To this end, we collected two different ROM measures (ROM_1 and ROM_2). A novel approach was used to obtain the ROM_1 measure. Two of the associates generated by the participants were selected for each ambiguous word of the stimuli set. One of the associates was the one with the highest response frequency (*modifier1*), and the other was the associate related to a different meaning than that for modifier1 with the highest response frequency (*modifier2*). For example, the two associates selected for the ambiguous word *siren* were *sea* (the associate of *siren* with the highest number of responses) and *ambulance* (the associate of *siren* related to a meaning different from that of *sea* with the highest number of responses). With these materials, a questionnaire was constructed. In each page of the questionnaire, the ambiguous word was paired in one line with modifier1, and in another line with modifier2 (e.g., *siren-sea* and *siren-ambulance*) followed by a 9-point scale, ranging from 1 (*unrelated meanings*) to 9 (*same meaning*). Participants were asked to rate to which degree the meaning of the ambiguous word paired with modifier1 (e.g., in *SIREN-sea*, the meaning *sea nymph*) and the meaning of the ambiguous word paired with modifier2 (e.g., in *SIREN-ambulance*, the meaning *warning alarm*) were related. It is important to note that we also included unambiguous words (e.g., *umbrella*), paired with the two associates with the highest response frequency (e.g., *rain* and *water*), with the aim of obtaining an additional validity measure of the judges' categorization.

All the words from the stimuli set (i.e., 386 ambiguous and 144 unambiguous words) were randomized and listed in eight questionnaire versions.[1] Each version consisted of, approximately, two thirds of ambiguous words and one third of unambiguous words. Detailed instructions and examples were included at the beginning of each version (see Appendix A). One hundred and eighty-three participants (mean age = 19.84 years, *SD* = 5.52, range = 18–68 years; 68 males) completed the questionnaires. Finally, the ROM_1 rating for each word was calculated by averaging the scores of the participants who had assessed it.

In addition, we collected a second ROM measure (ROM_2) by using the same approach used in a previous study (Hino et al., 2006). We randomized and listed the 386 ambiguous words in nine questionnaire versions, and then we asked 215 participants (mean age = 20.65 years, *SD* = 5.08, range = 18–56 years; 31 males) to think of all the meanings of each word and then to rate the relatedness of the meanings by choosing the appropriate ROM value on a 7-point scale, which ranged from *unrelated meanings* (1) to *highly related meanings* (7).

Other variables. We also collected values of concreteness, familiarity, emotional valence, arousal, and subjective age of acquisition for the 530 words of the database from different sources. Familiarity and concreteness ratings for 412 words were obtained from EsPal (Duchon et al., 2013). Given that the values of concreteness and familiarity for the remaining 118 words were not included in EsPal, they were provided by an additional group of participants. Regarding concreteness, the 118 words were randomized and listed in two questionnaires. Forty-four participants (mean age = 22.11 years, *SD* = 6.31, range = 18–56 years; 12 males) rated the words using a 7-point scale, which ranged from 1 (*minimum level of concreteness*) to 7 (*maximum level of concreteness*). With respect to familiarity, the 118 words were randomized and listed in three questionnaires. A group of 73 participants (mean age = 20.44 years, *SD* = 2.54 range = 17–30 years; 19 males) was asked to rate the familiarity of the words using a 7-point scale, which ranged from 1 (*not familiar at all*) to 7 (*very familiar*). To ensure comparability between our ratings and those provided by EsPal, we used the same instructions and rating scales.

In a similar way, emotional valence and arousal ratings were taken from different sources. The values for 191 words were obtained from Guasch, Ferré, and Fraga (2015). The values for the remaining 339 words were obtained from an additional sample of respondents, who were presented with the same instructions and rating scales as those used by Guasch et al. (2015). Concerning emotional valence, the 339 words were randomized and listed in four questionnaires. Then, 103 participants (mean age = 20.76 years, *SD* = 2.46, range = 18–29 years; 25 males) were asked to rate emotional valence using the self-evaluation valence scale of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994). The SAM provides a sequence of images to help participants to rate words in an affective scale. The scale for the emotional valence dimension ranges from 1 (*strongly negative*) to 9 (*strongly positive*). Regarding arousal, the 339 words were randomized and listed in four questionnaires. Participants were provided with the arousal scale of the SAM, which ranges from 1 (*not arousing at all*) to 9 (*strongly arousing*). Ninety-four participants (mean age = 21.72 years, *SD* = 4.87, range = 18–61 years; 24 males) answered the arousal questionnaires.

Finally, we collected subjective age of acquisition ratings. As for the above-mentioned variables, different sources were used. Age of acquisition ratings for 418 words were obtained from Alonso, Fernández, and Díez (2015). The remaining 112 words were listed in a single questionnaire and presented to a group of 30 participants (mean age = 18.8 years, *SD* = 1.06, range = 18–22 years; 11 males). We used the same instructions and scale as Alonso et al. (2015). Respondents were asked to estimate the age at which they learned each word by using an 11-point lineal scale. In this scale, a value of 1 indicates an age *lower than 2 years old*, numbers ranging from 2 to 10 indicate learning ages from *2* to *10 years*, and a value of 11 indicates that the learning age for the word was *11 years or older*.

RESULTS

The complete database can be downloaded from http://psico.fcep.urv.cat/exp/files/haro_et_al_database_2016.xls. A summary of the stimulus characteristics of the database is shown in Table 1. In what follows, we will provide some reliability and validity data for the measures included in the database, as well as the main results of the analyses conducted to fully characterize the normed words.

Data trimming

All the questionnaire versions for each measured variable were administered to at least 25 participants. Some respondents were rejected after applying a trimming procedure. To do so, we examined the data from the questionnaires in order to identify random or aberrant response patterns. For example, using graphical

Table 1. *Descriptive stimulus characteristics for the 386 ambiguous words and the 144 unambiguous words (classified according to the judges' categorization measure)*

	NOM	Ent	Sen	ROM_1	ROM_2	AoA	Fam	Con	Val	Aro	Freq	Lett
						Ambiguous Words						
Mean	1.66	1.30	9.77	3.10	3.66	6.68	5.49	4.39	5.07	4.42	70.5	6.20
SD	0.23	0.66	6.79	1.53	1.18	1.88	0.79	0.94	1.03	0.94	203.8	1.89
						Unambiguous Words						
Mean	1.13	1.03	3.38	7.19		6.92	5.31	5.60	4.94	4.37	20.7	7.53
SD	0.15	0.23	3.49	1.01		2.36	1.04	0.88	1.16	0.95	66.2	2.20

Note: NOM, Number of meanings; Ent, number of dictionary entries; Sen, number of dictionary senses; ROM, relatedness of meanings; AoA, age of acquisition; Fam, familiarity; Con, concreteness; Val, valence; Aro, arousal; Freq, frequency (number of occurrences per million); Lett, word length (in letters).

procedures for assessing person-fit, we discarded those participants with a pattern of responses with almost no variation, because this suggests that they used the same value of the scale for nearly all the words. Furthermore, we also computed a personal correlation coefficient between each participant's data and the mean, and we excluded those participants with negative values or values near to zero. Values near to zero would suggest that the participant responded randomly to the questionnaire, whereas negative values would suggest that the participant understood the scale in the opposite direction. As a result of applying this trimming procedure, the number of valid responses for each variable was, at least, 21 for NOM, 19 for ROM_1, 19 for ROM_2, 23 for emotional valence, 22 for emotional arousal, 24 for familiarity, and 22 for concreteness.

Reliability

We assessed the reliability of each variable through a split half intergroup procedure. This method consists of dividing the data from each questionnaire into two sets of scores, one set from even items and the other set from odd items. Then, the correlation between the two sets is calculated, giving an index of agreement between participants' scores. Correlations were corrected using the Spearman–Brown formula. Mean intergroup correlation values were $r = .85$ for NOM ratings, $r = .91$ for ROM_1, $r = .81$ for ROM_2, $r = .84$ for emotional valence, $r = .96$ for emotional arousal, $r = .98$ for familiarity, $r = .85$ for concreteness, and $r = .88$ for age of acquisition. All the correlations were positive and significant (all ps $<$ .001), supporting the reliability of the data.

Relationship between ambiguity measures

We obtained two objective measures of ambiguity (number of dictionary entries and number of dictionary senses) and two subjective measures of ambiguity (NOM

Table 2. *Correlations between the measures of ambiguity and the ROM measures*

Measure	1	2	3	4	5	6
1. Dictionary entries	—					
2. Dictionary senses	.14**	—				
3. Judges' categorization	.21***	.42***	—			
4. NOM	.24***	.50***	.75***	—		
5. ROM$_1$	−.27***	−.39***	−.79***	−.73***	—	
6. ROM$_2$	−.35***	.05	—	−.29***	.60***	—

Note: ROM, Relatedness of meanings; NOM, number of meanings.
**$p < .01$. ***$p < .001$.

ratings and the judges' ambiguous/unambiguous categorization based on word associates). As can be seen in Table 2, objective and subjective measures were positively and significantly correlated. This is of relevance because it supports the validity of the subjective measures that we collected.

In addition, it is important to note that ambiguous and unambiguous words categorized by judges differed in dictionary entries, t (528) = 4.82, $p < .001$, dictionary senses, t (528) = 10.76, $p < .001$, and NOM ratings, t (528) = 25.79, $p < .001$.

Relationship between ROM measures

Two ROM measures were collected: ROM$_1$ and ROM$_2$. ROM$_1$ ratings were obtained using a novel approach, whereas to get ROM$_2$ ratings, we followed the procedure used in a previous study (Hino et al., 2006). The validity of our novel approach was supported by the high correlation between ROM$_1$ and ROM$_2$, r (528) = .60, $p < .001$.

Moreover, it should be noted that ROM$_1$ ratings were requested not only for ambiguous words but also for unambiguous words. Unambiguous words were included in ROM$_1$ questionnaires with the aim of obtaining additional support for the judges' ambiguous/unambiguous categorization. Thus, we expected lower ROM$_1$ values (i.e., closer to the "unrelated meanings" value) for words classified as ambiguous than for words classified as unambiguous. Accordingly, ambiguous words had significantly lower ROM$_1$ ratings ($M = 3.10$) than unambiguous words ($M = 7.19$), t (528) = 29.69, $p < .001$. Likewise, this explains the negative correlation between ROM$_1$ and the ambiguity measures (see Table 2). In addition, ROM$_2$ was negatively correlated with both NOM ratings and number of dictionary meanings. Because ROM$_2$ ratings were only obtained for ambiguous words, this negative correlation indicates that ambiguous words with related meanings were considered by participants as having a lower NOM than ambiguous words with unrelated meanings.

Comparisons with other Spanish databases of ambiguous words

To assess the validity of the present database, we compared our ratings with those from other Spanish databases of ambiguity (Domínguez et al., 2001; Estevez, 1991; Gómez-Veiga et al., 2010), focusing on overlapping words across databases (range = 34–78 words). Because such norms only provide one subjective measure of ambiguity (i.e., NOM of the word) and no ROM measure, we could only examine the correlations between our NOM ratings and their measure of NOM. Significant correlations were observed between our data and those from Domínguez et al. (2001), r (32) = .46, p = .007, as well as between our data and those from Gómez-Veiga et al. (2010), r (57) = .39, p = .003. However, our ratings did not correlate with those of Estevez (1991), r (76) = .09, p = .44. This may be due to methodological differences between studies. In the norms of Estévez (1991), unlike in the other two studies, the responses of the participants were categorized according to the most common meanings provided by the dictionary. This is of relevance, because it might be possible that some responses related to other meanings were ignored, resulting in a bias in counting the meanings. In addition, Estevez relied on the 1984 edition of the RAE dictionary to categorize participants' responses. There is a difference of 30 years between that version and the current RAE edition. Thus, it is likely that the most common meanings provided by the 1984 edition are slightly different from those included in the current edition. Finally, because the meanings of the words may evolve by the speakers' use of language, it is also likely that the responses of participants in 1991 were somewhat different from those provided by respondents in the present study.

Conclusion

The aim of this study was to provide normative data for a set of ambiguous and unambiguous Spanish words. The resulting database is made up of 530 Spanish words rated on several subjective measures of ambiguity. It also contains dictionary measures of ambiguity (number of dictionary entries and number of dictionary senses), ROM measures, and values of some relevant lexicosemantic variables (i.e., concreteness, familiarity, emotional valence, arousal, and subjective age of acquisition).

The database includes two subjective measures of ambiguity and two subjective ROM measures. Although the two measures of ambiguity were obtained using distinct approaches (judges' categorization based on word associates and NOM ratings), they were highly intercorrelated. In addition, both measures were correlated with number of senses as well as with number of entries in the dictionary. Furthermore, some significant correlations were observed between our NOM ratings and the ratings from other Spanish databases of ambiguous words. Altogether, these correlations provide support for the validity of the subjective measures of ambiguity. Similarly, ROM measures were obtained using distinct methods. A novel approach was used for collecting ROM_1 ratings, whereas ROM_2 ratings were obtained following the method employed in a previous study (Hino et al., 2006). Despite the differences between both methods, the two ROM measures were highly intercorrelated, supporting the validity of the novel method. In light of this evidence, we consider that both ambiguity and ROM measures could be reliably used in semantic ambiguity research. In particular, they may be of value to assess the influence of the NOM and of the ROM on word recognition, helping to refine models of word recognition and to test their hypothesis about the processing and representation of ambiguous words.

Furthermore, the database includes some subjective ratings for concreteness, familiarity, emotional valence, arousal, and subjective age of acquisition that are not provided by any other available Spanish database. Namely, we collected new concreteness and familiarity ratings for 118 words, new emotional valence and arousal ratings for 339 words, and new age of acquisition ratings for 112 words. Given that, the present norms will be useful not only for researchers interested in studying semantic ambiguity but also for researchers looking for controlling such variables in their experiments.

In contrast, it should be highlighted that the present database might have some applications beyond psycholinguistic research. Ambiguous words have been used, among other situations, to study clinical populations. For example, by presenting homographs with both personal and impersonal meanings (e.g., *close* and *console*), Hertel and El-Messidi (2006) examined the tendency of people in depressed or dysphoric states to interpret ambiguous events as personally relevant. Similarly, semantic ambiguity has been employed as a tool to investigate the negative interpretation bias related to anxiety disorders. This has been done by examining the response of people with anxiety disorders when they were presented with ambiguous words having both threat and nonthreat meanings (e.g., *patient*; Hayes, Hirsch, Krebs, & Mathews, 2010). A further application can be found in the study of deficits in contextual processing in children with autism. For instance, Hala et al. (2007) investigated whether children with autism are able to make use of the meaning of semantic primes to interpret the meaning of ambiguous words (e.g., *pencil-lead*). Thus, researchers interested in those or similar applications of semantic ambiguity could benefit from the present database.

In conclusion, this normative study overcomes some of the limitations of the published databases of Spanish ambiguous words: the scarcity of measures of ambiguity, the lack of subjective ROM measures, and the absence of a set of unambiguous words. As such, the database will be useful for researchers interested in studying semantic ambiguity as well as in their applications in different situations and/or populations. It will be especially valuable to assist researchers in categorizing ambiguous and unambiguous words, in categorizing ambiguous words that differ in the relatedness of their meanings, and in preventing any experimental confound due to uncontrolled variables.

APPENDIX A

Instructions for the ROM$_1$ questionnaire, translated from the original instructions in Spanish

As you may already know, there are different types of ambiguous words. There are ambiguous words with unrelated meanings, and ambiguous words with related meanings. For example, *siren* is an ambiguous word with two unrelated meanings: it may refer to *a warning device* or to *a sea nymph*. On the contrary, the ambiguous word *balloon* has several related meanings: it may refer to *a rubber bag that can be inflated with gas* or to *a bag of strong, light material filled with a gas lighter than air so as to rise through the air*.

Below, we present you with a word assessment questionnaire. In this questionnaire, you will find both ambiguous and unambiguous words. Each word to be assessed will appear together with two other words, which we will call *modifiers*. Modifiers are used to direct your attention toward a specific meaning of the first word. Your task consists of comparing the meaning of the word to be evaluated when it is accompanied by the first modifier, to the meaning of that word when it is accompanied by the second modifier. In order to make your ratings, you have to use a scale from 1 to 9, where 1 indicates that the two meanings are completely different, and 9 that the two meanings are exactly the same. We encourage you to use all the values on the scale.

Original instructions in Spanish for the ROM$_1$ questionnaire

Como ya sabrás, hay distintos tipos de palabras ambiguas. Existen palabras ambiguas con significados no relacionados y palabras ambiguas con significados relacionados. Por ejemplo, *sirena* es una palabra ambigua con significados no relacionados: puede referirse a 1) *sonido de alerta* o a 2) *ninfa del mar*; en cambio, la palabra ambigua *globo* posee significados relacionados: puede referirse a 1) *bolsa de goma o de otro material flexible que se llena de gas o de aire* o a 2) *nave aerostática formada por una bolsa que, llena de un gas de menor densidad que el aire atmosférico, eleva una barquilla sujeta a su parte inferior*.

A continuación te presentamos un cuestionario de evaluación de palabras. En el cuestionario encontrarás tanto palabras ambiguas como palabras no ambiguas. Cada palabra a evaluar aparecerá acompañada por otras dos, a las cuales denominamos *modificadores*. Los modificadores sirven para dirigir tu atención hacia un significado concreto de la primera palabra. Tu tarea consiste en comparar el significado de la palabra a evaluar acompañada por el primer modificador con el significado de la misma acompañada por el segundo modificador. Expresarás tu respuesta en una escala de 1 a 9, donde 1 indica que los significados son completamente distintos, y 9 que se trata de exactamente el mismo significado. Intenta utilizar todos los valores de la escala.

Instructions for the ROM$_2$ questionnaire, translated from the original instructions in Spanish

Below you will find a list of ambiguous words, that is, words that have more than one meaning. We ask you to:

1. Read each word (e.g., “siren”).
2. Think of all the meanings of that word (e.g., “warning device” and “sea nymph”)
3. Rate the relatedness of these meanings.

We provide you a scale ranging from 1 (*unrelated meanings*) to 7 (*highly related meanings*) to make your ratings. You should use low relatedness values if the meanings are not related. For example, you can use a low relatedness value to rate the ambiguous word “siren,” because it has two unrelated meanings (i.e., “warning device” and “sea nymph”). Instead, you should use high relatedness values if the meanings are related. For example, you can use a high relatedness value to rate the ambiguous word “newspaper,” because it refers to multiple related meanings: (a) a publication, usually issued daily or weekly;

(b) a business organization that prints and distributes such a publication; (c) a single issue of such a publication; and (d) the paper on which a newspaper is printed.

Important: Although all the words of the questionnaire are ambiguous, it is possible that initially you only remember one meaning of a given word. If you find yourself in that situation, take as much time as you need to remember all its meanings. If you are still unable to remember any other meaning, you can leave that item unanswered.

Original instructions in Spanish for the ROM$_2$ questionnaire

A continuación verás una lista formada por palabras ambiguas, es decir, palabras que poseen más de un significado. La tarea que debes realizar consiste en:

1. Leer la palabra (p.ej., “sirena”).
2. Pensar en TODOS los significados de esa palabra (p.ej., “alarma” y “ninfa del mar”).
3. Indicar si los significados de esa palabra están relacionados entre ellos o no.

Para evaluar el grado de relación de los significados dispones de una escala con puntuaciones de 1 a 7. Utiliza puntuaciones bajas si los significados de la palabra no están relacionados. Por ejemplo, la palabra ambigua “sirena” deberías puntuarla con un valor bajo, ya que posee dos significados no relacionados (“sirena” y “ninfa del mar”). En cambio, utiliza puntuaciones altas si los significados de la palabra están relacionados. Por ejemplo, la palabra ambigua “diario” deberías puntuarla con un valor alto, puesto que tiene significados relacionados entre sí: (1) que se corresponde a cada día (p.ej., desayuno diario) y (2) periódico que se publica todos los días (p.ej., El País, El Mundo, etc.).

Importante: Aunque todas las palabras del cuestionario son ambiguas, es posible que en un primer instante sólo recuerdes un significado de alguna(s) de ellas. Si te encuentras en esa situación, tómate todo el tiempo que sea necesario para intentar recordar todos sus significados. Si aun así te es completamente imposible recordar ningún otro significado, puedes dejar la respuesta en blanco.

ACKNOWLEDGMENTS

This work was supported by the Ministerio de Economía y Competitividad of Spain and by the Fondo Europeo de Desarrollo Regional (PSI2015-63525-P MINECO/FEDER), as well as Ref. 2014PFR-URV-B2-37 from the Universidad Rovira i Virgili.

NOTE

1. Before administering the ROM$_1$ questionnaire, we assessed whether participants could easily understand the instructions of the task. To accomplish this, a group of 21 students was asked to complete a pilot version of the questionnaire. The questionnaire consisted of 15 ambiguous and 15 unambiguous words selected from the stimuli set. We calculated the reliability of participants’ scores by correlating each participant’s data with the mean of the whole sample. The results showed a very high correlation, $r\,(19) = .81, p < .001$. In light of these results, we used exactly the same instructions in the final rating study.

REFERENCES

Alonso, M. A., Fernández, A., & Díez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods, 47*, 268–274. doi:10.3758/s13428-014-0454-2

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In L. A. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2223–2228). Austin, TX: Cognitive Design Society.

Azuma, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments, and Computers, 28*, 109–124. doi:10.3758/BF03203645

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than fAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language, 36*, 484–504. http://doi.org/10.1006/jmla.1997.2502

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–221). Hillsdale, NJ: Erlbaum.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 49–59. doi:10.1016/0005-7916(94)90063-9

Britton, B. K. (1978). Lexical ambiguity of words used in English text. *Behavior Research Methods and Instrumentation, 10*, 1–7. doi:10.3758/BF03205079

Domínguez, A., Cuetos, F., & de Vega, M. (2001). 100 Palabras polisémicas con sus acepciones. *Revista Electrónica de Metodología Aplicada, 6*, 63–84.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, M. A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*, 1246–1258. doi:10.3758/s13428-013-0326-1

Durkin, K., & Manning, J. (1989). Polysemy and the subjective lexicon: Semantic relatednes and the salience of interword senses. *Journal of Psycholinguistic Research, 18*, 577–612.

Eddington, C., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review, 22*, 13–37. doi:10.3758/s13423-014-0665-7

Eich, E. (1984). Memory for unattended events: Remembering with and without awareness. *Memory & Cognition, 12*, 105–111. doi:10.3758/BF03198423

Estevez, A. (1991). Estudio normativo sobre ambigüedad en castellano. *Psicológica, 3*, 237–271.

Ferraro, F., & Kellas, G. (1990). Normative data for number of word meanings. *Behavior Research Methods, Instruments, and Computers, 22*, 491–498. doi:10.3758/BF03204432

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, and Computers, 26*, 5–25. doi:10.3758/BF03204557

Gee, N. R., & Harris, S. L. (2010). Homograph norms: An alternative approach to determining meaning dominance. *Behavior Research Methods, 42*, 976–986. doi:10.3758/BRM.42.4.976

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*, 256–281. doi:10.1037/0096-3445.113.2.256

Gómez Veiga, I., Carriedo López, N., Rucián Gallego, M., & Vila Cháves, J. O. (2010). Estudio normativo de ambigüedad léxica en castellano, en niños y en adultos. *Psicológica, 31*, 25–47.

Gorfein, D. S., Viviani, J. M., & Leddo, J. (1982). Norms as a tool for the study of homography. *Memory & Cognition, 10*, 503–509. doi:10.3758/BF03197654

Griffin, Z. M. (1999). Frequency of meaning use for ambiguous and unambiguous words. *Behavior Research Methods, Instruments, and Computers, 31*, 520–530.

Guasch, M., Ferré, P., & Fraga, I. (2015). Spanish norms for affective and lexico-semantic variables for 1,400 words. *Behavior Research Methods*. Advance online publication. doi:10.3758/s13428-015-0684-y

Hala, S., Pexman, P. M., & Glenwright, M. (2007). Priming the meaning of homographs in typically developing children and children with autism. *Journal of Autism and Developmental Disorders, 37*, 329–340. doi:10.1007/s10803-006-0162-6

Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy, 48*, 171–178.

Hertel, P. T., & El-Messidi, L. (2006). Am I Blue? Depressed mood and the consequences of self-focus for the interpretation and recall of ambiguous words. *Behavior Therapy*, *37*, 259–268. doi:10.1016/j.beth.2006.01.003

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: When does relatedness matter? *Canadian Journal of Experimental Psychology, 64*, 180–196. doi:10.1037/a0020475

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1331–1356. doi:10.1037/0096-1523.22.6.1331

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 686–713. doi:10.1037/0278-7393.28.4.686

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language, 55*, 247–273. doi:10.1016/j.jml.2006.04.001

Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language, 52*, 1–28. doi:10.1016/j.jml.2004.08.003

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*, 278–305. doi:10.1016/0010-0285(81)90011-6

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 601–609. doi:10.1037/0096-1523.14.4.601

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20*, 1–24. doi:10.1016/j.jneuroling.2006.02.001

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language, 123*, 11–21. doi:10.1016/j.bandl.2012.06.007

Lin, C.-J. C., & Ahrens, K. (2005). How many meanings does a word have? Meaning estimation in Chinese and English. In J. W. Minett & W. S.-Y. Wang (Eds.), *Language acquisition, change and emergence: Essays in evolutionary linguistics* (pp. 437–464). Hong Kong: City University of Hong Kong Press.

Lin, C.-J. C., & Ahrens, K. (2010). Ambiguity advantage revisited: Two meanings are better than one when accessing Chinese nouns. *Journal of Psycholinguistic Research, 39*, 1–19. doi:10.1007/s10936-009-9120-8

Mathews, A., Richards, A., & Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology, 98*, 31–34. doi:10.1037/0021-843X.98.1.31

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17*, 141–147. doi:10.3758/BF03197064

Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The University of South Florida homograph norms. *Behavior Research Methods and Instrumentation, 12*, 16–37. doi:10.3758/BF03208320

Nickerson, C. A., & Cartwright, D. S. (1984). The University of Colorado meaning norms. *Behavior Research Methods, Instruments, and Computers, 16*, 355–382. doi:10.3758/BF03202462

Panman, O. (1982). Homonymy and polysemy. *Lingua, 58*, 105–136. doi:10.1016/0024-3841(82)90059-6

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1252–1270. doi:10.1037/0278-7393.30.6.1252

Real Academia Española. (1984). *Diccionario de la lengua española* (20th ed.). Madrid: Author.

Real Academia Española. (2014). *Diccionario de la lengua española* (23rd ed.). Madrid: Author.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*, 245–266. doi:10.1006/jmla.2001.2810

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28*, 89–104. doi.10.1016/j.cogsci.2003.08.002

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*, 487–494. doi:10.1016/S0022-5371(70)80091-3

Simpson, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359–374). San Diego, CA: Academic Press.

Swinney, D. A. (1991). The resolution of interdeterminacy during language comprehension: Perspectives on modularity in lexical, structural and pragmatic process. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 367–385). North Holland: Elsevier Science.

Syssau, A., & Laxén, J. (2012). The influence of semantic richness on the visual recognition of emotional words. *Canadian Journal of Experimental Psychology, 66*, 70–78. doi:10.1037/a0027083

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes, 22*, 727–779. doi:10.1080/01690960601057068

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition, 22*, 111–126. doi:10.3758/BF03202766

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188–208. doi:10.1016/0749-596X(87)90123-9

Wollen, K. A., Cox, S. D., Coahran, M. M., Shea, D. S., & Kirby, R. F. (1980). Frequency of occurrence and concreteness ratings of homograph meanings. *Behavior Research Methods and Instrumentation, 12*, 8–15. doi:10.3758/BF03208319

2.2. Study 2: Haro, J., & Ferré, P. (2018). Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition? *Journal of Psycholinguistic Research*, *47*, 679-698.

2.2.1. Description of the study

There seems to be a relationship between the approach used for NOM estimation and the effect of NOM in LDT. While all studies that have used a dictionary approach have observed an ambiguity disadvantage (Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006), those that have employed subjective NOM have found an ambiguity advantage (Hino et al., 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004). Hence, it is important to determine whether these conflicting results are related to the approach used to estimate NOM. Namely, in past studies using a dictionary approach there might have been a misdistribution of subjectively unambiguous and ambiguous words into words with one/more than one dictionary entries. That is, some words with more than one dictionary entry might be unambiguous for speakers because the second (and subsequent) dictionary entries represent jargon, old fashioned or low-frequency meanings. In contrast, some words with one dictionary entry may be ambiguous for speakers because dictionaries do not include the new meanings that speakers have incorporated into their daily use of language.

To address this issue, in the present study three LDT experiments were conducted. In them, the approach for NOM estimation was manipulated. Namely, in the Experiment 1 we used a dictionary approach, that is, we assumed that each dictionary entry corresponded to a word meaning. On the contrary, in the Experiments 2 and 3 we used subjective NOM ratings. These ratings were obtained by asking participants to indicate if a string of letters has (0) no meaning, (1) one meaning, or (2) more than one meaning. As such, words with ratings closer to 1 were classified as unambiguous, and those with ratings closer to 2 as ambiguous.

2.2.2 Predictions

In line with previous studies that employed subjective NOM measures, we expected an ambiguity advantage when the categorization was made according to subjective NOM (Experiments 2 and 3). In contrast, there were no clear predictions regarding the ambiguity effect when words were categorized according to dictionary NOM (Experiment 1). However, assuming that subjectively unambiguous and ambiguous words do not entirely correspond to words with one/more than one dictionary entries, it is plausible to expect a null ambiguity advantage or even an ambiguity disadvantage when using dictionary NOM.

J Psycholinguist Res
https://doi.org/10.1007/s10936-017-9554-3

Semantic Ambiguity: Do Multiple Meanings Inhibit or Facilitate Word Recognition?

Juan Haro[1] · **Pilar Ferré**[1]

Abstract It is not clear whether multiple unrelated meanings inhibit or facilitate word recognition. Some studies have found a disadvantage for words having multiple meanings with respect to unambiguous words in lexical decision tasks (LDT), whereas several others have shown a facilitation for such words. In the present study, we argue that these inconsistent findings may be due to the approach employed to select ambiguous words across studies. To address this issue, we conducted three LDT experiments in which we varied the measure used to classify ambiguous and unambiguous words. The results suggest that multiple unrelated meanings facilitate word recognition. In addition, we observed that the approach employed to select ambiguous words may affect the pattern of experimental results. This evidence has relevant implications for theoretical accounts of ambiguous words processing and representation.

Keywords Semantic ambiguity · Ambiguity advantage · Ambiguity disadvantage · Multiple meanings · Word recognition

Introduction

Ambiguity is a characteristic feature of language: most of the words that we read or hear every day have more than one meaning (i.e., ambiguous words; e.g., *pupil*, *bat*, *newspaper*, etc.). As such, a complete theory of language comprehension must account for how ambiguous words are processed and represented. This is why semantic ambiguity has been extensively studied for the last 40 years by using different experimental paradigms and tasks, such as lexical decision task (LDT), reading task, semantic categorization task, or relatedness judgment task, among others (see Eddington and Tokowicz 2015, and Simpson 1984, for reviews). The results of these studies have been inconclusive concerning the role of semantic ambiguity

---

✉ Juan Haro
juan.haro@urv.cat

[1] Research Center for Behavior Assessment (CRAMC) and Department of Psychology, Universitat Rovira i Virgili, Crta. de Valls s/n, Campus Sescelades, 43007 Tarragona, Spain

in processing. The main aim of the present work was to shed further light on this issue by examining whether multiple meanings facilitate or inhibit word recognition. To do so, we focused on how ambiguous words are processed in LDT, which is the most used task to study word recognition (e.g., Balota et al. 2007).

Initial studies on semantic ambiguity showed that ambiguous words were recognized faster than unambiguous words in LDT (e.g., Borowsky and Masson 1996; Hino and Lupker 1996; Jastrzembski 1981; Jastrzembski and Stanners 1975; Kellas et al. 1988; Millis and Button 1989; but see Forster and Bednall 1976; Gernsbacher 1984). This so-called *ambiguity advantage* received different explanations. Originally it was argued that ambiguous words would have multiple lexical representations, each for one of their meanings. Consequently, the chance to quickly select one of the lexical representations of an ambiguous word in LDT would be higher than the chance to select the single representation of an unambiguous word (Jastrzembski 1981; Rubenstein et al. 1970). In contrast, subsequent accounts suggested that ambiguous words might not have multiple lexical representations, but rather multiple semantic representations. Under this assumption, the ambiguity advantage would arise from ambiguous words triggering a greater amount of semantic activation after their presentation than unambiguous words. This would facilitate their recognition either by an increase in the global activation at the semantic level (Borowsky and Masson 1996) or by a large semantic-to-orthographic feedback (Balota et al. 1991; Hino and Lupker 1996).

Whatever the explanation, the mentioned accounts agree in that ambiguous words benefit from having multiple meanings. However, several studies conducted during the last decade have strongly challenged that the source of the ambiguity advantage is the multiplicity of meanings (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Klepousniotou and Baum 2007; Rodd et al. 2002; Tamminen et al. 2006). These studies have distinguished two types of semantic ambiguity: (a) ambiguity between meanings, and (b) ambiguity between senses. On the one hand, ambiguity between meanings—also called *homonymy*—is the best known form of semantic ambiguity. It is observed in words referring to multiple unrelated meanings; for example, in words like *bat,* which can mean either *nocturnal flying mammal*, or *a club used in certain games to strike the ball*. On the other hand, ambiguity between senses is present in words referring to a wide range of related meanings; for instance, in the word *newspaper*, which means: (a) *a publication, usually issued daily or weekly*; (b) *a business organization that prints and distributes such a publication*; (c) *a single issue of such a publication*, and (d) *the paper in which a newspaper is printed*. This class of semantic ambiguity is called *polysemy,* and each of the related meanings of a polysemous word is named *sense*. Of note, in the following we will refer to unrelated meanings as *meanings* and to related meanings as *senses*.

The first study that examined these two types of semantic ambiguity was that of Rodd et al. (2002, Experiments 2 [visual] and 3 [auditory]), where number of meanings (i.e., one vs. many) and number of senses (few vs. many) were manipulated orthogonally. Contrary to previous evidence, the results showed that words with many meanings were recognized slower than words with one meaning (although only in the Experiment 3). In contrast, words with many senses were recognized faster than words with few senses. Thus, surprisingly, these authors found that multiple meanings inhibited word recognition (i.e., *ambiguity disadvantage*), whereas many senses facilitated it (i.e., *sense advantage*). Furthermore, Rodd et al. analysed previous studies reporting a clear ambiguity advantage (e.g., Azuma and Van Orden 1997; Millis and Button 1989), and found that their ambiguous words differed in number of senses, but not in number of meanings, with respect to their unambiguous words. According to Rodd et al., all these findings suggested that previous reports of a processing advantage for ambiguous words should be interpreted as a benefit for words with many senses rather than for words with many meanings. Thus, these authors concluded that the source of the ambiguity advantage is not the multiplicity of meanings, but the multiplicity of senses.

The results from Rodd et al. (2002) posed a challenge to previous accounts of ambiguous word recognition. In order to provide an explanation for them, Rodd et al. (2004) developed a PDP model of word recognition. In this model, each meaning or sense of a word is represented by an attractor basin located in a semantic network. During the processing of a word, the semantic network moves from an initial state towards the word's attractor basin. The word is recognized (e.g., in LDT) when the semantic network enters the word's attractor basin. Senses are represented by neighbouring attractor basins, forming a single, broad and shallow attractor basin; in contrast, meanings are represented by attractor basins located at different and distant regions of the semantic network. Given these assumptions, the model predicts a sense advantage: the semantic network should find the broad and shallow attractor basin of a word with multiple senses faster than the narrow attractor basin of a word with few senses. On the other hand, a distinct prediction is made for words with multiple meanings. In this case, a blend of all their meanings would be activated (i.e., a *blend state*) in the early stages of the semantic network's settling. After that, and in order to complete word processing, the semantic network should escape from this blend state and move towards one of the word's attractor basins. Assuming that moving away from the blend state involves a high processing cost, and that meanings compete during word processing, the model predicts a disadvantage in the recognition of words with multiple meanings in comparison to words with one meaning.

Rodd et al. (2004)'s model has gathered additional support from several LDT studies that showed a disadvantage for words with multiple meanings along with an advantage for words with many senses (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Tamminen et al. 2006). However, there is also evidence incompatible with the model, as other LDT studies have found an advantage of the same magnitude for both types of ambiguous words (e.g., Hino et al. 2010, 2006; Pexman et al. 2004). Importantly, these latter results are consistent with previous reports of an ambiguity advantage in suggesting that multiple meanings facilitate word processing (e.g., Borowsky and Masson 1996; Hino and Lupker 1996; Jastrzembski 1981; Jastrzembski and Stanners 1975; Kellas et al. 1988; Millis and Button 1989).

As can be seen, the above conflicting findings do not allow researchers to draw firm conclusions about the representation and processing of ambiguous words. The main question is whether multiple unrelated meanings (e.g., *bat*) inhibit or facilitate word processing. The way in which this question is answered has a significant theoretical relevance, as each possible answer would give support to a different account of the semantic ambiguity phenomenon. Indeed, the ambiguity disadvantage would suggest, according to the Rodd et al. (2004)'s account, that the different semantic representations of a word with multiple meanings compete during word recognition. In contrast, the ambiguity advantage would be in line with accounts of an enhanced semantic activation for ambiguous words, according to which words with multiple meanings benefit from triggering a greater amount of semantic activation than unambiguous words (e.g., Borowsky and Masson 1996; Hino and Lupker 1996). In what follows we suggest that the above opposite experimental findings could be explained, at least in part, by the different approaches employed to categorize ambiguous and unambiguous words across studies.

Importantly, all the LDT studies that obtained an ambiguity disadvantage employed a *dictionary approach* (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Rodd et al. 2002; Tamminen et al. 2006). This approach relies on the assumption that unrelated meanings are listed in separate dictionary entries, so that words with more than one dictionary entry are classified as ambiguous, whereas words with only one dictionary entry are classified as unambiguous. Moreover, four of these dictionary-based studies used exactly the same set of experimental stimuli (Armstrong and Plaut 2008; Beretta et al. 2005; Rodd et al. 2002; Tamminen et al. 2006), which were in turn those employed by Rodd et al. (2002, Experiments 2 and 3). Finally, it should be noted that the ambiguity disadvantage has been found only in studies conducted in English.

On the contrary, the studies that found an advantage for both types of ambiguous words (i.e., polysemes and homonyms) relied on a *subjective approach* to categorize their stimuli (e.g., Haro et al. 2017a; Hino et al. 2010, 2006; Lin and Ahrens 2010; Pexman et al. 2004), and they were conducted in different languages, such as English (e.g., Experiment 1 in Pexman et al. 2004), Japanese (e.g., Hino et al. 2006), Chinese (e.g., Lin and Ahrens 2010), and Spanish (Haro et al. 2017a). The subjective measures employed in these studies were mainly two: Number-Of-Meanings (hereafter, NOM) and Relatedness-Of-Meanings (hereafter, ROM). To obtain NOM ratings, participants are asked to indicate if a word has one or more than one meaning, so that this measure allows to determine if a word is ambiguous or not according to speakers' knowledge. On the other hand, to collect ROM ratings, participants are required to indicate whether the meanings of an ambiguous word are related or unrelated. As such, ROM can be employed to categorize ambiguous words into polysemes and homonyms on the basis of the linguistic knowledge of participants.

It is worth noting that depending on the approach employed, the same word might be classified either as unambiguous or as ambiguous. This becomes clear if we compare how words' meanings are represented in dictionaries to how they are likely to be represented in the speaker's memory. For instance, Gernsbacher (1984) showed that college professors could provide only 2 of the 30 dictionary definitions of the word *gauge*, 3 of the 15 dictionary definitions of the word *fudge*, and 1 of the 15 dictionary definitions of the word *cadet*. This led Gernsbacher to suggest that speakers do not store in their memory all the word definitions listed in the dictionary, but only a small sample of them. In addition, Ferraro and Kellas (1990) showed that the correlation between subjects' ratings of number of meanings of English words and dictionary entries only accounted for the 12% of the total variance. In a similar study, Lin and Ahrens (2005) compared subjective norms of word meanings and dictionary entries in Chinese and English. They observed that, although correlated, the number of word meanings provided by each measure was significantly different. In addition, they pointed out that the meanings obtained through each of these measures may also differ with regard to their content, since dictionaries are plenty of old-fashioned definitions and do not reflect the emerging novel meanings of words.

The above considerations indicate that a dictionary approach may have some limitations as a psychological measure of semantic ambiguity. Indeed, we believe that there are, at least, two potential problems with this approach. The first of them occurs when a word is classified as ambiguous according to the dictionary (i.e., because it has more than one dictionary entry), but the second and subsequent entries/meanings are unknown or unfamiliar for the majority of the speakers. This usually concerns words whose subordinate dictionary meanings are old-fashioned, jargon or have a very low frequency. For example, considering the Spanish Language Dictionary published by the "Real Academia Española" (RAE) (2014), the word *coleta* is ambiguous because it has two distinct entries. The first entry is related to *a tied-back hairstyle* (i.e., "ponytail" in English), whereas the second one refers to *a mix of glue and honey used to restore paintings*. Although many speakers know the first meaning of *coleta*, the second meaning is unknown or unfamiliar to almost all of them, as it is a very low frequency jargon meaning.

The second problem arises when the word is ambiguous for nearly all the speakers, but not according to the dictionary (i.e., it has only one dictionary entry). An example is the Spanish word *tronco*, which has two unrelated meanings for the majority of the speakers: *trunk* and *mate*. However, the RAE dictionary includes only one entry for the word *tronco*, which, in addition, does not cover the *mate* meaning, the most recent but widely known meaning of such word. Moreover, many dictionaries group together different meanings under the same entry according to etymological criteria. This leads to cases where unrelated meanings are put together within the same entry. For example, the RAE dictionary provides only one entry for the word *sirena* ("siren" in English), comprising their two unrelated meanings: *sea nymph* and *alarm device*. More importantly for the present study, several instances of this problem can be found in the stimuli of Rodd et al. (2002). For example, the word *hang* was considered as unambiguous in such study, although it has two unrelated meanings (i.e., *to hold* and *the idea of how to do something, i.e., knack*). The same was true for the words *belt* (i.e., *a strip of cloth* and *to sing loudly*) and *soap* (i.e., *a cleaning substance* and *soap opera*). Probably, the reason for this was that the two meanings of these words are listed under the same entry in the Wordsmyth dictionary (Parks et al. 1998), the one used by Rodd et al.

Given the differences between both approaches, it is very relevant to determine whether the criterion used to classify words as ambiguous or unambiguous has contributed to the inconsistent findings between dictionary-based and subjective-based studies. Thus, to contribute to a better understanding of how ambiguous words are processed, we conducted three LDT experiments in which we varied the measure used to categorize ambiguous and unambiguous words. This would particularly help to identify the most suitable- and psychologically valid-approach to classify words having one and multiple meanings, and thereby to properly assess whether multiple meanings facilitate or inhibit word recognition.

Experiment 1

The aim of this experiment was to examine how ambiguous words are processed in LDT, by using a dictionary approach. To do so, we manipulated orthogonally dictionary entries and dictionary senses. Thus, words with many dictionary entries were defined as ambiguous, and words with one dictionary entry were defined as unambiguous. In addition, words were classified as having few or many senses on the basis of their number of dictionary senses. In line with previous studies that used a similar dictionary approach, we predicted a lack of advantage for words with many dictionary meanings. Indeed, they would be recognized either as fast and accurately (e.g., Rodd et al. 2002 [Experiment 2]), or even slower and less accurately (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Tamminen et al. 2006) than words with one dictionary meaning. In contrast, we expected words with many senses to be recognized faster and more accurately than words with few senses (Rodd et al. 2002). Finally, we did several additional analyses by considering a subjective measure of ambiguity (NOM) in order to know whether the pattern of results varied depending on the criterion used to classify the stimuli.

Method

Participants

Thirty-nine Spanish speakers (34 women and 5 men, mean age 20.4 years, SD = 2.1) participated in the experiment. They were undergraduate students who received academic credits for their participation. All of them had either normal or corrected-to-normal vision.

Design and Materials

Following previous studies (Beretta et al. 2005; Rodd et al. 2002), we used a 2 × 2 factorial design, with the first factor ambiguity (unambiguous words vs ambiguous words) and the second, number of senses (words with few senses vs words with many senses). We selected 72 Spanish words from the RAE dictionary to be included in the factorial design. Regarding the ambiguity factor, words were categorized as unambiguous if they had only one entry in the RAE dictionary and as ambiguous if they had more than one entry. Overall, the average number of dictionary entries was 1 ($SD = 0$) and 2.14 ($SD = 0.49$) for unambiguous and ambiguous words, respectively, $t(70) = 14.03,\ p < .001$. Number of senses for each word was computed as the total number of definitions listed considering all the dictionary entries of that word. Words having 10 or less dictionary senses were classified as words with few senses, whereas those having more than 10 senses were classified as words with many senses. Words with few senses had 6.33 senses ($SD = 2.46$) on average, and words with many senses had 17.44 senses ($SD = 6.25$) on average, $t(70) = 9.93,\ p < .001$.

Each cell of the factorial design consisted of 18 words. Number of dictionary senses was matched between ambiguous and unambiguous words, $t(70) = 0.48,\ p = .63$, and number of dictionary entries was matched between words with few senses and words with many senses, $t(70) = 0.28,\ p = .86$. In addition, many lexical and semantic variables were matched as close as possible across experimental conditions (all ps $> .2$, see Table 1 for more details). These variables were obtained from distinct sources. On the one hand, word length, number of syllables, logarithm word frequency (log word frequency), mean Levenshtein distance of the 20 closest words (old20), number of neighbours, number of higher frequency neighbours, bigram frequency, trigram frequency, and logarithm contextual diversity (log contextual diversity) were obtained from the EsPal subtitle tokens' database (Duchon et al. 2013). On the other hand, familiarity, concreteness, subjective age of acquisition, arousal and emotional valence ratings were taken from the database of Haro et al. (2017b).

Finally, following the procedure of Rodd et al. (2002), 72 pseudohomophones (i.e., nonwords that have a similar pronunciation to words, such as "sircle", pronounced like "circle") matched to words in length were used as nonwords in the LDT. The reason for using pseudohomophones is that they make the distinction between non-words and real words more difficult. This may cause a greater activation of the meanings and senses of the words resulting in larger ambiguity effects (Eddington and Tokowicz 2015). We also decided to use pseudohomophones for the sake of comparability with the study of Rodd et al. (2002), as they employed this type of non-words.

Procedure

Participants completed a lexical decision task consisting of 144 experimental trials. Each trial started with a fixation point (i.e., "+") appearing in the middle of the screen for 500 ms. Next, a string of letters (a word or a nonword) replaced the fixation point, and then participants had to decide whether the string was a Spanish word or not. They were instructed to press the "yes" button of a keypad with the preferred hand if the string of letters was a word, and to press the "no" button of the keypad with the non-preferred hand if it was not a word. The string of letters remained on the screen until participant's response or timeout (i.e., 2000 ms). After responding, a feedback message (i.e., "ERROR" or "CORRECT") was displayed for 750 ms. The interval time between trials was 500 ms. The DMDX software (Forster and Forster 2003) was used for displaying the stimuli and recording the responses. The order of

Table 1 Characteristics of the stimulus used in the experiments (standard deviations are shown in parentheses)

	DIC	SEN	NOM	ROM1	ROM2	FRE	CTD	FAM	AoA	LNG	SYL	CON	VAL	ARO	OLD	NEI	NHF	BFQ	TFQ
Experiment 1																			
Unambiguous, few senses	1	6.28	1.6	–	–	1.35	0.93	5.51	6.58	5.67	2.39	4.66	4.98	4.32	1.62	7.06	1.11	5224	675
	(0)	(2.63)	(0.29)			(0.53)	(0.44)	(0.73)	(2.2)	(1.41)	(0.61)	(0.91)	(0.95)	(0.81)	(0.41)	(7.65)	(1.94)	(2885)	(633)
Unambiguous, many senses	1	18.33	1.71	–	–	1.25	0.88	5.68	6.25	5.5	2.33	4.6	4.96	4.27	1.55	8.72	1.61	5696	611
	(0)	(6.94)	(0.21)			(0.52)	(0.4)	(0.74)	(1.99)	(1.29)	(0.69)	(0.92)	(1.05)	(0.78)	(0.41)	(9.32)	(2.3)	(4185)	(793)
Ambiguous, few senses	2.17	6.39	1.54	–	–	1.11	0.74	5.34	6.74	5.56	2.5	4.8	5.23	4.43	1.62	7.78	1.56	4855	674
	(0.51)	(2.35)	(0.27)			(0.67)	(0.5)	(0.96)	(1.72)	(1.54)	(0.62)	(1.17)	(1.52)	(1.34)	(0.48)	(8.36)	(3.29)	(3843)	(731)
Ambiguous, many senses	2.11	16.56	1.79	–	–	1.35	0.96	5.45	5.97	5.67	2.56	4.51	5.31	4.53	1.44	8.89	0.89	6203	986
	(0.47)	(5.52)	(0.15)			(0.67)	(0.53)	(1.09)	(1.95)	(1.33)	(0.51)	(0.82)	(1.2)	(1.01)	(0.32)	(8.26)	(1.49)	(4295)	(1298)
Experiment 2																			
Unambiguous	–	4.63	1.13	–	–	0.99	0.63	5.36	6.43	6.69	2.78	5.03	5.12	4.53	1.97	4.38	0.94	5218	489
		(1.76)	(0.1)			(0.59)	(0.44)	(1.04)	(2.16)	(2.21)	(0.87)	(0.93)	(1.3)	(0.96)	(0.57)	(6.61)	(1.88)	(3420)	(471)
Ambiguous	–	4.91	1.78	–	–	1.04	0.7	5.32	6.73	6.41	2.69	4.84	5.38	4.63	1.83	4.13	0.63	5507	705
		(1.59)	(0.08)			(0.5)	(0.38)	(0.78)	(1.81)	(1.32)	(0.59)	(0.65)	(1.01)	(1.12)	(0.42)	(5)	(1.34)	(3543)	(707)
Experiment 3																			
Unambiguous	–	5.63	1.13	–	–	1.10	0.71	5.48	6.39	6.37	2.78	5.07	5.17	4.51	1.93	5.07	1.07	4300	433
		(2.75)	(0.10)			(0.66)	(0.49)	(1.07)	(2.34)	(2.51)	(0.97)	(1.02)	(1.29)	(0.95)	(0.69)	(6.69)	(2.04)	(3284)	(430)
Homonyms	–	6.19	1.76	1.76	2.06	0.97	0.63	5.31	6.94	6.41	2.78	4.76	4.87	4.55	1.78	6.48	1.37	5708	608
		(2.34)	(0.12)	(0.53)	(0.50)	(0.52)	(0.40)	(0.73)	(2.26)	(1.74)	(0.64)	(0.80)	(0.95)	(0.85)	(0.52)	(8.65)	(2.39)	(3696)	(522)

DIC number of dictionary entries, *SEN* number of dictionary senses, *NOM* subjective Number-Of-Meanings ratings, *FRE* log word frequency, *CTD* log contextual diversity, *FAM* familiarity, *AoA* subjective age-of-acquisition, *LNG* word length, *SYL* number of syllables, *CON* concreteness, *VAL* emotional valence, *ARO* emotional arousal, *OLD* old20, *NEI* number of substitution neighbors, *NHF* number of higher frequency substitution neighbors, *BFQ* mean bigram frequency, *TFQ* mean trigram frequency the experimental trials was randomized for each participant. Prior to the beginning of the experiment, a practice block consisting of 10 trials (5 words and 5 nonwords) was presented.

Data Analyses

Data on reaction times (RT) and error rates (%E) were analyzed separately using linear mixed-effect models (e.g., Baayen 2008; Baayen et al. 2008). For this we used the lme4 package of R (Bates et al. 2014). In each analysis, the word's ambiguity (unambiguous and ambiguous words), the number of senses of the word (few senses or many senses), and the ambiguity × number of senses interaction were included as fixed effects, and participants and words as random effects (adjusting for the intercept). To determine the significance of fixed effects, log-likelihood ratio tests were used (R function Anova). We assessed the contribution of each fixed effect by comparing a model that included the effect of interest with another model that did not include it. We also checked whether the fit of the model was significantly improved by adding any of the following additional variables: log word frequency, word length, familiarity, number of neighbours, and concreteness. Thus, we first created a model including the variables of interest and their interaction (i.e., ambiguity and number of senses), random effects (participants and words), and the additional variables. The model was then progressively simplified by excluding one additional variable in each step, this being the one making the lowest contribution at that point. As such, the remaining model only included those additional variables that significantly increased its fit. Of note, the *p* values for pairwise comparisons were estimated using the lmerTest package (Kuznetsova et al. 2014).

Results and Discussion

We removed the data of one participant with more than 15% of errors. Reaction times that exceeded 2 *SD* of each participant's mean were also removed (5.0% of the whole). In addition, we excluded from the analyses one unambiguous word with many senses due to a high percentage of errors ($> 70\%$). The mean of reaction times (RT) for correct responses and the mean of error rates (%E) across experimental conditions (averaged across participants) are shown in Table 2. The best fitting model for RTs included the additional variables log word frequency and familiarity, which significantly increased model's fit (all ps $< .05$). On the other hand, the best fitting model for %E included the additional variable log word frequency ($p < .05$).

Concerning ambiguity, there was a significant effect on RTs, $\chi^2(1) = 6.01,\ p = .014$. Ambiguous words were recognized slower than unambiguous words, $ß = 23.36,\ SE = 9.92,\ t = 2.35,\ p = .022$. Likewise, ambiguous words were recognized less accurately than unambiguous words, $\chi^2(1) = 8.34,\ p = .004$, $ß = .07,\ SE = .02,\ t = 2.87,\ p = .005$. With respect to number of senses, no differences were found between words with many

Table 2 Mean RT (in ms), and percentage of error rates (%E) in Experiment 1 (averaged across participants; standard deviations in parentheses)

Ambiguity	Senses	Mean RT	%E
Unambiguous	Few	595 (77)	1.81 (3.03)
Unambiguous	Many	591 (81)	3.92 (4.48)
Ambiguous	Few	634 (87)	12.24 (9.01)
Ambiguous	Many	601 (73)	4.41 (4.48)
Nonwords		673 (92)	9.4 (6.43)

senses and words with few senses, either on RTs, $\chi^2(1) = 2.56$, $p = .11$, or on %E, $\chi^2(1) = 0.99$, $p = .32$.

Regarding the interaction between ambiguity and number of senses, it was significant in %E, $\chi^2(1) = 4.33$, $p = .04$. Post-hoc comparisons showed that this interaction was produced because number of senses reduced the %E of ambiguous words, ß = −.06, *SE* = .04, $t = -1.49$, $p = .15$, but increased that of unambiguous words, ß = .03, *SE* = .01, $t = 1.93$, $p = .06$ (although the comparisons were not significant).

The results of this experiment showed a disadvantage for ambiguous words: words with more than one dictionary entry were recognized more slowly and less accurately than words with one entry. It should be noted, however, that this disadvantage seemed to be due mainly to ambiguous words with few senses, in that this was the condition with the slowest RTs and the highest %E, where a considerable difference was seen compared to the other conditions (although this difference failed to reach statistical significance). On the one hand, these results are incompatible with reports of an ambiguity advantage from LDT studies that relied on subjective norms (e.g., Hino et al. 2010, 2006; Pexman et al. 2004). Conversely, they are in line with several English LDT studies that employed a dictionary approach for selecting ambiguous and unambiguous words (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Rodd et al. 2002; Tamminen et al. 2006) and that mainly relied on the same stimulus set. The data from the present experiment replicate these latter findings by using the same approach for selecting ambiguous words, and, importantly, they extend them to another language and to another set of words. Hence, these results support that multiple unrelated meanings inhibit word recognition.

However, before drawing definite conclusions, we wanted to explore if the results of this experiment were affected by the type of criterion used to classify the ambiguous/unambiguous words. To that end, we obtained NOM ratings for the experimental words from Haro et al. (2017b). As described earlier, NOM ratings are a subjective measure of semantic ambiguity. To collect them, participants are usually required to think about all the meanings of a string of letters and then choose the appropriate value in a three-point scale: *the word has no meaning* (0), *the word has one meaning* (1), or *the word has more than one meaning* (2). Words with values close to 2 are considered as ambiguous, and words with values close to 1 are considered as unambiguous (e.g., Hino et al. 2010, 2006; Pexman et al. 2004).

The analysis of our experimental stimuli showed, on the one hand, that that there was no difference in NOM ratings for words with many dictionary entries (NOM = 1.66) and words with one dictionary entry (NOM = 1.65), $t(69) = 0.24$, $p = .81$ (see Table 1 for the NOM ratings of each condition). This may suggest that words with one/many dictionary entries would not correspond with words having one/many meanings according to the speakers' linguistic knowledge. On the other hand, we observed that NOM ratings and number of senses were significantly correlated, $r(71) = .29$, $p = .01$. Words with many senses had higher NOM values (NOM = 1.75) than words with few senses (NOM = 1.57), $t(69) = 3.22$, $p = .002$, from which it may be inferred that participants considered words with many senses as having more meanings than words with few senses.

After having characterized our stimuli in terms of a subjective measure, we analysed the relationship between that measure and participants' performance in the LDT. We found that NOM ratings were negatively correlated with both RTs, $r(71) = -.34$, $p = .004$, and %E, $r(71) = -.39$, $p = .001$. That is, words with higher NOM ratings were responded faster and more accurately than words with lower NOM ratings. In addition, we reclassified the words of Experiment 1 based on a median split of their NOM values, finding that words with high NOM ratings were associated with faster RTs (597 ms) and less errors (3.61%) than words with low NOM ratings (mean RTs = 624 ms; mean %E = 8%) (both *p*s < .05).

Importantly, it should be noted that number of dictionary senses did not differ between words with low NOM (mean senses = 10.66) and words with high NOM (mean senses = 13.06), $t(69) = 1.38$, $p = .17$. Thus, although NOM ratings and number of dictionary senses of the words were correlated, the NOM advantage may not be entirely explained by dictionary senses.

In sum, these results indicate that the direction of the ambiguity effect varied depending on the measure employed to classify the same stimuli. That is, we obtained an ambiguity disadvantage when the stimuli were classified according to the number of dictionary entries, whereas an ambiguity advantage emerged when the classification was based on NOM ratings. Apart from the methodological and theoretical implications of these findings, they suggest that the subjective number of meanings could have contributed to the ambiguity disadvantage found with the dictionary criterion. In particular, this latter result might be explained by a misdistribution of NOM between words with one/many dictionary entries. Indeed, although words with one/many dictionary entries did not differ in NOM ratings, it should be noted that words with many dictionary entries and few senses had the lowest NOM ratings (mean = 1.54, see Table 1). In addition, this was the condition showing the slowest RTs in the experiment. Interestingly, on examining the words of this condition we observed that several of them were unambiguous words according to NOM ratings. In addition, many words of the one dictionary entry and few senses ought to be considered ambiguous words, based on NOM ratings. Hence, this misdistribution of NOM ratings may have had a notable impact in these two conditions, increasing the RTs in the many dictionary entries and few senses condition, and reducing the RTs in the one dictionary entry and few senses condition. This would also serve to explain the reversed ambiguity effect when using dictionary entries as a measure of semantic ambiguity.

Taking the above into consideration, it seems that NOM have a relevant role in determining ambiguity effects. However, this conclusion is based on a post-hoc reclassification of the stimuli. In order to investigate in more detail this issue, we conducted a further experiment in which we manipulated NOM ratings. This led us to assess if words having multiple meanings according to the speakers' knowledge inhibit or facilitate word recognition.

Experiment 2

In this experiment, we employed NOM ratings as a subjective measure of semantic ambiguity. We compared words with high NOM ratings (i.e., ambiguous words) to words with low NOM ratings (i.e., unambiguous words). In line with several LDT studies that employed this measure (Hino et al. 2010, 2006; Pexman et al. 2004), and given the negative correlation between NOM and RTs found in Experiment 1, we expected a facilitation for ambiguous words with respect to unambiguous words. However, it is worth noting that neither of these NOM-based studies took into account the number of senses of their experimental stimuli. This is critical, since Rodd et al. (2002) demonstrated that past reports of an ambiguity advantage from studies relying on subjective norms might be due to ambiguous words having many senses rather than many meanings. In addition, it has been shown that words with many senses are recognized faster than words with few senses, (e.g., Jager and Cleland 2016). For that reason, in Experiment 2 we equated number of senses between conditions, as a way to avoid the possible experimental confounding between meanings and senses. Of note, we employed number of dictionary senses as a measure of the number of senses known by the speakers, given the high correlation between both measures (Fraga et al. 2017).

Method

Participants

Thirty-eight Spanish speakers (34 women and 4 men, mean age 21.8 years, SD = 4.3) participated in the experiment. They were undergraduate students who received academic credits for their participation. All of them had either normal or corrected-to-normal vision.

Design and Materials

Experimental stimuli consisted of 32 ambiguous words and 32 unambiguous words. Unlike Experiment 1, we relied on subjective measures to define ambiguity. We used two indices of semantic ambiguity taken from the Haro et al. (2017b)'s database. The first of them was number-of-meanings (NOM), which was described in the additional analyses included in Experiment 1. To obtain the second measure of ambiguity, Haro et al. used a free association task, by requiring participants to generate word associates for a set of words. Then, on the basis of these associates, four judges classified the words as unambiguous or ambiguous. Words were classified as unambiguous if all their associates were related to the same meaning (e.g., *major* was rated as unambiguous because participants produced the following associates for it: *council*, *town*, *politician*, *and president*), and as ambiguous if their associates were related to distinct meanings or senses (e.g., *mouse* was rated as ambiguous because participants produced the following associates for it: *computer*, *cheese*, *mickey*, *cat*, *hamster*, *and keyboard*).

Words with NOM values lower than 1.40 were classified as unambiguous, and words with NOM values higher or equal than 1.65 were classified as ambiguous. The average NOM rating was 1.13 ($SD = 0.10$) for unambiguous words and 1.78 ($SD = 0.08$) for ambiguous words, $t(62) = 27.74,\ p < .001$. In addition, when considering the associates' measure, all the unambiguous words had an "unambiguous" rating and all the ambiguous words had an "ambiguous" rating. As stated before, number of dictionary senses (measured as in Experiment 1) was matched between conditions, $t(62) = 0.67,\ p = .51$. In addition, we matched the same lexical and semantic variables between ambiguous and unambiguous words as in Experiment 1 (all $ps > .15$). Full details of the experimental words are shown in Table 1. Finally, a set of 64 pseudohomophones matched in length to words was included as nonwords in the LDT.

Procedure

The experimental procedure was the same as in Experiment 1.

Data Analyses

The analyses were the same as in Experiment 1, with the exception that the variable of interest in this experiment was word's ambiguity (unambiguous and ambiguous words).

Results and Discussion

We removed the data of two participants with more than 15% of errors. RTs that exceeded 2 *SD* of each participant's mean were also rejected (4.4% of the whole). In addition, we excluded from the analysis one unambiguous word due to a high percentage of errors (> 70%).

Table 3 Mean RT (in ms), and percentage of error rates (%E) in Experiment 2 (averaged across participants; standard deviations in parentheses)

Ambiguity	Mean RT	%E
Unambiguous	675 (107)	6.21 (4.71)
Ambiguous	627 (102)	1.71 (2.71)
Nonwords	724 (123)	6.81 (3.79)

The mean of RTs for correct responses and the mean of %E across experimental conditions (averaged across participants) are shown in Table 3. The best fitting model for RTs included the additional variable log word frequency, which significantly increased model's fit ($p < .05$). On the other hand, the best fitting model for %E included the additional variable concreteness ($p < .05$).

There was a significant facilitative effect for ambiguous words both in RTs, $\chi^2(1) = 13.31$, $p < .001$, $ß = -50.96$, $SE = 13.48$, $t = 3.78$, $p < .001$, and in %E, $\chi^2(1) = 12.15$, $p < .001$, $ß = -.06$, $SE = .02$, $t = 3.58$, $p < .001$.

The results showed a clear ambiguity advantage: ambiguous words were recognized faster and more accurately than unambiguous words, even when number of dictionary senses was matched between conditions. These results are at odds with the ambiguity disadvantage reported when a dictionary criterion is used (e.g., Armstrong and Plaut 2008; Beretta et al. 2005; Rodd et al. 2002; Tamminen et al. 2006, and Experiment 1 of this study). In contrast, this finding is in line with several LDT studies that have relied on subjective norms (e.g., Hino et al. 2010, 2006; Lin and Ahrens 2010; Pexman et al. 2004). It also agrees with the results from the additional analyses of Experiment 1, which suggested that NOM facilitates word processing, and that the observed disadvantage for words with many dictionary entries might be due to a misdistribution of NOM across conditions.

In light of the above, it could be concluded that the advantage for words with high NOM ratings suggests that multiple meanings facilitate word recognition. This conclusion would be based on the assumption that these words had multiple meanings according to participants. However, there are some reasons to question this assumption. First, Rodd et al. (2002) demonstrated that when participants are asked to provide word meanings, they also take into account related word senses. Second, NOM ratings and dictionary senses are highly correlated (e.g., $r = .50$ in the study of Haro et al. 2017b). Finally, a close look at the stimuli of Experiment 2 reveals that there are several examples of ambiguous words with related senses (e.g., *mouse* or *ecstasy*). Consequently, it is possible that unrelated meanings and related senses were somewhat confounded in Experiment 2. As such, the advantage for words of high NOM could be due to either multiple unrelated meanings or multiple related senses. To overcome this limitation and to elucidate whether multiple unrelated meanings facilitate word recognition, we conducted a new experiment in which we used a subjective measure of relatedness between meanings.

Experiment 3

In this experiment, we compared words having multiple unrelated meanings (homonyms) to unambiguous words. Homonyms were selected according to ROM ratings, a measure obtained by asking participants to indicate if the distinct meanings of an ambiguous word are related or unrelated. Thus, if multiple unrelated meanings facilitate word processing, homonyms should show an advantage over unambiguous words.

Method

Participants

Forty-four Spanish speakers (37 women; mean age 22.09 years, SD = 3.54) participated in the experiment. They were undergraduate students who received academic credits for their participation. All of them had either normal or corrected-to-normal vision.

Design and Materials

We selected 27 unambiguous words and 27 homonyms for the present experiment. Homonyms were chosen according to NOM and ROM ratings (ROM_1 and ROM_2) obtained from Haro et al. (2017b)'s database. The criteria for the selection of homonyms were that they should have NOM ratings above 1.5, as well as ROM_1 and ROM_2 ratings below 3. On the other hand, the criterion to select unambiguous words was that they should have NOM ratings below 1.4. Unambiguous words and homonyms differed in NOM ratings (1.13 and 1.76, respectively), $t(52) = 20.93$, $p < .001$. However, they did not differ in number of dictionary senses, $t(52) = 1.35$, $p = .18$.

We matched the same lexical and semantic variables as in Experiments 1 and 2 between unambiguous words and homonyms (all $ps > .15$). Full details of the experimental words are shown in Table 1. Finally, a set of 54 pseudohomophones matched in length to words was included as nonwords in the LDT.

Procedure

The experimental procedure was the same as in Experiments 1 and 2.

Data Analyses

The analyses were the same as in Experiments 1 and 2, with the exception that the variable of interest in this experiment was homonymy (homonyms and unambiguous words).

Results and Discussion

We removed the data of three participants with more than 15% of errors. RTs that exceeded 2 *SD* of each participant's mean were also rejected (4.6% of the whole). The mean of RTs for correct responses and the mean of %E across experimental conditions (averaged across participants) are shown in Table 4. The best fitting model for RTs included the additional variables log word frequency and familiarity, which significantly increased model's fit ($p < .05$), whereas the best fitting model for %E included the additional variable familiarity ($p < .05$).

Table 4 Mean RT (in ms), and percentage of error rates (%E) in Experiment 3 (averaged across participants; standard deviations in parentheses)

Ambiguity	Mean RT	%E
Unambiguous	646 (133)	5.86 (4.17)
Homonyms	617 (116)	4.42 (4.66)
Nonwords	706 (143)	8.22 (5.29)

Homonyms were recognized faster than unambiguous words, $\chi^2(1) = 9.61,\ p = .002$, $ß = -45.70,\ SE = 14.53, t = 3.15,\ p = .003$. On the other hand, the difference between homonyms and unambiguous words in %E did not reach significance, $\chi^2(1) = 0.73,\ p = .39$.

The results of the present experiment showed a facilitation in RTs for homonyms with respect to unambiguous words. This homonymy advantage agrees with past findings (e.g., Hino et al. 2010, 2006; Pexman et al. 2004), and supports the hypothesis that multiple unrelated meanings facilitate word recognition.

General Discussion

In the present study we aimed to assess whether multiple meanings facilitate or inhibit word recognition. To address this issue, we compared ambiguous words to unambiguous words in three LDT experiments. The results obtained suggest that multiple unrelated meanings facilitate word recognition. We also found that the results obtained in the LDT experiments depended on the approach used to select these words. Namely, when number of dictionary entries was used (Experiment 1), a disadvantage for ambiguous words (i.e., words with more than one dictionary entry) appeared. Such a result is in line with previous LDT studies that employed a similar approach to categorize their stimuli (Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Rodd et al. 2002; Tamminen et al. 2006). In contrast, when the classification was made according to subjective ratings of semantic ambiguity (i.e., NOM and ROM ratings, Experiments 2 and 3), a facilitation for ambiguous words was obtained, in agreement with early reports of an ambiguity advantage (e.g., Borowsky and Masson 1996; Hino and Lupker 1996; Jastrzembski 1981; Jastrzembski and Stanners 1975; Kellas et al. 1988; Millis and Button 1989), and with the findings of more recent studies that have relied on subjective norms (e.g., Haro et al. 2017a; Hino et al. 2010, 2002, 2006; Lin and Ahrens 2010; Pexman et al. 2004). It is very important to note that this opposite pattern of results was observed even when the same set of words was classified according to either dictionary measures or subjective ratings (i.e., NOM, Experiment 1).

In our opinion, there are two main points that deserve to be discussed in detail. The first one is related to the criterion to select ambiguous/unambiguous words (i.e., dictionary approach vs subjective ratings). The second one is the advantage in processing found for words with unrelated meanings. Regarding the first point, our findings suggest that researchers should be cautious when defining semantic ambiguity, since the type of definition employed (and the way to operationalize the distinction between ambiguous and unambiguous words) may influence the experimental results obtained. In turn, it would have consequences for our understanding and theoretical explanations of how ambiguous words are processed. There are several objective and subjective measures to operationalize semantic ambiguity, and each of them has some advantages and limitations. Dictionaries provide objective (i.e., standardized) measures of ambiguity. This is a clear advantage, as it allows comparability across studies. Concerning dictionary senses, it seems to be a reliable index of the number of word senses known by the speakers, as revealed by the high correlation between both measures (Fraga et al. 2017). This makes dictionaries a practical tool for researchers interested in analysing the effects of number of senses on word processing. However, there are limitations in the measure of number of meanings (i.e., entries) provided by dictionaries. As previously stated, dictionaries list some word meanings that are unknown to the majority of the speakers (e.g., jargon, old-fashioned and low-frequency meanings) while they do not include novel unrelated word meanings that are commonly used in everyday life. This may cause that words having one/many dictionary entries do not correspond to words having one/multiple unrelated meanings according to the speakers' knowledge, as we have observed in Experiment 1. In fact, the results of Experiment 1 showed that the pattern of results can be reversed depending on the criterion used to classify the words as unambiguous or ambiguous. We reasoned that the ambiguity disadvantage found in that experiment could be accounted for by the misdistribution of subjectively unambiguous and ambiguous words into words with one/more than one dictionary entries. The results of the present study do not allow us to draw firm conclusions on the extent to which the ambiguity disadvantage found in past studies could be explained by the same reason, as our Spanish stimuli are not directly comparable to the English stimuli used in those studies. However, to explore this point further, we decided to examine the results of those studies that used a dictionary measure (i.e., a total of 11 experiments, Armstrong and Plaut 2008, 2011; Beretta et al. 2005; Rodd et al. 2002). On the one hand, there was an interaction between meanings and senses in several of these experiments (e.g. Armstrong and Plaut 2008, Experiment 2). On the other, we observed that words with many dictionary entries and few senses had -numerically- the slowest RTs of all 11 experiments. By contrast, in only one experiment did words with many dictionary entries and many senses show slower RTs than words with one dictionary entry (Exp. 3 from Armstrong and Plaut 2008). Such evidence suggest that the ambiguity disadvantage found in these studies was due mainly to words with many meanings and few senses, as occurred in our Experiment 1. Thus, it could be that, as in our Experiment 1, subjective NOM ratings were misdistributed across conditions in these studies, having the highest impact in the condition of many meanings and few senses, and thus leading to a large inhibition for such words. Further research is needed on this issue.

The above considerations suggest that the use of objective measures has important limitations. Given that, an alternative for the study of the processing of ambiguity is the use of subjective measures. Nevertheless, such measures are not exempt either of some methodological concerns. Indeed, in line with what Rodd et al. (2002) pointed out, we have observed that high NOM ratings can be assigned to both homonymous and polysemous words (Experiment 2). Therefore, although NOM seems to be a useful measure to discriminate ambiguous and unambiguous words, it is not adequate as a way to separate related senses and unrelated meanings. Researchers interested in such distinction have to rely on additional measures, such as ROM ratings. Consequently, it becomes evident that a proper categorization of ambiguous words should involve the use of several measures of semantic ambiguity.

The second point worth to be discussed here is the facilitation found in the present study for words with multiple unrelated meanings, as it has relevant theoretical implications. Indeed, this finding provides support for the enhanced semantic activation accounts of the phenomenon. One of them is that of Hino and Lupker (1996), which is based on the following assumptions. According to this account, ambiguous words are represented by one orthographic representation linked to multiple semantic representations, one for each meaning. Furthermore, when a word is presented, activation would flow bidirectionally between the orthographic and semantic levels. Finally, lexical decisions are supposed to be made on the basis of the activation at the orthographic level. Given these assumptions, the ambiguity advantage is quite straightforward: Ambiguous words would benefit from a larger semantic-to-orthographic feedback than unambiguous words, facilitating their orthographic processing and thus speeding up their recognition in LDT.

According to enhanced semantic activation accounts, the ambiguity advantage can be related to the semantic richness effects reported in the literature (see Pexman et al. 2013, for a review). Research on semantic richness has been mainly devoted to explore which dimensions of meaning affect word recognition and to study how orthography and semantics interact during word processing. The main finding of this line of research is that words associated with more semantic information are processed faster and more accurately. Indeed, it has been found that words having more semantic features elicit faster responses in several experimental tasks, such as LDT, naming and semantic categorization, than words with a lower number of semantic features (Yap et al. 2012). In addition, words with more semantic neighbours or with denser semantic neighbourhoods are facilitated in naming and lexical decision tasks (Yap et al. 2012). Furthermore, there is an advantage in LDT for words having more and stronger visual associations (e.g., Hargreaves and Pexman 2012). Following this line, the number of unrelated meanings or related senses that a word has could be regarded as another semantic richness dimension, as having more meanings or senses lead to faster and more accurately responses in LDT.

Apart from semantic enhanced accounts, another possible explanation for the advantage is that, for a word with multiple meanings, there are more chances for it to be recognized as a word.[1] For example, if a word has four meanings, and the participant does not know one of these, there are still three meanings available to her/him as a means of deciding that the item is a word. However, if a word has only one meaning and this is not known to the participant, she/he will necessarily decide that the item is a non-word. This would be seen especially in the accuracy measures, and it does, since we found a better accuracy for ambiguous words in the present study. But it would also speed up RTs for words with multiple meanings, because with greater opportunities for encountering any of the meanings of a word, we might also predict that this would happen faster than when the sole meaning for a word has to be found.

Importantly, the ambiguity advantage it is incompatible with the predictions of Rodd et al. (2004), as their model proposes a disadvantage for words with multiple meanings. Nevertheless, we think that Rodd et al.'s model could account for the ambiguity advantage if some changes were made on it. In what follows we are going to develop this point.

As previously described, the model of Rodd et al. assumes that the disadvantage for ambiguous words is produced, in part, because a blend of all the meanings of these words would be activated in the early stages of the semantic network's settling. Given its unspecific semantic nature, this state would not be enough to recognize an ambiguous word in LDT. Thus, in order to make a lexical decision, the semantic network should escape from this state and settle into a stable semantic pattern representing only one of the meanings of the ambiguous word. As a result of the need to move away from such state, and assuming that the multiple semantic representations of an ambiguous word compete during this process, slower RTs would be expected for ambiguous words in comparison to unambiguous words. Nevertheless, we would like to point out that some data from simulation studies suggest that this blend state may be sufficient to recognize a word in LDT. Part from this evidence comes from the modelling work of Joordens and Besner (1994). These authors tried to simulate the ambiguity advantage using the distributed memory model of Masson (1991), which included two word processing modules, one representing the orthography of the word, and the other representing its meaning. The model was trained with stimuli having a one-to-many mapping between their orthographic and semantic representations (i.e., representing ambiguous words), and stimuli having a one-to-one relation (i.e., representing unambiguous words). After the training phase, the model was presented with ambiguous and unambiguous words, and the performance of the network was assessed. The results of the test phase showed that the model failed in most of the cases to retrieve only one of the meanings of ambiguous words. In those cases, the model

---

[1] We thank an anonymous reviewer for this suggestion.

settled into a pattern of semantic activation containing information from the two meanings, that is, a blend state. Interestingly, the model activated faster the blend state of an ambiguous word than the stable semantic pattern of an unambiguous word.

Although blend states were initially considered errors of the model (i.e., an incorrect response in LDT), this assumption was later discussed in several commentaries (e.g., Besner and Joordens 1995; Masson and Borowsky 1995; Rueckl 1995). Overall, these commentaries suggested the possibility that, as blend states are thought to include information from multiple meanings, they would provide a strong familiarity cue to discriminate between words and nonwords. Accordingly, blend states would be enough to make lexical decisions. If we consider that blend states were reached faster by the network, the ambiguity advantage might be explained by a PDP model as that employed by Joordens and Besner (1994). Importantly, this possibility was later confirmed in the simulation study of Borowsky and Masson (1996), which used the same PDP model as that of Joordens and Besner. The authors assessed the level of familiarity during word recognition by using a feature of Hopfield networks called *energy*. The energy of the network was measured as the sum of activation of the orthographic and the semantic levels during the processing of an input, and such measure indicated the distance towards a stable pattern of activation. When the energy of the network reached a given criterion, a positive lexical decision was thought to occur. The training and test phases were similar to those conducted by Joordens and Besner. As a result, the authors observed that when the model was presented with an ambiguous word, the meaning units settled faster into a blend state in which both meanings of the ambiguous word were activated. As these states were similar to two stable patterns of semantic activation, they elicited a larger level of semantic activation than that produced by unambiguous words. Therefore, ambiguous words reached faster the energy criterion for a positive response than unambiguous words. In this way, the PDP model of Borowsky and Masson (1996) was able to account for the ambiguity advantage.

In sum, all this evidence lead us to think that the model of Rodd et al. (2002) could explain the ambiguity advantage if it assumes that the blend state is enough to make a lexical decision. However, it is important to note that this blend of meanings would not be sufficient to perform tasks requiring meaning specification, and thus no advantage would be expected for ambiguous words in such tasks. In this line, it has been found that ambiguous words are associated with slower responses in semantic categorization tasks (Hino et al. 2002). To account for this effect, Hino et al. (2002) pointed out that in order to determine if an ambiguous word belongs to a given category (e.g., is *novel* a living object?), each of its meanings must be activated, evaluated, and compared with the specific semantic category. As a result, more time would be consumed for these words in comparison to unambiguous words. A similar explanation can be given for the finding that participants tend to fixate longer on ambiguous words than on unambiguous words during text comprehension (e.g., Duffy et al. 1988). This is because the correct interpretation of a sentence containing an ambiguous word requires the disambiguation of its meaning. In contrast, this disambiguation process does not seem necessary to perform a LDT.

To conclude, the findings of the present study support the idea that multiple unrelated meanings facilitate word recognition. This pattern of results can be explained by semantic enhanced accounts, which assume that ambiguous words benefit during recognition either by an increase in the global activation at the semantic level (Borowsky and Masson 1996) or by a large semantic-to-orthographic feedback. On the contrary, these results are, at a first glance, incompatible with the model of Rodd et al. (2004). Nevertheless, we have argued that this model could account for the ambiguity advantage by making an additional assumption: namely, that a blend state is enough to make lexical decisions. On the other hand, the evidence presented in this study suggest that the approach employed to select ambiguous words may affect the experimental results obtained. As such, researchers should consider carefully the advantages and disadvantages of the available ambiguity measures before selecting their experimental items.

Funding Funding was provided by Spanish Ministry of Economy and Competitiveness (Grant No. PSI2015-63525-P), Research Promotion Program of the Universitat Rovira i Virgili (Grant No. 2014PFR-URV-B2-37), and Universitat Rovira i Virgili (Grant No. 2015PMF-PIPF-16).

References

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 273–278).

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2223–2228).

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, *36*(4), 484–504.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Hillsdale, NJ: Erlbaum.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/bf03193014.

Bates D., Maechler M., Bolker B., & Walker S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1, 1–10. http://CRAN.R-project.org/package=lme4.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research*, *24*(1), 57–65.

Besner, D., & Joordens, S. (1995). Wrestling with ambiguity—further reflections: Reply to Masson and Borowsky (1995) and Rueckl (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 515–519.

Borowsky, R., & Masson, M. E. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 63–85.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, *45*(4), 1246–1258.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of memory and language*, *27*(4), 429–446.

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review*, *22*(1), 13–37.

Ferraro, F. R., & Kellas, G. (1990). Normative data for number of word meanings. *Behavior Research Methods, Instruments, & Computers*, *22*(6), 491–498. https://doi.org/10.3758/bf03204432.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition*, *4*(1), 53–61.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124.

Fraga, I., Padrón, I., Perea, M., & Comesaña, M. (2017). I saw this somewhere else: The Spanish Ambiguous Words (SAW) database. *Lingua*, *185*, 1–10.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281.

González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods*, *46*(2), 517–525.

Hargreaves, I. S., & Pexman, P. M. (2012). Does richness lose its luster? Effects of extensive practice on semantic richness in visual word recognition. *Frontiers in Human Neuroscience*, *6*(234), 1–11.

Haro, J., Demestre, J., Boada, R., & Ferré, P. (2017). ERP and behavioral effects of semantic ambiguity in a lexical decision task. *Journal of Neurolinguistics*, *44*, 190–202.

Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics*, *38*(2), 457–475.

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: When does relatedness matter? *Canadian Journal of Experimental Psychology*, *64*(3), 180–196.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1331.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 686–713.

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language*, *55*(2), 247–273.

Jager, B., & Cleland, A. A. (2016). Polysemy advantage with abstract but not concrete words. *Journal of psycholinguistic research*, *45*, 143–156.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*(2), 278–305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior*, *14*(5), 534–537.

Joordens, S., & Besner, D. (1994). When banking on meaning is not (yet) money in the bank: Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1051–1062.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 601–609.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, *20*(1), 1–24.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models: R package version 2.0-6. http://CRAN.R-project.org/package=lmerTest.

Lin, C.-J. C., & Ahrens, K. (2005). How many meanings does a word have? Meaning estimation in Chinese and English. In J. W. Minett & W. S.-Y. Wang (Eds.), *Language acquisition, change and emergence: Essays in evolutionary linguistics* (pp. 437–464). Hong Kong: City University of Hong Kong Press.

Lin, C.-J. C., & Ahrens, K. (2010). Ambiguity advantage revisited: Two meanings are better than one when accessing Chinese nouns. *Journal of Psycholinguistic Research*, *39*(1), 1–19.

Masson, M. E., & Borowsky, R. (1995). Unsettling questions about semantic ambiguity in connectionist models: Comment on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 509–514.

Masson, M. E. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Erlbaum.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition*, *17*(2), 141–147.

Parks, R., Ray, J., & Bland, S (1998). Wordsmyth English dictionary—Thesaurus. http://www.wordsmyth.net/ [2016, May].

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1252–1270.

Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience*, *7*, 1–3.

Real Academia Española. (2014). Diccionario de la lengua española (23rd ed.). Madrid, Spain.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*(2), 245–266.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, *28*(1), 89–104.

Rueckl, J. G. (1995). Ambiguity and connectionist networks: Still settling into a solution: Comment on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 501–508.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487–494.

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, *96*(2), 316–340.

Tamminen, J., Cleland, A. A., Quinlan, P. T., & Gaskell, M. G. (2006). Processing semantic ambiguity: Different loci for meanings and senses. In *Proceedings of the twenty-eighth annual conference of the cognitive science society* (pp. 2222–2227).

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*(72), 1–10.

2.3. Study 3: Haro, J., Demestre, J., Boada, R., & Ferré, P. (2017). ERP and behavioral effects of semantic ambiguity in a lexical decision task. *Journal of Neurolinguistics*, *44*, 190-202.

2.3.1 Description of the study

There is some inconsistent evidence regarding the effect of ROM in word recognition. Some studies have observed a facilitation for words with related meanings and an inhibition for words with unrelated meanings (e. g., Rodd et al., 2002), while others have reported a similar facilitation for both types of ambiguous words (e.g., Hino et al., 2010). The aim of the present study was to examine the role of ROM in word recognition, by testing the assumptions of Rodd et al. (2004)'s model, according to which polysemes and homonyms are differently represented and processed. To do so, we conducted an LDT experiment comparing polysemes and homonyms in which we also registered neurophysiological correlates (i.e., ERPs). This was the first time that ERP correlates of ROM were registered during word processing in isolation. In particular, we focused on an ERP component associated with semantic processing (i.e., N400).

On the other hand, we also tested some assumptions of the most relevant models that can account for the ambiguity (NOM) advantage: the model of Kawamoto et al. (1994), the model of Borowsky & Masson (1996) and the semantic feedback hypothesis (Balota et al., 1991; Hino & Lupker, 1996). These models represent two different points of view. While Kawamoto et al. (1994) suggest that ambiguous words benefit from creating stronger orthographic links, the other two models (Borowsky & Masson, 1996; Hino & Lupker, 1996) suggest that the ambiguity advantage is due to the fact that ambiguous words trigger a larger amount of semantic activation in comparison to unambiguous words. To contrast both points of view, we compared EEG amplitudes elicited by ambiguous and unambiguous words in a component associated with orthographic processing (i.e., N200) and in a component associated with semantic processing (i.e., N400).

2.3.2 Predictions

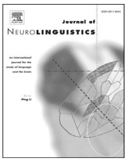
In accordance to previous studies that have relied on subjective norms, we expected a similar facilitation on RTs for polysemes and homonyms with respect to unambiguous words. More importantly, we also expected similar N400 amplitudes for both types of words. Regarding the models' predictions in relation to the ambiguity advantage, if Kawamoto et al. (1994) were right, we would observe differences between ambiguous and unambiguous words in the modulation of the N200. In contrast, according to semantic enhanced activation accounts (i.e., Borowsky & Masson, 1996; Hino & Lupker, 1996), ambiguous words were expected to differ with respect to unambiguous words in N400 amplitudes.

Journal of Neurolinguistics 44 (2017) 190–202

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling



ERP and behavioral effects of semantic ambiguity in a lexical decision task

Juan Haro*, Josep Demestre, Roger Boada, Pilar Ferré

Research Center for Behavior Assessment (CRAMC), Department of Psychology, Universitat Rovira i Virgili, Tarragona, Spain

ARTICLE INFO

Article history:
Received 7 November 2016
Received in revised form 25 May 2017
Accepted 21 June 2017
Available online 30 June 2017

Keywords:
Semantic ambiguity
Ambiguity advantage
Meanings relatedness
Polysemy
Homonymy
Word recognition
ERP
N400

ABSTRACT

In the present study we examined electrophysiological and behavioral correlates of ambiguous word processing. In a lexical decision task, participants were presented with ambiguous words with unrelated meanings (i.e., homonyms; e.g., *bat*), ambiguous words with related meanings (i.e., polysemes; e.g., *newspaper*), and unambiguous words (e.g., *guitar*). Ambiguous words elicited larger N400 amplitudes than unambiguous words and showed an advantage in RTs. Importantly, no differences were found between homonyms and polysemes, on either N400 amplitudes or in RTs. These results suggest that ambiguous words, regardless of the relatedness between their meanings, benefit from enhanced semantic activation in comparison to unambiguous words during word recognition.

1. Introduction

Understanding how meaning is retrieved from printed words and how it is represented in the mind are two primary goals of word recognition research. A fruitful line of research has been devoted to elucidate how orthography and semantics interact during word recognition, and to examine which semantic variables play a role in this process. Among such variables, semantic ambiguity has been one of the most studied. Semantic ambiguity refers to a linguistic phenomenon in which an orthographic form is mapped to more than one meaning (e.g., the word *pupil*, which means both *a student* and *the opening in the iris of the eye*). Given this one-to-many relation between orthography and meaning, semantic ambiguity poses intriguing questions for word recognition research. One central issue is whether ambiguous words have one or multiple lexical/semantic representations. For instance, are both meanings of the word *pupil* (e.g., *student* and *part of the eye*) included in the same lexical/semantic representation, or are they listed in separate lexical/semantic representations? A further crucial question is how orthography and semantics interact during the recognition of ambiguous words. Do the meanings *student* and *part of the eye* compete during the recognition of the word *pupil*? Or rather, does having two meanings, and thus more semantic information, facilitate the recognition of such a word? The aim of the present study was to shed some light on these questions by examining the behavioral and electrophysiological correlates of ambiguous word processing.

---

* Corresponding author. Research Center for Behavior Assessment (CRAMC), Department of Psychology, Universitat Rovira i Virgili, Crta. de Valls s/n, Campus Sescelades, 43007, Tarragona, Spain.
E-mail address: juan.haro@urv.cat (J. Haro).

http://dx.doi.org/10.1016/j.jneuroling.2017.06.001
0911-6044/

Rubenstein, Garfield, and Millikan's (1970) were the first to address some of these issues. Its main finding was that ambiguous words were recognized faster than unambiguous ones in a lexical decision task (LDT; a task in which participants decide whether a string of letters is a real word or not). Since the pioneering work of Rubenstein et al. (1970), there have been many reports of such a facilitation for ambiguous words in LDT (i.e., the *ambiguity advantage*) (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Pexman, Hino, & Lupker, 2004).

The ambiguity advantage appears to be a consistent effect in the literature (see, however, Rodd, Gaskell, & Marslen-Wilson, 2002). For this reason, it has had significant implications for models of word recognition, and has also received different explanations. Some accounts propose that ambiguity effects are located at the surface level of the representation of words (i.e., orthography/phonology), whereas others suggest that they are located at the semantic level of representation (see Armstrong & Plaut, 2016, for an overview). With respect to the former, it is worth mentioning the Parallel Distributed Processing (PDP) model of word recognition proposed by Kawamoto, Farrar, and Kello (1994). This model consists of two processing modules representing the orthography and semantics of words. The model was trained with pairs of activation patterns representing the form and meaning of the words. After the training phase, the authors assessed the performance of the network by presenting just the orthographic pattern of the words, observing that ambiguous words reached the criterion for a lexical decision faster than unambiguous words (i.e., the orthographic units of the model achieved their maximum level of activation faster when they were presented with an ambiguous word). To explain this behavior, the authors showed that the network tried to compensate for the inconsistent orthographic-to-semantic relation for ambiguous words (i.e., one orthographic form associated with multiple meanings) by strengthening the connection weights between their orthographic units. These stronger connection weights between orthographic units would serve to speed up the settling of the orthographic representation of ambiguous words, hence facilitating lexical decisions.

With respect to those accounts that have focused on semantics, it has been suggested that there would be an advantage for ambiguous words during word recognition because they elicit a larger amount of semantic activation (i.e., *semantic-based accounts*; e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996). For instance, based on interactive activation principles, several authors have proposed that after the presentation of an orthographic input, the activation would flow bidirectionally between the orthographic and semantic levels (Balota, Ferraro, & Connor, 1991; McClelland & Rumelhart, 1981). In addition, they assumed that a word would be recognized in a LDT when the activation of its orthographic representation reached a recognition threshold. With these assumptions in place, the explanation of the ambiguity advantage is straightforward: because ambiguous words have more than one semantic representation, they would cause a larger semantics-to-orthography feedback than unambiguous words, and thus would reach the orthographic recognition threshold faster. A similar account was provided by the PDP model of Borowsky and Masson (1996). In this model, words were represented as patterns of activation across orthographic, phonological and semantic processing units. Additionally, a word was thought to be recognized when the level of activation of the network reached a given threshold. The level of activation of the network indicated the distance from the current state of the network to the pattern of orthographic and semantic activation corresponding to a known word; that is, the higher the activation of the network, the lower the distance to a learned pattern. The simulation data showed interesting behavior when ambiguous words were presented to the model, because in those cases the meaning units of the network settled faster into a state in which the two meanings of the ambiguous word were partially activated. Since these blend states were similar to both learned semantic patterns of the word, ambiguous words elicited more semantic activation and reached the criterion for a lexical decision faster than unambiguous words.

It should be noted that according to semantic-based accounts, the ambiguity advantage is closely related to the so-called semantic richness effects reported in word recognition research. Work on semantic richness is devoted to examine to what extent the amount of semantic information of a word influences its recognition (Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008; Pexman, Siakaluk, & Yap, 2013). Semantic richness effects in behavioral responses are quite homogeneous, in that words having more or richer semantic information (e.g., number of semantic features, number of semantic neighbors, or number of word associates) are associated with faster response latencies in a number of experimental tasks, such as LDT, naming and semantic categorization (Pexman et al., 2008). In addition, semantic richness effects have also been found in EEG studies. Particularly, the amount of semantic information a word contains seems to modulate the N400 component, a negative-going potential that is thought to reflect mainly semantic processing (see Kutas & Federmeier, 2011 for a review). For example, there is evidence, a) that concrete words elicit larger N400 amplitudes than abstract words (Kounios & Holcomb, 1994; West & Holcomb, 2000), b) that words with many semantic features are associated with larger N400 amplitudes than words with few semantic features (Amsel, 2011; Rabovsky, Sommer, & Rahman, 2012), and c) that words with many associates show a larger N400 than words with few associates (Laszlo & Federmeier, 2011; Müller, Duñabeitia, & Carreiras, 2010).

Taking into account the above evidence, it follows that the more or richer semantic information a word has, the more semantic activation it engages, and the larger the N400 it elicits (see, however, Taler, Kousaie, & Lopez Zunini, 2013). In fact, it has been suggested that the N400 component may reflect the amount of semantic activity before the orthographic and semantic levels have settled, thus providing a temporal window into the activity generated by a stimulus in a distributed, cascaded, semantic system (Laszlo & Federmeier, 2011). Therefore, it is reasonable to think that if semantic-based accounts of the ambiguity advantage are correct, ambiguous words would cause a larger N400 than unambiguous words, as the former would engage a larger amount of semantic activation during word recognition than the latter. In contrast, if ambiguity effects are located at the orthographic level of representation (i.e., ambiguous words benefit from having stronger orthographic-to-orthographic connections), as suggested by Kawamoto et al. (1994), one would expect differences between ambiguous and unambiguous words on ERP components associated with orthographic processing. One of these components is the N200, a negative-going component peaking at about 200 ms and which seems sensitive to orthographic processing (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Kramer & Donchin, 1987; Simon, Bernard, Lalonde, & Rebaï, 2006). For instance, N200 amplitudes are larger for orthographic stimuli (e.g., consonant strings and words) than for non-orthographic stimuli (e.g., symbols) (Bentin et al., 1999). Thus, following Kawamoto et al. (1994)'s model, ambiguous and unambiguous words should elicit a distinct pattern in the N200. The main aim of the present study was to test these two hypotheses regarding the source of the ambiguity advantage. To do so, we compared the amplitude of the N200 and the N400 elicited by ambiguous and unambiguous words while participants performed a LDT.

A second aim relates to the existence of distinct types of ambiguity. Indeed, semantic ambiguity is not a homogenous phenomenon, as not all ambiguous words are qualitatively similar. In particular, the degree of relatedness between the different meanings of an ambiguous word can vary widely. In the linguistics literature, ambiguous words have been categorized into at least two main classes: homonyms and polysemes. Homonyms have been defined as ambiguous words with unrelated meanings; for example, the homonym *yard* means both *a unit of measure* and *the ground that surrounds a house*, meanings that are clearly unrelated. On the other hand, polysemes have been defined as ambiguous words with related meanings (also known as *senses*); for instance, the polyseme *newspaper* refers to a wide range of related meanings or senses: (a) *a publication, usually issued daily or weekly*; (b) *a business organization that prints and distributes such a publication*; (c) *a single issue of such a publication*, and (d) *the paper on which a newspaper has been printed*. Given this distinction, one issue for word recognition research is whether such a linguistic categorization has psychological validity.

There is no consensus as to how relatedness of meanings (hereafter, ROM) affects ambiguous word recognition. On the one hand, some experimental data indicate that homonyms and polysemes are processed differently. A strong piece of evidence for this distinction can be found in Rodd et al. (2002)'s work, where the authors observed a facilitation for polysemes (i.e., polysemy or sense advantage) along with an inhibition for homonyms (i.e., homonymy or ambiguity disadvantage) in LDT. To account for these results, Rodd, Gaskell, and Marslen-Wilson (2004) developed a model of ambiguous word recognition, according to which polysemes would benefit during word recognition from having a single, richer semantic representation containing all their senses, whereas the separate semantic representations for homonyms would compete during word recognition. Importantly, Rodd et al.'s model obtained further support from subsequent LDT studies (Armstrong & Plaut, 2008, 2011; Klepousniotou & Baum, 2007; Tamminen, Cleland, Quinlan, & Gaskell, 2006). In addition, there is some neurophysiological evidence supporting it. For instance, Beretta, Fiorentino, and Poeppel (2005) found differences between polysemes and homonyms on the M350, a MEG component that reflects lexical processing and whose latencies are thought to be comparable to N400 amplitudes (Pylkkänen & Marantz, 2003). Specifically, in that study words with multiple related senses (i.e., polysemes) were seen to elicit earlier M350 peak latencies than words with few related senses. Furthermore, words with more than one meaning (i.e., homonyms) showed later M350 peak latencies than words with a single meaning. In contrast to the above findings, there are reports showing that polysemes and homonyms are processed similarly. In particular, several LDT studies have found that both polysemes and homonyms are recognized faster, and equally so, compared to unambiguous words (Hino, Kusunose, & Lupker, 2010; Hino, Pexman, & Lupker, 2006; Pexman et al., 2004). These authors, then, suggest that having multiple meanings, regardless of their ROM, leads to a stronger semantic-to-orthographic feedback during word recognition, facilitating orthographic processing and thus speeding up lexical decisions (Hino et al., 2010).

The second aim of the present study was to further explore the distinction between polysemes and homonyms by using ERP. If ROM does not affect the semantic activation of ambiguous words, as some of the above mentioned behavioral studies suggest (Hino et al., 2010, 2006; Pexman et al., 2004), no differences on the N400 between homonyms and polysemes should be expected, given that the N400 seems to be sensitive to semantic activation during word processing (e.g., Laszlo & Federmeier, 2011; Rabovsky et al., 2012). However, it might be that differences between homonyms and polysemes can be observed with electrophysiological measurements, as they are more sensitive than RTs (e.g., Chen, Shu, Liu, Zhao, & Li, 2007). In this case, we might expect that ROM modulates the amplitude of the N400.

To sum up, the purpose of the present study was to examine the behavioral and EEG correlates of ambiguous word processing by using a LDT. To do so, 1) we compared behavioral responses (RTs and %E) and EEG responses (the N200 and the N400) between ambiguous and unambiguous words, and 2) we compared behavioral responses (RTs and %E) and EEG responses (N400 amplitudes) between ambiguous words that differ in their ROM (i.e., homonyms vs polysemes). It should be noted that the present ERP study is not the first to examine ambiguous word processing. Indeed, some previous ERP studies have analyzed the neural correlates of ambiguous word processing by using a semantic priming paradigm (e.g. Klepousniotou, Pike, Steinhauer, & Gracco, 2012; MacGregor, Bouwsema, & Klepousniotou, 2015). For instance, Klepousniotou et al. (2012) compared the N400 elicited by polysemes (e.g., *arm*) and homonyms (e.g., *ball*) when they were preceded by a related prime (e.g., *wrist-arm*, *green-mold*) relative to when an unrelated word served as prime (e.g., *reef-arm*, *energy-mold*). In addition, they manipulated the dominance of the prime, using words related either to the dominant meaning of the ambiguous word (e.g., *hit-ball*) or to its subordinate meaning (e.g., *dance-ball*). By doing so, they were able to examine the time course of the activation of the distinct meanings of ambiguous words during processing. In contrast, the present study was designed to explore whether ambiguity benefits lexical access and whether this benefit is modulated by the degree of relatedness between the distinct meanings of the ambiguous words. To our knowledge, the present work is the first ERP study to compare the processing of polysemes and homonyms in isolation.

Table 1
Characteristics of the stimuli used in the experiment (standard deviations are shown in parentheses).

	NOM	ROM	FRE	CTD	FAM	AoA	LNG	SYL	CON	OLD	NEI	NHF	BFQ	TFQ
Unambiguous words	1.13	–	1.17	0.8	5.39	6.37	5.72	2.47	4.87	1.57	7.74	1.23	5442.95	617.39
	(0.19)		(0.66)	(0.51)	(1.11)	(2.38)	(1.8)	(0.81)	(1.21)	(0.45)	(8.03)	(2.36)	(3207.86)	(708.59)
Ambiguous words	1.74	2.76	1.18	0.82	5.51	6.49	5.53	2.31	4.55	1.5	9.33	1.28	5553.1	803.9
	(0.19)	(1.17)	(0.44)	(0.34)	(0.75)	(1.8)	(1.03)	(0.55)	(0.73)	(0.39)	(9.52)	(2.06)	(3190.51)	(695.74)
Polysemes	1.71	3.76	1.19	0.84	5.45	6.47	5.53	2.27	4.54	1.49	8.88	0.97	5098.96	790.45
	(0.18)	(0.93)	(0.45)	(0.34)	(0.81)	(1.92)	(0.93)	(0.51)	(0.71)	(0.35)	(8.79)	(1.88)	(2299.69)	(516.31)
Homonyms	1.77	1.86	1.17	0.79	5.56	6.51	5.53	2.34	4.56	1.51	9.74	1.55	5959.44	815.94
	(0.19)	(0.34)	(0.44)	(0.34)	(0.69)	(1.71)	(1.13)	(0.58)	(0.75)	(0.44)	(10.22)	(2.2)	(3802.3)	(831.13)

Note. NOM = subjective Number-Of-Meanings ratings; ROM = subjective Relatedness-Of-Meanings ratings; FRE = log word frequency; CTD = log contextual diversity; FAM = familiarity; AoA = subjective age-of-acquisition; LNG = word length; SYL = number of syllables; CON = concreteness; OLD = old20; NEI = number of substitution neighbors; NHF = number of higher frequency substitution neighbors; BFQ = mean bigram frequency; TFQ = mean trigram frequency.

2. Method

2.1. Participants

Twenty-five Spanish speakers (21 women; mean age 20.6 years, $SD = 3.1$) from the Universitat Rovira i Virgili (Tarragona, Spain) participated in the experiment. They were undergraduate students and were paid 10€ for their participation. All had either normal or corrected-to-normal vision, had no language difficulties or history of neurological disease, and 24 were right-handed. Prior to the experiment, participants signed an informed consent.

2.2. Design and materials

The experimental stimulus set consisted of 152 Spanish words: 76 ambiguous words and 76 unambiguous words[1] (see the Appendix). Stimuli were categorized as ambiguous or unambiguous according to Number-Of-Meanings (NOM) ratings (e.g., Kellas, et al., 1988; Pexman et al., 2004). The common procedure to obtain NOM ratings is as follows. Participants are asked to indicate how many meanings a particular string of letters has. They make their ratings by using a 3-point Likert scale: (0) *the word has no meaning*, (1) *the word has one meaning*, or (2) *the word has more than one meaning*. Words with values close to 2 are classified as ambiguous, and words with values close to 1 are classified as unambiguous. We employed different sources to obtain NOM ratings. NOM ratings for 125 words were taken from Haro, Ferré, Boada, and Demestre (2017). NOM ratings for the remaining 27 words were provided by a group of 20 participants (15 women; mean age 22.3 years, $SD = 3.5$). According to this measure, unambiguous words had one meaning (NOM = 1.13, $SD = 0.19$) and ambiguous words had more than one meaning (NOM = 1.74, $SD = 0.19$), $t(144) = 19.68$, $p < 0.001$.

The set of 76 ambiguous words comprised 38 homonyms and 38 polysemes. The homonym/polyseme categorization was made on the basis of subjective ROM ratings, which were obtained from Haro et al. (2017). In that study, participants were asked to judge how related were the meanings of pairs of words, each pair containing the same ambiguous word and an associate related to one of its meanings (e.g., SIREN-*ambulance* [*warning alarm*] and SIREN-*sea* [*sea nymph*]). Participants were provided with a 9-point scale, ranging from 1 (unrelated meanings) to 9 (same meaning), to make their ratings. Using such a measure, homonyms are expected to have low ROM ratings, and polysemes high ROM ratings (for similar approaches, see Hino et al., 2010; Hino et al., 2006). Words with ROM ratings below 2.5 were categorized as homonyms, and those with ROM ratings above 2.5 were categorized as polysemes. Overall, the homonyms selected for this experiment averaged 1.86 ($SD = 0.34$) and the polysemes averaged 3.76 ($SD = 0.93$) on ROM ratings, $t(70) = 11.76$, $p < 0.001$. Importantly, homonyms and polysemes did not differ in NOM ratings, $t(70) = 1.38$, $p = 0.17$, which indicates that both types of ambiguous words had a similar number of meanings.

A large number of lexical and semantic variables that are known to affect word recognition were matched between ambiguous and unambiguous words, as well as between homonyms and polysemes (all $ps > 0.05$, see Table 1 for more details). These variables were drawn from several different sources. On the one hand, number of letters, number of syllables, logarithm of word frequency (log word frequency), mean Levenshtein distance of the 20 closest words (OLD20), number of neighbors, number of higher frequency neighbors, bigram frequency, trigram frequency, and logarithm of contextual diversity (log contextual diversity) were taken from EsPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). On the other hand, familiarity, concreteness and subjective age of acquisition were taken from Haro et al. (2017). Given that subjective age of acquisition ratings for 27 words were not available from Haro et al.'s database, we asked a sample of 20 participants (15 women; mean age 22.3 years, $SD = 3.5$) to provide them.

---

[1] Due to data loss, 4 ambiguous words and 2 unambiguous words were not included in the analyses.

Finally, we created a set of 152 pronounceable nonwords from the 152 experimental words, by using the Wuggy nonword generator (Keuleers & Brysbaert, 2010). Words and nonwords were matched in length, number of syllables, subsyllabic structure and transition frequencies.

2.3. Procedure

Participants performed a lexical decision task. Each trial began with an image of an eye displayed for 2000 ms, which indicated to participants that they were allowed to blink. The image was followed by a fixation point (i.e., "+") appearing in the center of the screen for 500 ms. Immediately after this, a string of letters (a word or a nonword) replaced the fixation point, and participants then had to decide whether the string was a Spanish word or not. They were instructed to press the "yes" labelled key of a keyboard with the right hand if the string of letters was a word, and to press the "no" labelled key of the keyboard with the left hand if it was not a word. The string of letters remained on the screen until participant's response or timeout (after 2000 ms). After responding, a feedback message (i.e., "ERROR" or "CORRECT") was displayed for 750 ms. The DMDX software (Forster & Forster, 2003) was used to display the stimuli and record the responses. The order of the experimental trials was randomized for each participant. Prior to the experiment, a practice block consisting of 12 trials (6 words and 6 nonwords) was presented. There were two brief breaks during the experiment.

2.4. EEG recording

Participants were seated in a comfortable chair in a sound attenuated and dimly illuminated room. The EEG was recorded from 32 Ag/AgCl electrodes attached to an elastic cap (ActiCap, Brain Products, Gilching, Germany) that was positioned according to the 10–20 system. One electrode was placed beneath the left eye to monitor blinking and vertical eye movements (VEOG), and another at the outer canthus of the right eye to monitor horizontal eye movements (HEOG). All scalp electrodes were referenced online to the right mastoid and re-referenced off-line to the average of the right and left mastoids. Electrode impedances were kept below 5 kΩ. All EEG and EOG channels were amplified using a actiCHamp amplifier (Brain Products Gilching, Germany).

Data was processed using BrainVision Analyzer 2 (Brain Products, Gilching, Germany). EEG was refiltered offline with a bandpass of 0.1–30 Hz 12 dB/oct, zerophase shift digital filter. Average ERPs were calculated per condition per participant from −100 to 800 ms relative to the onset of the word. A 100 ms pre-target period was used as baseline. Trials were rejected if the amplitude on any channel exceeded ±75 μV, and also if deflections on any channel exceeded ±150 μV. Less than 5% of trials were rejected after applying such trimming procedures. Only correct response trials were included in the averages.

3. Results

3.1. Behavioral results

The data from one participant with more than 15% of errors were discarded from both the behavioral and ERP analyses. RTs that exceeded 2 *SD* of each participant's mean were also rejected (3.7% of the data). In addition, we excluded two unambiguous words from the analyses due to a high percentage of errors (>70%). We then calculated the mean of RTs for correct responses and the mean %E across experimental conditions (see Table 2). Mean RTs and mean %E were analyzed with separated t-tests (paired t-tests for participants' analyses, and unpaired t-tests for items' analyses).

Ambiguous words were recognized faster than unambiguous words, $t_1(23) = 7.03$, $p < 0.001$, $t_2(142) = 3.05$, $p = 0.003$. Likewise, ambiguous words were recognized more accurately than unambiguous words, $t_1(23) = 4.79$, $p < 0.001$, $t_2(142) = 3.19$, $p = 0.002$. On the other hand, no differences were found between homonyms and polysemes, either in RTs, $t_1(23) = 1.04$, $p = 0.31$, $t_2(70) = 1.01$, $p = 0.32$, or in %E, $t_1(23) = 0.86$, $p = 0.40$, $t_2(70) = 0.67$, $p = 0.50$.

3.2. ERP results

ERP analyses were focused on the N200 and N400 components. N200 was measured by computing mean amplitudes between 150 and 250 ms after word onset, whereas the time range for the N400 component was established between 350 and 450 ms after word onset. Several repeated-measures analyses of variance (ANOVAs) were performed to examine

Table 2
Mean RT (in ms), and %E (percentage of error rates) (standard error in parentheses).

Ambiguity	ROM	RT	%E
Unambiguous words		590.51 (12.69)	8.33 (1.09)
Ambiguous words		572.70 (12.53)	3.25 (0.47)
	Polysemes	575.66 (11.92)	3.63 (0.57)
	Homonyms	570.20 (13.53)	2.92 (0.67)

differences between ambiguous and unambiguous words on the N200 and the N400 (i.e., ambiguity effects), and to examine differences between homonyms and polysemes on the N400 (i.e., ROM effects).

3.2.1. *Ambiguity effects*

An ANOVA was conducted with the factors of ambiguity (ambiguous and unambiguous words) and electrode site (28 electrodes). We also carried out other ANOVAs to examine separately midline electrodes: ambiguity (ambiguous and unambiguous words) x electrode site (Fz, Cz, Pz, Oz), and lateral electrodes: ambiguity (ambiguous and unambiguous words) x hemisphere (left/right) x electrode site (Fp1/Fp2, F3/F4, F7/F8, FC1/FC2, FC5/FC6, C3/C4, T7/T8, CP1/CP2, CP5/CP6, P3/P4, P7/P8, O1/O2). All factors were within-subjects. For effects involving more than one degree of freedom, Greenhouse-Geisser correction was applied (corrected p-values are reported). Grand average waveforms for ambiguous and unambiguous words are shown in Fig. 1.

3.2.1.1. N200. The analysis of the data from all the electrodes failed to show any difference between ambiguous and unambiguous words on N200 amplitudes, $F(1,23) = 0.08$, $MSE = 1.36$, $p = 0.78$. No other significant effects or interactions were found (all $ps > 0.1$).

3.2.1.2. N400. The analysis including data from all the electrodes revealed a main effect of ambiguity, $F(1,23) = 5.07$, $MSE = 84.22$, $p = 0.034$. Ambiguous words elicited larger N400s (−1.70 μV) than unambiguous words (−1.20 μV). No interaction was found between ambiguity and electrode site, $F(27,621) = 1.53$, $MSE = 5.26$, $p = 0.20$. The main effect of ambiguity on the N400 was also found in the analysis of midline electrodes, $F(1,23) = 4.95$, $MSE = 21.21$, $p = 0.036$, as well as in the analysis of lateral electrodes, $F(1, 23) = 4.95$, $MSE = 64.51$, $p = 0.036$. Of note, no significant interaction was found between ambiguity and hemisphere, $F(1,23) = 0.16$, $MSE = 0.29$, $p = 0.69$.

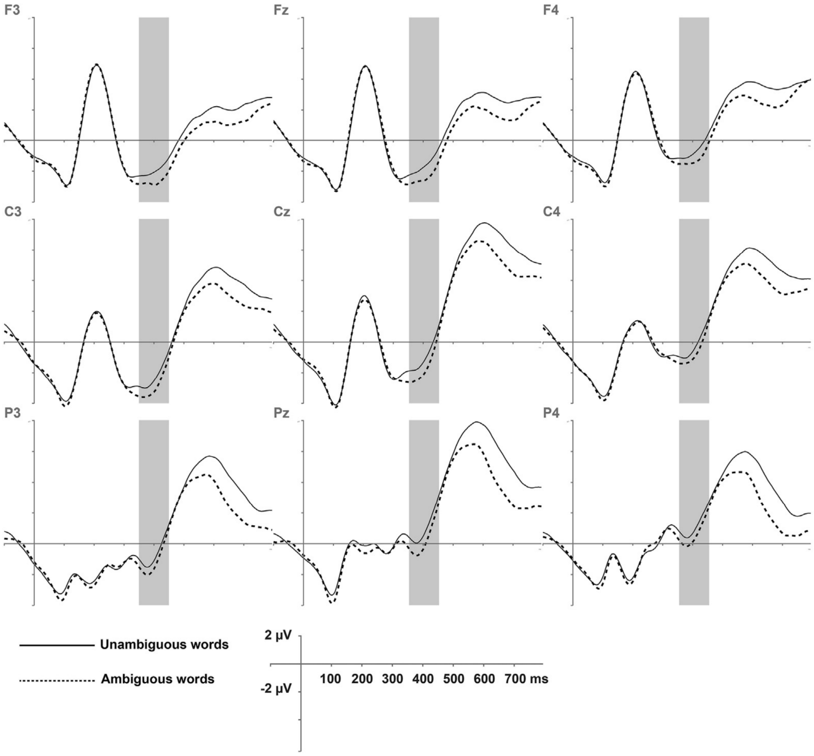


Fig. 1. Grand average waveforms for ambiguous and unambiguous words for nine representative electrodes (negativity is plotted down). The shaded area represents the time range for the N400 component (350–450 ms).

3.2.2. ROM effects

The same analyses as those conducted to examine ambiguity effects were conducted to compare the N400 elicited by homonyms and polysemes (i.e., ROM factor). Grand average waveforms for homonyms, polysemes and unambiguous words are shown in Fig. 2. The main effect of ROM did not reach significance in the analysis including data from all the electrodes, $F(1,23) = 0.16$, $MSE = 5.62$, $p = 0.70$. Homonyms and polysemes showed similar N400s (−1.75 μV vs −1.62 μV). No interaction was observed between ROM and electrode site, $F(27,621) = 1.41$, $MSE = 11.55$, $p = 0.24$. Concerning midline and lateral separate analyses, no main effect of ROM was found in the analysis of midline electrodes, $F(1,23) = 0.21$, $MSE = 2.06$, $p = 0.65$, nor in the analysis of lateral electrodes, $F(1,23) = 0.14$, $MSE = 3.90$, $p = 0.71$. Finally, the interaction between ROM and hemisphere was not significant, $F(1,23) = 0.01$, $MSE = 0.30$, $p = 0.93$. No other relevant effects or interactions were found.

4. Discussion

The aim of the present study was to obtain behavioral and EEG correlates of ambiguous word processing. As far as we know, this is the first ERP study to examine the processing of polysemes and homonyms in isolation. On the one hand, we compared ambiguous words to unambiguous words in a LDT. The results showed faster and more accurate behavioral responses and larger N400 amplitudes for ambiguous words in comparison to unambiguous words. In contrast, there were no differences in the N200 between both types of words. On the other hand, we examined ambiguous words differing in ROM. In particular, we compared ambiguous words with unrelated meanings (i.e., homonyms) to ambiguous words with related meanings (i.e., polysemes). The results showed that homonyms and polysemes exhibited a similar degree of facilitation in behavioral responses relative to unambiguous words. Furthermore, the two types of ambiguous words did not differ in the N400. In what follows, we will discuss separately the ambiguity effects and the ROM effects.

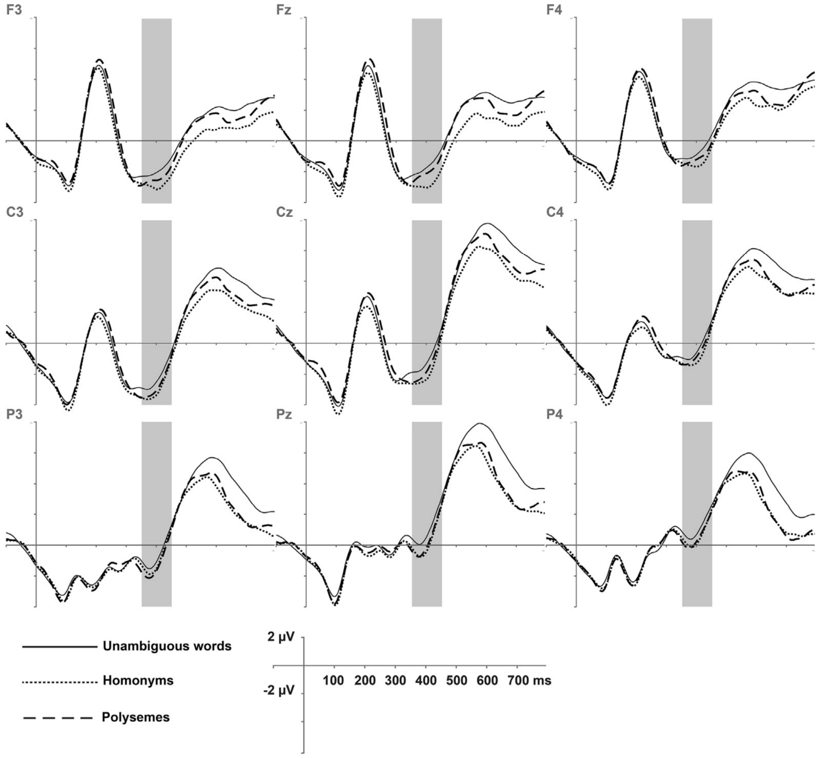


Fig. 2. Grand average waveforms for homonyms, polysemes and unambiguous words for nine representative electrodes (negativity is plotted down). The shaded area represents the time range for the N400 component (350–450 ms).

4.1. Ambiguity effects

In line with previous reports of a facilitation for ambiguous words in LDT, behavioral data showed that ambiguous words were recognized faster and more accurately than unambiguous words (Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 2002; Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Kellas et al., 1988; Millis & Button, 1989; Pexman et al., 2004). More importantly, EEG data revealed that semantic ambiguity modulated the N400, but not the N200. These electrophysiological findings contribute to a better understanding of the advantage for ambiguous words over unambiguous words in LDT. Thus, the absence of differences between ambiguous and unambiguous words on the N200 does not provide support for the models that locate the source of the ambiguity advantage at the orthographic level (e.g., Kawamoto et al., 1994). In contrast, considering that the N400 seems to reflect activity taking place at the semantic level of representation (e.g., Laszlo & Federmeier, 2011), the larger N400 observed for ambiguous words would indicate that they engage more semantic activation than unambiguous words during word recognition. This evidence is compatible with the above described semantic accounts of the ambiguity advantage. Such accounts suggest that as ambiguous words are represented by more than one semantic representation, they would benefit during word recognition either from an increase in the global activation at the semantic level of representation (Borowsky & Masson, 1996) or from a larger semantic-to-orthographic feedback (Hino & Lupker, 1996).

It is also worth noting here that the facilitation effect found for ambiguous words resembles the effects of semantic richness in word recognition; that is, the more or richer semantic information a word has, the faster is its recognition (Pexman et al., 2008). This facilitative effect has been demonstrated with different manipulations and variables, such as the number of semantic features (e.g., Amsel, 2011), number of associates (e.g., Müller et al., 2010), and concreteness (e.g., Kounios & Holcomb, 1994), among others. In addition, the N400 seems to be modulated by some of these variables, with increased N400 amplitudes for semantically richer words (Rabovsky et al., 2012), in agreement with the present electrophysiological data. In light of these findings, it seems that the advantage for ambiguous words fits well within semantic richness effects in word recognition and may be explained by the same mechanisms. To explore this issue in greater depth, we measured the correlation between NOM and some relevant semantic richness variables (i.e., number of word associates and contextual diversity measures). NOM ratings and number of word associates were obtained from Haro et al. (2017), whereas log contextual diversity was obtained from Duchon et al. (2013). We found a significant correlation between NOM ratings and semantic richness variables: number of word associates, $r = 0.12$, $p = 0.004$, and log contextual diversity, $r = 0.40$, $p < 0.001$. In addition to that, we also conducted a hierarchical regression analysis to examine if NOM influences lexical decision times beyond the effect of those semantic richness variables to which it is correlated. Of note, given that the experimental stimuli of our study were matched on a large number of variables, we conducted this analysis with a more heterogeneous (and larger) set of words. Hence, the regression analysis was conducted with the 260 words of the database of Haro et al. (2017), for which lexical decision data were available from González-Nosti, Barbón, Rodríguez-Ferreiro, and Cuetos (2014), a large LDT study of Spanish words. The dependent variable was the RT for lexical decisions. The independent variables were entered as predictors of RTs on three steps. Word length, number of syllables, log word frequency, number of neighbors, OLD20, bigram frequency, and trigram frequency were entered as predictors in the first step. Number of associates, concreteness, and log contextual diversity were entered in the second step. Finally, NOM ratings were entered in the third step. Number of syllables, log word frequency and number of neighbors, OLD20, bigram frequency, and trigram frequency were obtained from Duchon et al. (2013), whereas the ratings for the rest of variables were taken from Haro et al. (2017). The results revealed a significant facilitative effect of NOM ratings on RTs, $\beta = -0.14$, $t = 2.37$, $p = 0.019$ (detailed results of the regression analysis are presented in Appendix B). In sum, these findings show that number of meanings is correlated with some semantic richness variables and that the facilitative effect of NOM on lexical decision times cannot be explained by the effect of such variables. We consider that these findings provide further support for considering NOM as a semantic richness variable and for its unique role in contributing to LDT performance.

4.2. ROM effects

The results of the present study failed to show any effect of ROM either on behavioral or electrophysiological measures. This null effect is in line with several studies showing that homonyms and polysemes are similarly processed in LDT, with faster responses for both types of ambiguous words in comparison to unambiguous words and with no differences between them (Hino et al., 2006, 2010; Pexman et al., 2004). In contrast, this finding is incompatible with other studies that reported a distinct response pattern for homonyms and polysemes, both in behavioral performance (Armstrong & Plaut, 2008, 2011; Klepousniotou & Baum, 2007; Rodd et al., 2002; Tamminen et al., 2006) and in neurophysiological data (Beretta et al., 2005). A possible explanation for these divergent results may be the different approach employed to the categorization of ambiguous words across studies. In the present study, as well as in those where no differences were found between both types of ambiguous words, homonyms and polysemes were categorized according to subjective ROM measures. This represents a crucial methodological difference with respect to those LDT studies showing that ROM affects word processing, given that they mainly relied on dictionary definitions to classify the words (see, however, Rodd et al., 2002, Experiment 1). Such an approach is based on the assumption that unrelated meanings are listed in separate dictionary entries, whereas related meanings are listed under the same dictionary entry. Within this approach, then, homonyms are taken to be words with more than one dictionary entry, whereas polysemes are words having many dictionary senses within a single entry.

Although more research is needed to compare directly the experimental effects of using these two distinct criteria, an interesting finding is that subjective measures of semantic ambiguity seem to be better predictors of lexical decision times than dictionary measures (Fraga, Padrón, Perea, & Comesaña, 2017).

Another possible explanation for the null ROM effect is that LDT is a task that does not engage very much semantic processing. Indeed, there is some evidence showing that ambiguous word processing may be modulated by the requirements of the experimental task. For instance, in contrast to the ambiguity advantage commonly found in LDT, ambiguous words are usually responded to more slowly than unambiguous words in more semantically engaging tasks, such as semantic categorization, sense judgement and semantic relatedness tasks (see Eddington & Tokowicz, 2015, for a review). These tasks, unlike LDT, usually require a specific meaning of the ambiguous word to be activated. Consequently, this may increase the competition between the multiple meanings of the ambiguous words, leading to slower responses for ambiguous words in comparison to unambiguous words. Furthermore, a significant ROM effect has also been observed in these tasks. For example, Hino et al. (2006) found than homonyms were responded to more slowly and less accurately than unambiguous words in two semantic categorization tasks (using the semantic category “living thing” [Experiment 2] and “human related” [Experiment 5]). In contrast, polysemes showed faster and more accurate responses than homonyms. In a similar vein, Brown (2008) reported that pairs of homonym verb phrases (e.g., *banked the plane – blanked the money*) were responded to more slowly than pairs of polyseme verb phrases (e.g., *broke the glass – broke the radio*) in a sense judgment task. So, this evidence, although limited, suggests that ROM effects may emerge in tasks requiring exhaustive semantic activation.

Taking all the above into consideration, the results of the present study suggest that the number of meanings (i.e., ambiguity), but not ROM (i.e., the distinction between polysemes and homonyms), influences word recognition when it is assessed with a LDT. These findings have implications for models of semantic ambiguity processing and representation. In particular, the null ROM effect is a challenge for Rodd et al. (2004)'s model, since it postulates that polysemes are represented and processed in a different way from homonyms. Namely, it assumes that all the related meanings of a polyseme are stored in a single, richer semantic representation (i.e., single shared representation for polysemes), whereas the unrelated meanings of a homonym would be stored in separated semantic representations. Hence the model predicts an advantage for polysemes in LDT, as these words would benefit from having a single, richer semantic representation, and a disadvantage for homonyms, since their multiple, unrelated semantic representations are thought to compete during word processing. Importantly, in contrast to Rodd et al. (2004), some authors have suggested that both types of ambiguous words may be represented similarly, that is, with each homonym or polyseme meaning having a separate entry in the mental lexicon (i.e., separated representation for polysemes; e.g., Klein & Murphy, 2001, 2002). Thus, according to interactive activation principles, if each separate meaning of a homonym or polyseme provides an independent stream of feedback to its linked orthographic representation, both types of ambiguous words would trigger a similar amount of semantic feedback and, thus, no differences should be expected between them in LDT (Hino et al., 2010). Hence, the null ROM effect found in the present study seems compatible with proposals claiming that homonyms and polysemes are represented similarly, that is, through separate representations.

To sum up, in the present study we have shown that ambiguous words elicit faster and more accurate behavioral responses and larger N400s than unambiguous words in a LDT. This suggests that the cause of the ambiguity advantage is that ambiguous words engage a larger amount of semantic activation during word recognition than unambiguous words. In addition, we have observed no differences between homonyms and polysemes, either in behavioral or electrophysiological data. This seems to indicate that ROM does not affect ambiguous word recognition in lexical decision tasks, and that both types of ambiguous words benefit from triggering a similar amount of semantic activation.

Acknowledgements

This research was funded by the Spanish Ministry of Economy and Competitiveness (PSI2015-63525-P) and by the Research Promotion Program of the Universitat Rovira i Virgili (2015PFR-URV-B2-37). The first author has a grant from the Universitat Rovira i Virgili (2015PMF-PIPF-16).

Appendix A

Experimental stimuli

Word	English trans.	Condition	NOM	ROM
ácido	sour/acid	homonym	1.54	2.31
acuario	aquarius/aquarium	homonym	1.58	1.64
baja	fall in price/time off sick	homonym	1.90	1.68
burbuja	bubble	homonym	1.56	1.96
campaña	campaign/countryside	homonym	1.91	1.56
caña	rod/beer	homonym	1.88	1.67
código	code	homonym	1.25	2.28

(*continued*)

Word	English trans.	Condition	NOM	ROM
colonia	colony/cologne	homonym	1.74	2.25
cómoda	chest of drawers/comfortable	homonym	1.76	2.04
copa	cup/crown	homonym	1.96	1.96
ficha	piece/card	homonym	1.70	1.96
físico	physical/physicist	homonym	1.88	2.44
fuente	fountain/source	homonym	1.92	1.33
guion	hyphen/script	homonym	1.79	2.00
heroína	heroine/heroin	homonym	1.61	1.87
jota	J (letter)/jota (Spanish dance)	homonym	1.83	1.32
lima	lime/rasp	homonym	1.81	1.26
manto	cape/blanket	homonym	1.08	1.85
matriz	womb/matrix	homonym	1.65	1.71
mona	female monkey/pretty	homonym	2.00	1.74
monitor	monitor/instructor	homonym	1.62	2.10
notas	mark/memo	homonym	1.90	2.41
palma	palm	homonym	1.92	1.57
partida	departure/round (game)	homonym	1.74	2.17
pasta	pasta/money	homonym	2.00	1.89
patrón	boss/pattern	homonym	1.91	1.95
pensión	pension/hostel	homonym	1.74	2.24
perfil	profile	homonym	2.00	2.37
pipa	pipe/seed	homonym	1.87	1.48
plancha	sheet/iron	homonym	1.90	1.88
planear	plan/glide	homonym	1.90	1.85
recto	straight/rectum	homonym	1.65	1.52
rollo	roll/bore	homonym	1.86	1.72
sirena	siren	homonym	1.78	1.16
tabla	board/table	homonym	1.79	1.84
tanque	tank	homonym	1.71	2.39
tono	tone	homonym	1.63	1.57
vale	OK/voucher	homonym	1.90	1.61
acento	accent/emphasis	polyseme	1.67	3.38
activo	active	polyseme	1.81	5.15
aguja	needle	polyseme	1.39	2.57
armar	arm/assemble	polyseme	1.65	3.05
barra	bar	polyseme	1.81	2.63
bestia	beast/brute	polyseme	1.76	5.47
billete	bill/ticket	polyseme	1.75	3.52
bombón	chocolate/beauty	polyseme	1.43	3.38
brote	sprout/outbreak	polyseme	1.86	4.11
busto	bust	polyseme	1.30	3.91
capa	layer/cape	polyseme	1.87	2.74
cartas	cards/letters	polyseme	1.76	3.30
cólera	anger/cholera	polyseme	1.58	3.44
damas	draughts (game)/ladies	polyseme	1.81	3.00
fracción	part/section/split/fraction	polyseme	1.61	5.48
genio	genie/genius	polyseme	1.52	2.76
globo	balloon/globe	polyseme	1.75	3.65
grano	grain/spot	polyseme	1.92	2.89
letra	letter	polyseme	1.62	4.26
listo	ready/clever	polyseme	1.65	4.32
manual	manual	polyseme	1.74	3.21
marca	mark/brand	polyseme	1.95	2.50
pasajero	passenger/temporary	polyseme	1.50	4.30
pluma	feather/quill	polyseme	2.00	4.17
rango	rank/status	polyseme	1.43	5.39
rosa	rose/pink	polyseme	1.91	3.56
señal	gesture/signal/mark	polyseme	1.85	4.24
solar	solar/site	polyseme	1.83	2.61
sólido	solid/strong	polyseme	1.46	3.59
talla	size/height	polyseme	1.67	6.08
titular	title/principal	polyseme	1.88	3.82
tronco	trunk (tree)/trunk (body)	polyseme	1.65	3.91
virgen	virgin	polyseme	1.91	3.08
vocal	vocal/vowel	polyseme	1.75	4.39
abeja	bee	unambiguous	1.05	
acabar	to finish	unambiguous	1.30	
aceite	oil	unambiguous	1.17	
acero	steel	unambiguous	1.00	
agua	water	unambiguous	1.04	

(*continued on next page*)

(*continued*)

Word	English trans.	Condition	NOM	ROM
alcalde	mayor	unambiguous	1.05	
alma	soul (of a person)	unambiguous	1.15	
almirante	admiral	unambiguous	1.15	
amar	to love	unambiguous	1.10	
bandera	flag (symbol of a country)	unambiguous	1.48	
barranco	ravine	unambiguous	1.09	
baúl	trunk (storage)	unambiguous	1.04	
bayeta	baize	unambiguous	0.83	
biólogo	biologist	unambiguous	1.07	
bruma	mist	unambiguous	0.50	
caballero	gentleman	unambiguous	1.81	
calor	hot (temperature)	unambiguous	1.08	
camioneta	van	unambiguous	1.04	
caos	chaos	unambiguous	1.45	
casta	lineage	unambiguous	1.05	
cerilla	match (stick for lighting fire)	unambiguous	1.04	
cerveza	beer	unambiguous	1.00	
clan	clan	unambiguous	1.40	
coágulo	clot (blood)	unambiguous	1.08	
cofre	chest (box)	unambiguous	1.09	
coleta	pigtail (hairstyle)	unambiguous	1.35	
contusión	contusion	unambiguous	1.09	
cuestionario	questionnaire	unambiguous	1.04	
década	decade	unambiguous	1.09	
domingo	Sunday	unambiguous	1.13	
ecuación	equation (mathematical expression)	unambiguous	1.13	
error	error	unambiguous	1.09	
fe	faith (religious belief)	unambiguous	1.00	
flores	flowers	unambiguous	1.19	
gama	spectrum	unambiguous	1.50	
geología	geology	unambiguous	1.00	
guitarra	guitar	unambiguous	1.04	
hallar	to find	unambiguous	1.10	
hélice	propeller	unambiguous	1.13	
hijo	son	unambiguous	1.04	
himno	anthem	unambiguous	1.00	
hito	milestone	unambiguous	0.80	
humo	smoke	unambiguous	1.22	
ira	anger	unambiguous	1.00	
jabón	soap (bar of soap)	unambiguous	1.00	
jeringa	syringe	unambiguous	1.00	
junio	June	unambiguous	1.04	
labor	labour (work)	unambiguous	1.30	
legado	legacy	unambiguous	1.15	
lencería	lingerie	unambiguous	1.17	
llegar	to arrive	unambiguous	1.15	
lograr	to achieve	unambiguous	1.10	
mar	sea	unambiguous	1.76	
martillo	hammer	unambiguous	1.43	
mente	mind (brain)	unambiguous	1.05	
miel	honey (sweet fluid made by bees)	unambiguous	1.07	
modo	mode (manner)	unambiguous	1.35	
neutrón	neutron	unambiguous	1.00	
optar	to opt	unambiguous	1.10	
pan	bread	unambiguous	1.04	
paraguas	umbrella	unambiguous	1.09	
pensar	to think	unambiguous	1.05	
rato	little while	unambiguous	1.05	
recado	errand	unambiguous	1.20	
riñón	kidney	unambiguous	1.04	
sede	headquarters	unambiguous	1.30	
sobrina	niece	unambiguous	1.00	
tarea	homework	unambiguous	1.25	
teclado	keyboard	unambiguous	1.26	
usar	to use	unambiguous	1.15	
vejez	old age	unambiguous	1.04	
zona	zone	unambiguous	1.10	

Appendix B

Response times regression coefficients for 260 words from the database of Haro et al. (2017), for which lexical decision data were available from González-Nosti et al. (2014). P-values are represented with asterisks. The reported regression coefficients correspond to the variables entered in that particular step.

Predictor		
Step 1: Lexical variables		
Word length	−0.05	
Number of syllables	0.18	*
Log word frequency	−0.49	***
OLD20	0.16	
Number of neighbors	0.00	
Bigram frequency	−0.04	
Trigram frequency	0.07	
Adjusted R^2	0.40	***
Change in R^2	0.40	***
Step 2: Semantic variables		
Number of associates	−0.02	
Log contextual diversity	−0.57	*
Concreteness	−0.14	*
Adjusted R^2	0.41	***
Change in R^2	0.01	*
Step 3: Ambiguity measures		
NOM	−0.14	*
Adjusted R^2	0.42	***
Change in R^2	0.01	*

Note. $^{*}p < 0.05$; $^{***}p < 0.001$.

References

Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia, 49*(5), 970–983.

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 273–278).

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2223–2228).

Armstrong, B. C., & Plaut, D. C. (2016). *Semantic ambiguity effects in lexical processing: A neural-network account based on semantic settling dynamics*. Unpublished manuscript.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Hillsdale, NJ: Erlbaum.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*(3), 235–260.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research, 24*(1), 57–65.

Borowsky, R., & Masson, M. E. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 63–85.

Brown, S. W. (2008). Polysemy and the mental lexicon. *Colorado Research in Linguistics, 21*, 1–12.

Chen, L., Shu, H. U. A., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition, 10*(2), 161–174.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*(4), 1246–1258.

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review, 22*(1), 13–37.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124.

Fraga, I., Padrón, I., Perea, M., & Comesaña, M. (2017). I saw this somewhere else: The Spanish Ambiguous Words (SAW) database. *Lingua, 185*, 1–10.

González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods, 46*(2), 517–525.

Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics, 38*(2), 457–475.

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: When does relatedness matter? *Canadian Journal of Experimental Psychology, 64*(3), 180–196.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1331–1356.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 686–713.

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language, 55*(2), 247–273.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13*(2), 278–305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior, 14*(5), 534–537.

Kawamoto, A. H., Farrar, W. T., & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1233–1247.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 601–609.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627–633.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language, 45*(2), 259–282.
Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: Conceptual relations of polysemous senses. *Journal of Memory and Language, 47*(4), 548–570.
Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20*(1), 1–24.
Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language, 123*(1), 11–21.
Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 804–823.
Kramer, A. F., & Donchin, E. (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 76–86.
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.
Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology, 48*(2), 176–186.
MacGregor, L. J., Bouwsema, J., & Klepousniotou, E. (2015). Sustained meaning activation for polysemous but not homonymous words: Evidence from EEG. *Neuropsychologia, 68*, 126–138.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407.
Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17*(2), 141–147.
Müller, O., Duñabeitia, J. A., & Carreiras, M. (2010). Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology, 47*(3), 455–466.
Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review, 15*(1), 161–167.
Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1252–1270.
Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience, 7*, 723.
Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences, 7*(5), 187–189.
Rabovsky, M., Sommer, W., & Rahman, R. A. (2012). Depth of conceptual knowledge modulates visual processes during word reading. *Journal of Cognitive Neuroscience, 24*(4), 990–1005.
Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*(2), 245–266.
Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28*(1), 89–104.
Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 487–494.
Simon, G., Bernard, C., Lalonde, R., & Rebaï, M. (2006). Orthographic transparency and grapheme–phoneme conversion: An ERP study in Arabic and French readers. *Brain Research, 1104*(1), 141–152.
Taler, V., Kousaie, S., & Lopez Zunini, R. (2013). ERP measures of semantic richness: The case of multiple senses. *Frontiers in Human Neuroscience, 7*, 5.
Tamminen, J., Cleland, A. A., Quinlan, P. T., & Gaskell, M. G. (2006). Processing semantic ambiguity: Different loci for meanings and senses. In *Proceedings of the twenty-eighth annual conference of the cognitive science society* (pp. 2222–2227).
West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024–1037.

2.4. Study 4: Haro, J., Comesaña, M, & Ferré, P. (submitted). Is there an orthographic boost for ambiguous words during their processing?

2.4.1. Description of the study

There are two main differences between the two models that suggest that ambiguous words benefit from triggering a large amount of semantic activation (Balota et al., 1991; Borowsky & Masson, 1996; Hino & Lupker, 1996). The first is that the Borowsky and Masson model assumes that the link between the orthographic and semantic levels is unidirectional, so that the activation at the semantic level has no influence on lower levels. Instead, the semantic feedback hypothesis (Balota et al., 1991; Hino & Lupker, 1996) suggests that the link between the orthographic and semantic levels is bidirectional. This leads to the second difference. According to the semantic feedback hypothesis, the mechanism of the facilitation for ambiguous words is that the semantic activation triggered by ambiguous words influences lower levels, causing the orthographic activation threshold required for word recognition to be reached faster. In contrast, in the Borowsky and Masson model, the advantage is due to the fact that ambiguous words increase global activation at the semantic level. Therefore, a straightforward way to contrast these two models is to test whether ambiguous words benefit from an orthographic boost or not.

To address the above issue, three experiments were performed. In Experiment 1, we verified that the words selected for this experimental series showed the typical ambiguity advantage in a LDT. After that, we examined whether ambiguous words benefit from orthographic boost in two more experiments, in which a Two-Alternative Forced-Choice task (2AFC) was used. In this task, which is considered to directly tap into orthographic processing (e.g., Gomez, Ratcliff, &

Perea, 2008), a word (ambiguous or unambiguous) was briefly presented (i.e., 50 ms) and then it was replaced by two response options corresponding to two different words. In the Experiment 2, these words were the previous presented word and an orthographic neighbor of that word, whereas in the Experiment 3, the words were an orthographic neighbor of the previous presented word and a control word for that neighbor. Participants were asked to select either the word that was presented previously (Experiment 2; e.g., *banco – banco/manco*) or the word orthographically related to that presented before (Experiment 3; e.g., *banco – manco/perro*).

2.4.2. Predictions

In agreement with both models, we expected an ambiguity advantage in the LDT (Experiment 1). However, the predictions for the 2AFC experiments (Experiments 2 and 3) differ between models. According to the semantic feedback hypothesis (Balota et al., 1991; Hino & Lupker, 1996), we expected differences between ambiguous and unambiguous words in the 2AFC tasks. On the contrary, in line with Borowsky and Masson (1996)'s model, no differences should be expected between both types of words in these tasks.

Abstract

The present study explores the issue of why ambiguous words are recognized faster than unambiguous ones during word recognition. To this end we contrasted two different hypotheses: the *semantic feedback* hypothesis (Hino & Lupker, 1996), and the hypothesis proposed by Borowsky and Masson (1996). Although both hypotheses agree that ambiguous words benefit during recognition in that they engage more semantic activation, they disagree as to whether or not this greater semantic activation feeds back to the orthographic level, hence speeding up the orthographic coding of ambiguous words. Participants were presented with ambiguous and unambiguous words in two tasks, a lexical decision task (LDT) and a two-alternative forced-choice task (2AFC). We found differences between ambiguous and unambiguous words in both the LDT and the 2AFC tasks. These results suggest that the orthographic coding of ambiguous words is boosted during word processing. This finding lends support to the semantic feedback hypothesis.

Keywords

Lexical ambiguity; ambiguity advantage; word recognition; orthographic processing; two-alternative forced-choice task

Many studies have shown that ambiguous words (that is, words having more than one meaning, such as *bank*) are recognized faster than unambiguous words (words having only one meaning, like *tennis*) in a lexical decision task (hereafter, LDT; e.g., Borowsky & Masson, 1996; Fraga, Padrón, Perea, & Comesaña, 2017; Haro, Demestre, Boada, & Ferré, 2017; Haro & Ferré, 2018; Hino, Kusunose, & Lupker, 2010; Hino, Lupker, & Pexman, 2002; Hino & Lupker, 1996; Hino, Pexman, & Lupker, 2006; Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Lin & Ahrens, 2010; Millis & Button, 1989; Pexman, Hino, & Lupker, 2004; Rubenstein, Garfield, & Millikan, 1970). Despite such a large body of evidence, the source of the so-called *ambiguity advantage* has not been fully clarified. Some early accounts claimed that the cause of the facilitation for ambiguous words in LDT was that these words are represented by multiple lexical entries, one for each of their meanings. As such, the likelihood of finding a match for an ambiguous word during the scanning of lexical entries is higher than for an unambiguous word (e.g., Forster & Bednall, 1976). More recent accounts, by contrast, have suggested that ambiguous words do not have multiple lexical entries, but rather multiple semantic representations (i.e., Borowsky & Masson, 1996; Hino & Lupker, 1996). Thus, the cause of the facilitation for ambiguous words in LDT would be that these words engage a large amount of semantic activation during processing.

Although an interesting proposition, it is not clear how such enhanced semantic activation might boost ambiguous word recognition. Indeed, two different hypotheses have been suggested. On the one hand, Hino and Lupker's (1996) *semantic feedback* hypothesis relies on principles of interactive activation (e.g., Balota, Ferraro & Connor, 1991; McClelland & Rumelhart, 1981). Within this framework, the visual word processing system consists of at least two linked, bidirectional levels of processing, one devoted to the orthography and the other to the meaning of the word. When the system is presented with a word, activation spreads forward (from the orthographic to the semantic level) and backwards (from the semantic to the orthographic level), and the word is recognized when the activation at the orthographic level reaches a given threshold. Accordingly, the activation at the semantic level modulates the activation at the orthographic level during word processing, so that the more semantic information a word has (e.g., number of meanings), the higher is the impact on its orthographic processing. The ambiguity advantage, then, is because the multiple semantic representations of ambiguous words provide a large amount of semantic feedback for their orthographic representation, leading to ambiguous words reaching the threshold for word recognition faster than unambiguous words.

The alternative hypothesis for the ambiguity advantage was provided by Borowsky and Masson (1996). They developed and tested a Parallel Distributed Processing (hereafter, PDP) model consisting of three levels of processing units, orthographic, phonologic and semantic. The model differs in two significant aspects with respect to that of Hino and Lupker. First, Borowsky and Masson assigned a unidirectional link between orthographic and semantic levels, so that activation can only flow forward (i.e., from the orthographic to the semantic level). Second, they considered that word recognition not only depends on the amount of activation reached at the orthographic level, as Hino and Lupker suggested, but also at the semantic level. Thus, a word is recognized when the summed activation of both orthographic and semantic levels reaches a given value. Despite these restrictions, simulation data from Borowsky and Masson's model clearly replicated the ambiguity advantage, as ambiguous words reached the criterion for word recognition faster than unambiguous words. This was the case because all the different meanings of ambiguous words were partially activated during word processing, eliciting more semantic activation than unambiguous words. However, it is important to note that since the link between orthographic and semantic levels was not bidirectional, this increased semantic activation for ambiguous words had no effect on orthographic processing. Therefore, when ambiguous words reached the criteria for word recognition during the simulations, no differences in the amount of orthographic activation were found between these words and unambiguous words.

In light of the above considerations, it seems clear that the main discrepancy between the two accounts is related to whether or not orthographic processing is boosted during the recognition of ambiguous words. To test this hypothesis, in the present study we compared ambiguous and unambiguous words in a task that taps perceptual aspects of word processing. For this we employed a two-alternative forced-choice paradigm (hereafter, 2AFC). In this task, a target stimulus was presented briefly (e.g., 50 ms), and immediately afterwards the participant was asked to decide which of two strings of letters, (e.g., the flashed word or a lexical neighbor word) was the one previously presented. According to the semantic feedback account (Hino & Lupker, 1996), since the orthographic representation of ambiguous words benefits from a great amount of semantic feedback, we might expect an ambiguity advantage in the 2AFC task with respect to unambiguous words. By contrast, based on Borowsky and Masson (1996)'s model, because the enhanced semantic activation for ambiguous words does not have any influence on orthographic processing, we should not observe an advantage for these words in the 2AFC task. Finally, before conducting the main experiment we verified that our experimental stimuli showed the typical ambiguity advantage in LDT. Thus, the experimental stimuli to be presented in the 2AFC task (Experiment 2) were first tested in a LDT (Experiment 1).

EXPERIMENT 1

Method

Participants

Twenty-two Spanish speakers (18 women and 4 men, mean age = 22 years) participated in the experiment. These were undergraduate students who received academic credits for their participation. All had either normal or corrected-to-normal vision.

Design and materials

Experimental stimuli consisted of 50 Spanish words: 25 ambiguous words and 25 unambiguous words. The ambiguous/unambiguous categorization was based on Number-Of-Meanings (NOM) ratings (c.f., Kellas, et al., 1988; Pexman et al., 2004). The NOM ratings were obtained from the normative study of Haro, Ferré, Boada, and Demestre (2017). To obtain NOM, participants were required to indicate how many meanings a string of letters has, on a 3-point scale: (0) *the word has no meaning*, (1) *the word has one meaning*, or (2) *the word has more than one meaning*. Words with NOM ratings below 1.3 were classified as unambiguous, and words with NOM ratings above 1.4 were classified as ambiguous. This criterion was similar to that used in previous studies (e.g., Hino et al., 2006). The average NOM rating was 1.11 (*SD* = 0.08) for unambiguous words and 1.71 (*SD* = 0.16) for ambiguous words, $t(48) = 17.20$, $p < .001$. In addition, stimuli were matched on several lexical and semantic variables that influence word recognition (see Table 1). Specifically, they were matched in terms of number of letters, number of syllables, logarithm of word frequency (log word frequency), mean Levenshtein distance of the 20 closest words (OLD20), number of neighbors, number of higher frequency neighbors, bigram frequency, trigram frequency, and logarithm of contextual diversity (log contextual diversity) (all *p*s > .13). These values were taken from EsPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). Ambiguous and unambiguous words were also matched in terms of familiarity, concreteness, valence, and subjective age of acquisition (all *p*s > .48). The values for these variables were taken from Haro et al. (2017). Finally, 50 pseudohomophones matched in length to words were included as nonwords in the LDT. All the materials are presented in the Appendix.

Table 1. *Characteristics of the stimulus used in both experiments (standard deviations are shown in parentheses).*

	NOM	FRE	CTD	FAM	AoA	LNG	SYL	CON	VAL	OLD20	NEI	NHF	BFQ	TFQ
Unambiguous	1.11 (0.08)	1.18 (0.60)	0.75 (0.43)	5.51 (0.95)	6.83 (2.63)	6.16 (1.7)	2.60 (0.82)	5.43 (0.86)	4.78 (1.38)	1.87 (0.51)	4.76 (5.19)	0.80 (1.12)	6,100 (4,017)	718.39 (756.48)
Ambiguous	1.71 (0.16)	1.03 (0.47)	0.71 (0.37)	5.55 (0.80)	6.37 (1.89)	6.36 (1.35)	2.52 (0.59)	5.35 (0.68)	5.01 (1.09)	1.68 (0.35)	5.60 (6.80)	0.88 (1.23)	5,117 (3,840)	793.37 (756.85)

Note. NOM = subjective Number-Of-Meanings ratings; FRE = log word frequency; CTD = log contextual diversity; FAM = familiarity; AoA = subjective age-of-acquisition; LNG = word length; SYL = number of syllables; CON = concreteness; VAL = emotional valence; NEI = number of substitution neighbors; NHF = number of higher frequency substitution neighbors; BFQ = mean bigram frequency; TFQ = mean trigram frequency.

Procedure

Participants completed a LDT consisting of 100 experimental trials. Each trial started with a fixation point (i.e., "+") appearing in the middle of the screen for 500 ms. Next, a string of letters (a word or a pseudoword) replaced the fixation point, and then participants had to decide whether the string was or was not a Spanish word. They were instructed to press the "yes" button of a keypad with the preferred hand if the string of letters was a word, and to press the "no" button of the keypad with the non-preferred hand if it was not a word. The string of letters remained on the screen until participant's response or timeout (2000 ms). After responding, a feedback message (i.e., "ERROR" or "CORRECT") was displayed for 750 ms. The interval time between trials was 500 ms. We used DMDX software (Forster & Forster, 2003) to present the stimuli and to record the responses. The order of the experimental trials was randomized for each participant. Prior to the beginning of the experiment, a practice block consisting of 10 trials (5 words and 5 nonwords) was presented.

Results and Discussion

RTs that exceeded 2 SD of each participant's mean were rejected (4.9%). The mean of reaction times (RT) for correct responses and the mean of error rates (%E) across experimental conditions (averaged across participants) are shown in Table 2.

Table 2. *Mean RT (in ms), and percentage of error rates (%E) in Experiment 1 per experimental condition (standard deviations in parentheses)*

Ambiguity	Mean RT	%E
Unambiguous	628 (113)	5.19 (4.29)
Ambiguous	591 (102)	2.48 (3.62)
Pseudowords	672 (119)	8.46 (6.77)

The results showed that ambiguous words were faster and more accurately recognized than unambiguous words, $t_1(21) = 5.81$, $p < .001$, $t_2(48) = 2.41$, $p = .02$, $t_1(21) = 2.37$, $p = .028$, $t_2(48) = 1.88$, $p = .067$, for latency and error data respectively. Therefore, the selected stimuli produced a robust ambiguity advantage, resembling that observed in previous studies (e.g., Haro et al., 2017; Haro & Ferré, 2018; Hino et al., 2002; Hino & Lupker, 1996; Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Kellas et al., 1988; Lin & Ahrens, 2010; Millis & Button, 1989; Pexman et al., 2004; Rubenstein et al., 1970). The stimuli were thus suitable to be tested in the 2AFC task, which was the task used in the Experiment 2 to assess the two theoretical accounts of the ambiguity advantage mentioned above.

EXPERIMENT 2

Method

Participants

Thirty-one Spanish speakers (22 women and 9 men, mean age = 22 years) from the same population as those in the first experiment carried out the task. They were undergraduate students who received academic credits for their participation, and all of them had either normal or corrected-to-normal vision.

Design and materials

Experimental stimuli comprised 50 pairs of words, each pair consisting of a word from Experiment 1 and a lexical neighbor differing in one or two letters. For example, the unambiguous word *techo* (“roof”) was paired with its neighbor *pecho* (“chest”), and the ambiguous word *fuente* (“fountain” or “source”) was paired with its neighbor *puente* (“bridge”). Thus, there were two conditions: one formed by 25 pairs of words containing an unambiguous word, and the other formed by 25 pairs of words containing an ambiguous word. Experimental conditions were matched for a large number of variables (all *p*s > .28; see Table 3). First, conditions were matched for the Levensthein distance, and number of different letters between the target and its neighbour. Levensthein distance and orthographic similarity were computed using NIM (Guasch, Boada, Ferré, & Sánchez-Casas, 2013). Furthermore, since deviant letter position (i.e., the position occupied by the letter that varies between the target and the neighbor) can influence word recognition (see Comesaña, Coelho, Oliveira, & Soares, 2017, for more detail), this variable was matched between conditions. There was a similar number of pairs between conditions having a deviant letter in the first, middle, last and other positions. Finally, the lexical neighbor of each pair was matched between conditions in log word frequency, number of letters and syllables, number of neighbors, number of higher frequency neighbors, OLD20, and trigram and bigram frequency. All these variables were obtained from EsPal (Duchon et al., 2013).

Table 3. *Characteristics of the pairs of stimulus used in the 2AFC task (standard deviations are shown in parentheses).*

	LD	OS	DIFF	FRE	LNG	SYL	OLD20	NEI	NHF	BFQ	TFQ
Unambiguous	0.75	0.70	1.52	1.01	6.16	2.52	1.71	5.52	1.04	7,368	1,117
	(0.08)	(0.11)	(0.51)	(0.55)	(1.70)	(0.77)	(0.43)	(5.92)	(1.43)	(5,878)	(2,373)
Ambiguous	0.75	0.71	1.52	1.11	6.36	2.56	1.62	5.52	1.16	5,805	1,190
	(0.08)	(0.11)	(0.51)	(0.67)	(1.35)	(0.58)	(0.28)	(5.99)	(2.10)	(4,002)	(1,140)

Note. LD = Levensthein distance between the target and its neighbor; OS = orthographic similarity between the target and its neighbor; DIFF = number of different letters between the target and its neighbour; FRE = log word frequency of the neighbour; LNG = word length of the neighbour; SYL = number of syllables of the neighbour; NEI = number of substitution neighbors of the neighbour; NHF = number of higher frequency substitution neighbors of the neighbour; BFQ = mean bigram frequency of the neighbour; TFQ = mean trigram frequency of the neighbour.

Procedure

The stimuli were presented using a 2AFC paradigm. The sequence of each trial was as follows. First, a fixation point (“+”) was displayed for 500 ms in the center of the screen. Then, a word (i.e., an ambiguous or unambiguous word) was presented for 50 ms, and was then immediately masked with segments of letters. When the mask appeared, two lowercase words were displayed below it, one on each side. These words were the flashed ambiguous or unambiguous word and its lexical neighbor (e.g., *cerveza*-*certeza*). Then, participants were asked to decide which of the two words was the flashed one. Participants had to press the right button of a keypad if the flashed word was the one located on the right, and left button if it was the one located on the left. The next trial started automatically after response or timeout (3000 ms). There were two different versions of the experiment to counterbalance the position of the target (i.e., left or right) across participants. Participants were presented with 10 practice trials and 50 experimental trials. The order of the experimental trials was randomized for each participant.

Results and Discussion

Following the usual procedure for analyzing the 2AFC data, we calculated the mean %E across experimental conditions (see Table 4).

Table 4. *Percentage of error rates (%E) in Experiment 2 (standard deviations in parentheses)*

Ambiguity	%E
Unambiguous	16.39 (11.88)
Ambiguous	12.52 (9.84)

The results showed that ambiguous words were identified more accurately than unambiguous words, $t_1(30) = 2.27$, $p = .031$, although the effect was marginal in the analysis by items, $t_2(98) = 1.67$, $p = .097$.

The advantage for ambiguous words in the 2AFC task suggests that orthographic processing is boosted during ambiguous word processing, in accordance with the semantic feedback account (Hino & Lupker, 1996). However, there is the possibility that the ambiguity advantage observed in the 2AFC task was not caused exclusively by an orthographic boost for ambiguous words. Indeed, although this task is considered to tap perceptual aspects of word processing (e.g., Gomez, Ratcliff, & Perea, 2008), there is also evidence showing that 2AFC responses are somewhat influenced by semantic processing (e.g., Bell, Forster, & Drake, 2015; Marcel, 1983). For example, in the study of Marcel (1983), participants conducted a 2AFC task in which the flashed word and one of the two targets were related semantically (e.g., *dog* - wallet/animal). Participants had to indicate which of them was semantically related to the preceding flashed word. The results showed that although participants did not consciously perceive the flashed word, they were able to select the correct option above chance.

Taking into account the above, one could argue that the results of Experiment 2 do not strongly prove that ambiguous words benefit from an orthographic boost. For this reason, we designed a new 2AFC experiment in which targets and flashed words were only orthographically related. In this experiment, the targets were a lexical neighbor of the flashed word and a control of that neighbor. Participants were asked to decide which of the two targets was orthographically related with the previously flashed word.

EXPERIMENT 3

Method

Participants

Forty-one Spanish speakers (35 women and 6 men, mean age = 21 years) from the same population as those of the previous experiments participated in this experiment. They were undergraduate students who received academic credits for their participation, and all of them had either normal or corrected-to-normal vision.

Design and materials

The experimental stimuli were the same as those employed in Experiment 2, that is, 25 ambiguous and 25 unambiguous words, as well as a lexical neighbor for each critical word (i.e., 50 words). In addition, we selected 50 control words for the lexical neighbors. They were pairwise matched with the lexical neighbors in log frequency, number of letters, number of neighbors, number of higher frequency neighbors, and OLD20 (all *p*s > .32). The values for these variables were obtained from EsPal (Duchon et al., 2013).

Procedure

The procedure of the 2AFC was similar to that employed in Experiment 2, but with some changes that are detailed as follows. Unlike Experiment 2, the words presented after the unambiguous or ambiguous flashed word (e.g. *faro*) were its lexical neighbor (e.g., *foro*) and a control for that neighbor (e.g., *lona*). In addition, participants were asked to decide which of the two words was orthographically related with the flashed word. Finally, all 50 trials were presented three times (in three different randomized blocks) to each participant.

Results and Discussion

As in Experiment 2, we calculated the mean %E across experimental conditions (see Table 5).

Table 5. *Percentage of error rates (%E) in Experiment 3 (SD in parentheses)*

Ambiguity	%E
Unambiguous	6.57 (7.80)
Ambiguous	8.62 (7.98)

The results showed that lexical neighbors preceded by ambiguous words were identified less accurately than those preceded by unambiguous words in the analysis by participants, $F_1(1,40) = 14.88$, $p < .001$, although the effect did not reach significance in the analysis by items, $F_2(1,48) = 1.97$, $p = .17$.

Hence, ambiguous words caused an inhibitory effect in this experiment. At a first glance, this result might seem to contradict the facilitation effect found for ambiguous words in Experiments 1 and 2. However, this is not the case: The inhibition effect found here is similar to that observed in other studies that employ a masked form priming procedure, where participants are required to respond to a target word preceded by an orthographically related subliminal word. Using this procedure, some studies reported that target words are recognized slower and less accurately when they are preceded by a lexical neighbor in the LDT (e.g., De Moor & Brysbaert, 2000; Segui & Grainger, 1990). The explanation for this effect, according to the interactive activation model (McClelland & Rumelhart, 1981), is that the orthographic representation of the neighbor, that was presented as a prime, is strongly activated while participants try to recognize the target. Consequently, the activation of the neighbor interferes with the recognition of the target word, resulting in slower reaction times and more errors for these words in the LDT. Taking this into account, the inhibition found in the 2AFC can be explained in a similar way. Assuming that semantic-to-orthographic feedback is larger for ambiguous words than for unambiguous words, the orthographic representation of an ambiguous word would be more active after its presentation than that of an unambiguous word. As such, when participants were required to decide which of the two displayed words (i.e., a lexical neighbor of the flashed word or a control of that neighbor) was orthographically related to the one presented before, more interference would be expected when the flashed word was an ambiguous word than when it was an unambiguous word. Thus, the results of this experiment provide further support to the hypothesis that ambiguous words benefit from an orthographic boost (Hino & Lupker, 1996).

General Discussion

The aim of the present study was to investigate the source of the ambiguity advantage, that is, the reason why ambiguous words are recognized faster than unambiguous words during word recognition. We contrasted two hypotheses here: i) the semantic feedback hypothesis (Hino & Lupker, 1996), and ii) the hypothesis developed by Borowsky and Masson (1996). Although both agree that the facilitation for ambiguous words is due to the fact that these words elicit more semantic activation than unambiguous ones, they differ in whether such enhanced semantic activation boosts orthographic coding (Hino & Lupker, 1996) or not (Borowsky & Masson, 1996). To examine this question, we analyzed the processing of ambiguous and unambiguous words using a task that taps perceptual aspects of word processing (i.e., the 2AFC task). A LDT was also used to verify that the typical ambiguity advantage reported in previous LDT studies (e.g., Hino et al., 2002; Lin & Ahrens, 2010; Rubenstein et al., 1970) was also observed here.

The results showed a facilitation of ambiguous words in the LDT as well as differences between ambiguous and unambiguous words in the 2AFC tasks. Therefore, the results of the 2AFC tasks are incompatible with the PDP model of

Borowsky and Masson (1996). This model assumes that the link between the orthographic and the semantic level is unidirectional, so that semantic-to-orthographic feedback is not allowed, and thus no differences should be expected between ambiguous and unambiguous words in tasks that tap perceptual aspects of word processing. In contrast, the evidence obtained in the 2AFC tasks suggests that orthographic processing is boosted during ambiguous word processing, thus giving support to the semantic feedback account (Hino & Lupker, 1996). This account is based on interaction activation principles, in which activation flows bidirectionally between orthographic and semantic levels after presenting the input word (e.g., Balota et al., 1991; McClelland & Rumelhart, 1981). Hence, because ambiguous words have multiple semantic representations, their orthographic representation would receive a large amount of semantic feedback during word processing. This would eventually speed up the orthographic coding of these words, allowing them to reach the orthographic activation criteria for word recognition faster. Of note, the same semantic feedback mechanism was proposed to accommodate other effects found in word recognition research. For instance, it could explain why words with many synonyms are recognized slower in LDT than those with few synonyms (Hino et al., 2002; Pecher, 2001); that is, such a synonymy effect would be due to a single semantic representation spreading its activation to multiple orthographic representations (i.e., one for each synonym), thus increasing competition at the orthographic level. Hence, although the effects are inhibitory rather than facilitative in this case, they would also be produced by the activation of orthographic units as a consequence of semantic activation. In sum, the present study suggests that ambiguous words benefit from a boost in their orthographic coding during word processing, and this would explain why such words are usually recognized faster than unambiguous words in LDT.

ACKNOWLEDGMENTS:

This research was funded by the Spanish Ministry of Economy and Competitiveness (PSI2015-63525-P) and by the Research Promotion Program of the Universitat Rovira i Virgili (2016PFR-URV-B2-37). This has also been partially supported by the FCT (Foundation for Science and Technology) through the state budget with reference IF / 00784/2013 / CP1158 / CT0013. The first author also holds a grant from the Universitat Rovira i Virgili (2015PMF-PIPF-16).

References

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187-222). Hillsdale, NJ: Erlbaum.

Bell, D., Forster, K., & Drake, S. (2015). Early semantic activation in a semantic categorization task with masked primes: Cascaded or not?. *Journal of Memory and Language, 85*, 1-14.

Buland, O., Casalis, S., & Comesaña, M. (2017, 5-8 April). Cross-language transposed-letter effects during native and non-native reading. Poster presented at the XIII International Symposium of Psycholinguistics, Braga, Portugal.

Borowsky, R., & Masson, M. E. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 63-85. doi: 10.1037/0278-7393.22.1.63

Comesaña, M., Coelho, R., Oliveira, H., & Paula Soares, A. (in press). How letter order is encoded in bilingual reading? The role of deviant-letter position in cognate word recognition. *Speech, Language and Hearing*. doi: 10.1080/2050571X.2017.1369049

De Moor, W., & Brysbaert, M. (2000). Neighborhood-frequency effects when primes and targets are of different lengths. *Psychological Research*, *63*(2), 159-162.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246-1258. doi: 10.3758/s13428-013-0326-1

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition, 4,* 53-61. doi: 10.3758/BF03213255

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116-124. doi: 10.3758/BF03195503

Fraga, I., Padron, I., Perea, M., & Comesaña, M. (2017). I saw this somewhere else: The Spanish Ambiguous Words (SAW) database. *Lingua, 185,* 1-10. doi: 10.1016/j.lingua.2016.07.002

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological review*, *115*, 577-600. doi: 10.1037/a0012667

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior research methods*, *45,* 765-771. doi: 10.3758/s13428-012-0296-8

Haro, J., Demestre, J., Boada, R., & Ferré, P. (2017). ERP and behavioral effects of semantic ambiguity in a lexical decision task. *Journal of Neurolinguistics*, *44*, 190-202. doi: 10.1016/j.jneuroling.2017.06.001

Haro, J., & Ferré, P. (2018). Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition? *Journal of Psycholinguistic Research*, *47*, 679-698. doi: 10.1007/s10936-017-9554-3

Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics*, *38*, 457-475. doi: 10.1017/S0142716416000266

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: When does relatedness matter? *Canadian Journal of Experimental Psychology*, *64*, 180-196. doi: 10.1037/a0020475

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1331-1356. doi: 10.1037/0096-1523.22.6.1331

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 686-713. doi: 10.1037/0278-7393.28.4.686

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language*, *55*, 247-273. doi: 10.1016/j.jml.2006.04.001

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*, 278-305. doi: 10.1016/0010-0285(81)90011-6

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior*, *14*, 534-537. doi: 10.1016/S0022-5371(75)80030-2

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 601-609. doi: 10.1037/0096-1523.14.4.601

Lin, C-J. C., & Ahrens, K. (2010). Ambiguity advantage revisited: Two meanings are better than one when accessing Chinese nouns. *Journal of psycholinguistic research*, *39*, 1-19. doi: 10.1007/s10936-009-9120-8

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive psychology, 15*(2), 197-237.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*, 375–407. doi:10.1037/0033-295X.88.5.375

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition*, *17*, 141-147. doi: 10.3758/BF03197064

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1252-1270. doi: 10.1037/0278-7393.30.6.1252

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*, 487-494. doi: 10.1016/S0022-5371(70)80091-3

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 65.

APPENDIX

Experimental stimuli

Condition	Target	Target trans.	Neighbor	Neighbor trans.	Pseudohomophone	Pseudohomophone original word
unambiguous	abeja	bee	queja	complaint	arina	harina
unambiguous	aceite	oil	agente	agent / officer	bajina	vagina
unambiguous	alcalde	mayor	alcance	range /significance	tunva	tumbas
unambiguous	almirante	admiral	aspirante	candidate	bervo	verbo
unambiguous	cal	lime (calcium oxide)	col	cabbage	ardiya	ardilla
unambiguous	calor	hot	color	colour	gayo	gallo
unambiguous	camión	truck	cartón	cardboard	fayo	fallos
unambiguous	cerveza	beer	certeza	certainty	monio	moño
unambiguous	cirugía	surgery	ciruela	plum	omosexual	homosexual
unambiguous	contusión	bruise	confesión	confession	rodiya	rodilla
unambiguous	cueva	cave	curva	curve	llate	yate
unambiguous	demencia	dementia	decencia	decency	raia	raya

Condition	Target	Target trans.	Neighbor	Neighbor trans.	Pseudohomophone	Pseudohomophone original word
unambiguous	ecuación	equation	erupción	eruption	berso	verso
unambiguous	electrón	electron	elección	choice	jobentut	joventut
unambiguous	enzima	enzyme	encina	holm oak	beneno	veneno
unambiguous	hijo	son	hilo	thread	hoso	oso
unambiguous	humo	smoke	zumo	juice	urvano	urbano
unambiguous	jabón	soap	jamón	ham	vaía	bahía
unambiguous	lealtad	loyalty	fealdad	ugliness	poyo	pollo
unambiguous	lencería	lingerie	mercería	haberdashery	anvición	ambición
unambiguous	miel	honey	piel	skin	havuso	abuso
unambiguous	modestia	modesty	molestia	bother	amariya	amarilla
unambiguous	techo	ceiling	pecho	chest	idrójeno	hidrógeno
unambiguous	tenis	tennis	tesis	thesis	rovo	robo
unambiguous	vejez	old age	veloz	quick	orario	horario
ambiguous	activo	active / assets	altivo	arrogant	bisual	visual
ambiguous	acuario	Aquarius / aquarium	armario	cupboard	elado	helado
ambiguous	asistir	help / assist / attend	existir	to exist	varvilla	barbilla

Condition	Target	Target trans.	Neighbor	Neighbor trans.	Pseudohomophone	Pseudohomophone original word
ambiguous	botones	buttons / bellboy	balones	balls	viología	biología
ambiguous	canario	canary / Canarian (demonym of Canary Islands)	catarro	cold (illness)	abenida	avenida
ambiguous	churro	fritter / mess	charco	puddle	bela	vela
ambiguous	cómoda	comfortable / chest of drawers	comida	food	vurguesía	burguesía
ambiguous	complejo	complex	completo	full	enbra	hembras
ambiguous	faro	lighthouse / headlamp	foro	forum	vevé	bebé
ambiguous	ficha	piece / ticket	fecha	date (time)	corvata	corbata
ambiguous	fracción	part / fraction	tracción	traction	alva	alba
ambiguous	fuente	source / fountain	puente	bridge	envra	hembra
ambiguous	golpe	hit / robbery	gripe	flu	imbasión	invasión
ambiguous	herencia	legacy / heredity	carencia	lack	novle	noble
ambiguous	lima	lime (tool) / rasp	liga	league / garter	erida	herida
ambiguous	navaja	knife / razor shell	baraja	deck of cards	yabe	llave
ambiguous	pasador	bolt (security) / hairclip	paladar	palate / taste	ierva	hierba
ambiguous	pensión	pension / hostel	presión	pressure	bibienda	vivienda

Condition	Target	Target trans.	Neighbor	Neighbor trans.	Pseudohomophone	Pseudohomophone original word
ambiguous	plato	plate	plazo	period / deadline	onbro	hombro
ambiguous	postal	postal / postcard	portal	hallway / website	beículo	vehículo
ambiguous	ratón	mouse	razón	reason / reasoning	biernes	viernes
ambiguous	resolución	resolution / decision	revolución	revolution	notavle	notable
ambiguous	segundo	second	seguido	followed	voteya	botella
ambiguous	tanque	tank	parque	park	dever	deber
ambiguous	tronco	trunk / mate	trofeo	trophy	jenética	genética

3. General discussion

This doctoral thesis was developed to address some unresolved questions about how ambiguous words are processed in isolation. The starting point was an empirical review of the studies published since 1970 on ambiguous word recognition. From this review, I found that many studies reported a facilitation for ambiguous words with respect to unambiguous ones in LDT (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Ferraro & Hansen, 2002; Hino et al., 2002, 2006, 2010; Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Lin & Ahrens, 2010; Millis & Button, 1989; Pexman et al., 2004; Piercey & Joordens, 2000; Rubenstein et al., 1970, 1971). The explanation given for this so-called *ambiguity advantage* was that ambiguous words benefit during lexical processing from having as many lexical entries or semantic representations as meanings.

However, over the decades, several criticisms have called into question the existence of such an advantage. Although most of these criticisms were overcome by later evidence, such as the possible confusion between familiarity and the number of meanings (Gernsbacher, 1984), the one raised by Rodd et al. (2002) is still relevant today. In particular, Rodd et al. (2002) showed that number of meanings (NOM) does not facilitate word recognition, but rather makes it more difficult. In addition, they found that the relatedness of meanings (ROM) has a facilitative effect on ambiguous word recognition: the higher the ROM, the faster the recognition. Based on these findings, Rodd et al. suggested that, instead of an ambiguity advantage, there would be an ambiguity disadvantage (inhibition for ambiguous words with multiple meanings) along with a sense advantage (facilitation for words with multiple related senses). Although further studies have supported the findings of Rodd et al. (Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Tamminen et al., 2006), there is also evidence showing a similar facilitation for ambiguous words differing in their ROM with respect to unambiguous words (e.g., Hino et al., 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004), suggesting that NOM, but not ROM, influences word recognition.

In light of the above, the present thesis was focused on the three following main objectives: 1) to assess whether the approach to categorize ambiguous and unambiguous words has any influence on ambiguous word recognition; 2) to examine the processing of ambiguous words that differ in their meaning relatedness by recording both behavioural and neurophysiological measures, and 3) to test the predictions of word recognition models that account for the ambiguity advantage.

The starting point to address these objectives was to gather different objective and subjective measures of semantic ambiguity, as well as some other lexical and semantic measures, for a large set of words. Thus, in Study 1 we developed a database of ambiguous and unambiguous Spanish words made up of 530 words. This database included several NOM and ROM measures, both subjective and objective (i.e., based on dictionary meanings), and it also provided ratings for several lexical and semantic variables. In this way, the database overcomes some of the weaknesses of the existing databases of ambiguous Spanish words, that is, the limited number of words and the scarcity of semantic ambiguity variables (Domínguez et al., 2001; Estevez, 1991; Fraga et al., 2017; Gómez-Veiga et al., 2010).

The database developed in the Study 1 allowed us to address the next objective, that is, to assess whether the approach to categorize ambiguous and unambiguous words has any influence on ambiguous word recognition. My initial hypothesis was that some methodological differences could explain why an ambiguity disadvantage has been observed in some studies, whereas an ambiguity advantage has been found in others. This hypothesis was based on the fact that all the studies that reported a disadvantage for ambiguous words used dictionary definitions to estimate NOM (e.g., Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006), while those that observed an ambiguity advantage employed subjective NOM ratings (e.g., Hino et al., 2002, 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004). Given the particularities of each approach (see the introduction for a detailed explanation), it is plausible to expect that using one approach or the other may modulate the effect of NOM in LDT. In brief, the most remarkably differences between both approaches are that dictionaries list many word meanings that are unknown to speakers (as they include meanings that are outdated or belong to jargon), and that they do not usually list the new meanings that speakers have incorporated into their daily use of language.

The objective described above was addressed in a series of LDT experiments included in the Study 2. On the one hand, we found that ambiguous words were recognized slower and with more errors than unambiguous words when unambiguous/ambiguous words were categorized using dictionary NOM (Experiment 1). This result was in line with previous studies that employed a dictionary approach, since all of them reported an ambiguity disadvantage (e.g., Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006). In contrast, we observed a facilitation for ambiguous words when the unambiguous/ambiguous categorization was made according to subjective NOM (Experiments 2 and 3). This facilitation was similar to that found in previous studies that used subjective NOM approaches (e.g., Hino et al., 2002, 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004). Importantly, this ambiguity advantage was also observed in two additional experiments included in the present thesis (i.e., in the

Experiment reported in Study 3 and in the Experiment 1 of the Study 4). In all of them, subjective ratings were used for NOM estimation.

The findings of the Study 2 suggest that the approach employed to estimate NOM has an influence in ambiguous word recognition. The cause of these opposite results when using subjective or dictionary NOM measures may be that there is not a strong correspondence between the words that have one/many entries in the dictionary and the words that have one/multiple meanings for the speakers. It should be noted that this possibility was supported by the post-hoc analyses of the Experiment 1 in the Study 2, where we found that the same set words could be categorized as ambiguous or unambiguous depending on the criterion chosen. Nevertheless, the findings from Study 2 do not allow to conclude that the misdistribution between subjective NOM and dictionary NOM is the cause of the ambiguity disadvantage found in previous studies, since all of them were conducted in English. In any case, I would like to highlight the significant similarity between the results of the Experiment 1 of the Study 2 and those reported in previous studies that found an ambiguity disadvantage (Armstrong & Plaut, 2008, 2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006). In such experiments, as well as in Experiment 1, words with many dictionary entries and few senses showed the slowest RTs (although only numerically), suggesting that the ambiguity disadvantage usually reported when using dictionary NOM was mainly produced by these words. In addition, in the Experiment 1 we found that the misdistribution between subjective NOM and dictionary NOM was larger in words within that condition (i.e., words with many meanings and few senses) than in the rest of conditions. Thus, this misdistribution could also have occurred in past experiments, leading to a larger inhibition for words with many meanings and few senses with respect to the rest of conditions, and so causing a disadvantage for ambiguous words.

Since using different approaches for NOM estimation is associated with distinct and opposite experimental results in LDT, we should briefly discuss which approach is psychologically more valid in order to provide a proper explanation of how NOM affects word recognition. Gernsbacher (1983) was the first author who questioned the validity of number of dictionary definitions as a measure of NOM: "How psychologically valid is the dictionary count definition of polysemy? Consider, as illustration, the words, gauge, cadet, and fudge. These three words were considered highly familiar by an average of more than 65% of the undergraduate raters (Gernsbacher, 1983). Yet in reality, how many of these subjects are likely to have stored in memory all 30 dictionary meanings of the word gauge, all 15 dictionary meanings of the word cadet, or even all 15 dictionary meanings of the word fudge?" (pages 20 and 21). Of note, the objection raised by Gernsbacher (1983) obtained further support. For instance, Ferraro and Kellas (1990) showed that the correlation between the number of word meanings known by participants and the number of dictionary definitions was only of 0.12. In addition, Lin and Ahrens (2005) compared the number of definitions provided by several dictionaries for the same set words, observing that such number differed significantly between dictionaries. Importantly, this difference was found for both English and Chinese words, which suggests that this issue is not produced by the characteristics of a particular language, but to the way in which dictionaries are made. In sum, it seems that there is a clear disparity between the definitions listed for a given word in the dictionary and the meanings that a speaker knows for that word. This indicates that the number of dictionary definitions is not a psychologically valid measure of NOM and, thus, subjective NOM should be the preferred measure to estimate how many meanings a word has. Consequently, the data obtained in the present thesis (i.e., Study 2, the behavioural results of the Study 3 and the LDT data of the Study 4) provide further support for the facilitating effect of NOM in word recognition.

After having established the relevance of the criterion to define ambiguous words in the pattern of results obtained, in the Study 3 we examined whether ROM has any influence on ambiguous word recognition. We selected unambiguous words and ambiguous words differing in their ROM. Word selection was made according to subjective ratings of NOM and ROM. The results showed a similar facilitation in LDT for ambiguous words with unrelated meanings (homonyms) and ambiguous words with related meanings (polysemes) with respect to unambiguous words. This null effect of the degree of semantic relatedness between ambiguous word meanings is consistent with previous studies that employed subjective ROM (e.g., Hino et al., 2006, 2010; Lin & Ahrens, 2010; Pexman et al., 2004). In addition, no differences were observed in the Study 3 between homonyms and polysemes in the amplitude of the N400 component, which is an ERP component associated with semantic processing (see Kutas & Federmeier, 2011 for a review). These ERP results are of great relevance, since as far as I know, this is the first time that EEG correlates of the processing of polysemes and homonyms in isolation have been examined. Thus, this piece of evidence supports the absence of ROM effects observed in the behavioral data and, at the same time, suggests that there are no neurophysiological differences between polysemes and homonyms in terms of their processing.

The results above described are at odds, however, with those obtained in studies that used dictionary measures for ROM estimation, since they found that high ROM facilitates recognition, whereas low ROM inhibits it (Armstrong & Plaut, 2008, 2011; Klepousniotou & Baum, 2007; Rodd et al., 2002; Tamminen et al., 2006). Again, these conflicting results regarding ROM suggest that the approach used to measure semantic ambiguity influences ambiguous word recognition. Thus, if we assume that subjective measures of semantic ambiguity are psychologically more valid than dictionary ones, the results found in the Study 3 indicate that ROM does not affect ambiguous word recognition. This finding is at odds with the model of Rodd et al. (2004), as it predicts an advantage for polysemes and a disadvantage for homonyms in LDT. This prediction is grounded on the assumption that polysemes and homonyms are represented differently. According to the model, all the related senses of a polyseme are contained within the same single shared representation, forming a broad and rich semantic representation. In contrast, the unrelated meanings of a homonym are represented in separate and distant semantic representations, which compete during word processing. However, the lack of a ROM effect in Study 3 suggests that polysemes and homonyms might be represented in a similar vein. This possibility was first raised by Klein and Murphy (2001). These authors examined the processing of polysemes and homonyms using a sensicality judgment task, a task in which participants were asked to decide if a phrase made sense or not. Polysemes and homonyms were embedded in noun phrases and presented twice to the participants. For example, the polyseme *paper* (i.e., a *sheet of material* and *a publication printed on such material*) appeared the first time embedded in the phrase *wrapping PAPER*, and then it was presented the second time in a consistent-sense phrase (e.g., *shredded PAPER*) or in an inconsistent-sense phrase (e.g., *daily PAPER*). Decision times were compared between inconsistent- and consistent-sense conditions, observing a sense consistency effect for both polysemes and homonyms; that is, decision times were faster when the second appearance of the polyseme or homonym was congruent with the sense of the first presentation (e.g., *shredded-PAPER* after *wrapping-PAPER*) than when it was incongruent with such sense (e.g., *daily-PAPER* after *wrapping-PAPER*). More importantly, the consistency effect was similar for polysemes and homonyms. The authors interpreted these findings as evidence in favour of a separated entries account, that is, the hypothesis that polyseme senses are stored separately, in a similar fashion as homonym meanings. Accordingly, both related senses and unrelated meanings would be represented in separate entries and would contribute similarly to the recognition of ambiguous words.

There is also the possibility, however, of explaining the null effect of ROM without assuming that polysemes and homonyms are represented in the same way. This explanation is based on the evidence that the semantic richness of a word influences its recognition. Semantic richness is defined as the quantity and quality of semantic information associated with a word. Some of the semantic variables included in this dimension are the number of semantic features (Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012), the number of semantic neighbors or the density of semantic neighborhoods (Yap et al., 2012), and the strength of visual associations of a word (e.g., Hargreaves & Pexman, 2012). Research in this field has shown that words with a high semantic richness are more easily recognized than those with a low semantic richness (e.g., Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008). Taking into account the above, one could argue that both the related senses and the unrelated meanings of a word contribute to its semantic richness, as they reflect the amount of semantic information the word has. So that, a word with either multiple senses or meanings would have a richer representation than a word with few senses or with a single meaning. Therefore, even if polysemes and homonyms were represented differently, that is, one representation for polysemes and multiple for homonyms, polysemes would benefit during word recognition from having a rich semantic representation containing all their senses, in a similar way as homonyms benefit from having multiple representations.

So far, the results of the Studies 2 and 3 suggest that NOM, regardless of ROM, facilitates word recognition. On the one hand, such a facilitation for NOM is at odds with Rodd et al. (2004)'s model, since it predicts that words with multiple meanings would be recognized slower than words with one meaning. However, in my opinion, the effect could be explained by this model if we take into account the following. According to Rodd et al.'s model, in the earlier stages of ambiguous word processing, a state containing semantic information from all the meanings of the word is quickly reached (i.e., a *blend state*). Reaching this state is not supposed to be enough to recognize a word in LDT, so that the network should escape from it and reach a semantic representation corresponding to one of the ambiguous word meanings. Given that escaping from this state implies a processing cost for the network, and that meanings compete during word processing, the model predicts a NOM disadvantage. However, some PDP simulations by Joordens and Besner (1994) and Borowsky and Masson (1996) demonstrated that a blend state like that one may be sufficient to recognize the word in LDT, as it seems to provide a strong familiarity cue to discriminate between words and nonwords. Thus, considering that the network reaches faster the blend state of an ambiguous word than the single semantic representation of an unambiguous word (Joordens & Besner, 1994), the model of Rodd et al. could also account for the ambiguity advantage. It should be noted, however, that this explanation would be limited to tasks in which it is not necessary to specify the meaning of the word (e.g., LDT), so that the information contained in the blend state would be sufficient to respond. Therefore, no ambiguity advantage is expected in more semantic engaging tasks, which require that a particular meaning of the word is accessed. This has been supported by the fact that ambiguous words show slower or similar RTs in semantic categorization tasks with respect to unambiguous words (Hino et al., 2002).

The NOM advantage, apart from the possibility of being explained by the model of Rodd et al. if we take the above into account, is compatible with the model of Kawamoto et al. (1994), the model of Borowsky & Masson (1996) and the semantic feedback account (Balota et al., 1991; Hino & Lupker, 1996). Although all these accounts assume that the cause of the facilitation for ambiguous words is that they have multiple semantic representations, they differ in the mechanism through which this multiplicity of meanings facilitates their recognition. For instance, under the PDP model of Kawamoto et al. (1994), the one-to-many inconsistency between orthography and semantics for ambiguous words would cause that these words develop strong orthographic links. This particularity of ambiguous words would speed up their orthographic settling, facilitating thus their recognition. In contrast, Borowsky and Masson (1996)'s model and the semantic feedback account (Balota et al., 1991; Hino & Lupker, 1996) suggest that ambiguous words benefit from triggering a large amount of semantic activation (i.e., semantic enhanced activation accounts). Considering these two explanations of the NOM advantage, we tried to determine the mechanism by which NOM facilitates recognition. To do so, we examined the EEG correlates of ambiguous and unambiguous word processing during a LDT. The results failed to show a NOM effect regarding the N200 (an ERP component associated with orthographic processing). By contrast, we found a NOM effect in the N400 (an ERP component associated with semantic processing), as ambiguous words elicited larger N400 amplitudes than unambiguous words. Accordingly, this modulation of the N400 provided support for semantic enhanced activation accounts of the NOM advantage.

Nevertheless, semantic enhanced activation models differ in their explanation of how this large amount of semantic activation elicited by ambiguous words facilitate their recognition. The semantic feedback account (e.g., Hino & Lupker, 1996) holds that the semantic activation triggered by the multiple meanings of ambiguous words feedbacks to the orthographic level, boosting orthographic processing and thus making these words to reach faster the orthographic recognition threshold. In contrast, Borowsky & Masson's (1996) model assume that such semantic activation does not boost orthographic processing, so that ambiguous words are recognized faster because they increase the global semantic activation level. To test whether ambiguous words benefit from an orthographic boost, in the Study 4 we compared ambiguous and unambiguous words in a task that taps orthographic processing, in particular, a two-alternative forced-choice task (2AFC) (e.g., Gomez et al., 2008). We conducted two different 2AFC experiments, where a word (ambiguous or unambiguous) was briefly presented, and then participants were asked to select the word that was presented previously (Experiment 2) or the word orthographically related to that presented before (Experiment 3). We observed significant differences between ambiguous and unambiguous words in both experiments. This piece of evidence suggests that ambiguous and unambiguous words differ in their orthographic processing, providing further support for the hypothesis that the former benefit from an orthographic boost (e.g., Hino & Lupker, 1996).

In sum, the evidence obtained in the present thesis suggests that NOM facilitates word recognition regardless of ROM. It should be noted, however, that this facilitation effect for NOM only appears when subjective NOM measures are used. The best explanation for this effect is that the multiple semantic representations of ambiguous words boost orthographic processing through semantic feedback. In this way, ambiguous words would generate a large amount of semantic activation during processing, which would affect lower (i.e., word and letter) processing levels and thus accelerate their recognition. Of note, the same mechanism can accommodate the interactions observed between NOM and word frequency (Hino & Lupker, 1996; Pexman et al., 2004), and between NOM and nonword difficulty (Azuma & Van

Orden, 1997; Borowsky & Masson, 1996; Piercey & Joordens, 2000). For instance, high-frequency words do not benefit from a NOM advantage because their letter-to-word links are very strong, so that semantic feedback may have a marginal influence on orthographic processing. Similarly, the interaction between NOM and nonword difficulty would be due to the fact that the system would set a higher recognition threshold when nonwords are more wordlike. As such, by increasing the nonword difficult, semantic feedback would have more time to influence inferior processing levels. Finally, the semantic feedback account has been also proposed to explain the lack of a ROM effect. Namely, if polysemes and homonyms do not differ at the representational level, both polyseme and homonym meanings might be represented separately (Hino et al., 2010). Under this view, each ambiguous word meaning, regardless of their ROM, would spread its activation to its linked orthographic representation. Another possibility would be that, despite being represented differently from homonyms, polysemes could benefit from semantic richness effects by having a single rich semantic representation. Either way, polysemes and homonyms would elicit a comparable amount of semantic-to-orthographic feedback and, thus, would show a similar facilitation during word recognition.

4. Conclusions

To sum up, the conclusions that can be drawn from the present thesis are the following:

1. The approach employed for estimating number of meanings modulates the effect of NOM in LDT. In particular, a NOM disadvantage is found when NOM is estimated according to dictionary definitions, whereas a NOM advantage is observed when it is estimated on the basis of the speakers' knowledge.
2. Dictionary meanings are not representative of the word meanings known by the speakers. For this reason, subjective NOM is a psychologically more valid measure than dictionary NOM.
3. The recognition of words with more than one meaning is facilitated only when their meanings are known by the speakers.
4. NOM facilitates word recognition by boosting orthographic processing via semantic feedback
5. In contrast, ROM does not seem to influence word recognition. This may suggest that: a) polysemes and homonyms are represented similarly, or b) polysemes benefit during word recognition from having a single rich representation containing all their senses, whereas homonyms benefit from having multiple semantic representations.

References

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 273-278).

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2223-2228).

Azuma, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 109–124.

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, *36*(4), 484-504.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research*, *24*(1), 57-65.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187-222). Hillsdale, NJ: Erlbaum.

Borowsky, R., & Masson, M. E. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 63-85.

Britton, B. K. (1978). Lexical ambiguity of words used in english text. *Behavior Research Methods & Instrumentation*, *10*(1), 1–7.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior, 12*(4), 335-359.

Durkin, K., & Manning, J. (1989). Polysemy and the subjective lexicon: Semantic relatednes and the salience of interword senses. *Journal of Psycholinguistic Research*, *18*(6), 577–612.

Domínguez, A., Cuetos, F., & de Vega, M. (2001). 100 Palabras polisémicas con sus acepciones. *Revista Electrónica de Metodología Aplicada*, *6*(2), 63–84.

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: the current state of the literature. *Psychonomic bulletin & review*, *22*(1), 13-37.

Estevez, A. (1991). Estudio normativo sobre ambigüedad en castellano. *Psicológica*, *3*(2), 237–271.

Ferraro, F. R., & Hansen, C. L. (2002). Orthographic neighborhood size, number of word meanings, and number of higher frequency neighbors. *Brain and Language, 82*(2), 200-205.

Ferraro, F. R., & Kellas, G. (1990). Normative data for number of word meanings. *Behavior Research Methods, Instruments, & Computers, 22*(6), 491–498.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition*, *4*(1), 53-61.

Fraga, I., Padrón, I., Perea, M., Comesaña, M. (2017). I saw this somewhere else: The Spanish Ambiguous Words (SAW) database. *Lingua*, *185*, 1-10.

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, & Computers*, *26*(1), 5–25.

Gee, N. R., & Harris, S. L. (2010). Homograph norms: an alternative approach to determining meaning dominance. *Behavior Research Methods*, *42*(4), 976–986.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281.

Gómez Veiga, I., Carriedo López, N., Rucián Gallego, M., & Vila Cháves, J. O. (2010). Estudio normativo de ambigüedad léxica en castellano, en niños y en adultos. *Psicológica*, *31*(1), 25–47.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological review*, *115*, 577-600.

Gorfein, D. S., Viviani, J. M., & Leddo, J. (1982). Norms as a tool for the study of homography. *Memory & Cognition*, *10*(5), 503–509.

Griffin, Z. M. (1999). Frequency of meaning use for ambiguous and unambiguous words. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 520–530.

Hargreaves, I. S., & Pexman, P. M. (2012). Does richness lose its luster? Effects of extensive practice on semantic richness in visual word recognition. *Frontiers in human neuroscience*, *6*(234), 1-11.

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: when does relatedness matter? *Canadian Journal of Experimental Psychology*, *64*(3), 180–196.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1331.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 686-713.

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language*, *55*(2), 247-273.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*(2), 130.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive psychology*, *13*(2), 278-305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of verbal learning and verbal behavior, 14*(5), 534-537.

Joordens, S., & Besner, D. (1994). When banking on meaning is not (yet) money in the bank: Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1051-1062.

Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory & Cognition, 29*(2), 363-372.

Kawamoto, A. H., Farrar, W. T., & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1233-1247.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 601-609.

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language, 81*(1-3), 205-223.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, *20*(1), 1-24.

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language, 123*(1), 11–21.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology, 62*, 621-647.

Lin, C-J. C., & Ahrens, K. (2005). How many meanings does a word have? Meaning estimation in Chinese and English. In J. W. Minett & W. S.-Y. Wang (Eds.), *Language acquisition, change and emergence: Essays in evolutionary linguistics* (pp. 437–464). Hong Kong: City University of Hong Kong Press.

Lin, C-J. C., & Ahrens, K. (2010). Ambiguity advantage revisited: Two meanings are better than one when accessing Chinese nouns. *Journal of psycholinguistic research*, *39*(1), 1-19.

MacGregor, L. J., Bouwsema, J., & Klepousniotou, E. (2015). Sustained meaning activation for polysemous but not homonymous words: Evidence from EEG. *Neuropsychologia, 68*, 126-138.

Masson, M. E. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*, (pp. 233-263). Hillsdale, NJ: Erlbaum.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *The Quarterly Journal of Experimental Psychology*, *52*(4), 813-839.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition*, *17*(2), 141-147.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In *Processing of visible language* (pp. 259-268). Springer, Boston, MA.

Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The

University of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, *12*(1), 16–37.

Nickerson, C. A., & Cartwright, D. S. (1984). The University of Colorado meaning norms. *Behavior Research Methods, Instruments, & Computers*, *16*(4), 355–382.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, *9*(3), 225-236.

Panman, O. (1982). Homonymy and polysemy. *Lingua*, *58*(1-2), 105–136.

Parks, R., Ray, J., & Bland, S (1998). Wordsmyth English dictionary—Thesaurus. [ONLINE]. Available: http:// www.wordsmyth.net/ [2016, May].

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*(1), 161-167.

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1252-1270.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology*, *53*(4), 323.

Piercey, C. D., & Joordens, S. (2000). Turning an advantage into a disadvantage: Ambiguity effects in lexical decision versus reading tasks. *Memory & Cognition*, *28*(4), 657-666.

Real Academia Española. (2014). Diccionario de la lengua española (23rd ed.). Madrid, Spain.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*(2), 245–266.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, *28*(1), 89-104.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of verbal learning and verbal behavior*, *9*(5), 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of verbal learning and verbal behavior*, *10*(6), 645-657.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review, 89*(1), 60.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive psychology, 14*(4), 489-537.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of verbal learning and verbal behavior, 20*(1), 120-136.

Swaab, T., Brown, C., & Hagoort, P. (2003). Understanding words in sentence contexts: The time course of ambiguity resolution. *Brain and Language, 86*(2), 326-343.

Swinney, D. A. (1991). The resolution of interdeterminacy during language comprehension: perspectives on modularity in lexical, structural and pragmatic process. In G. B. Simpson (Ed.), *Understanding Word and Sentence* (pp. 367–385). North Holland: Elsevier Science Publishers B.V.

Tamminen, J., Cleland, A. A., Quinlan, P. T., & Gaskell, M. G. (2006). Processing semantic ambiguity: Different loci for meanings and senses. In *Proceedings of the twenty-eighth annual conference of the cognitive science society* (pp. 2222-2227).

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, *22*(1), 111–126.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language*, *26*(2), 188–208.

Wollen, K. A., Cox, S. D., Coahran, M. M., Shea, D. S., & Kirby, R. F. (1980). Frequency of occurrence and concreteness ratings of homograph meanings. *Behavior Research Methods & Instrumentation*, *12*(1), 8–15.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*(72), 1-10.

Universitat Rovira i Virgili
2018