Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition?

Juan Haro[1] and Pilar Ferré[1]

[1]Research Center for Behavior Assessment (CRAMC) and Department of Psychology. Universitat Rovira i Virgili. Tarragona. Spain.

Juan Haro

Research Center for Behavior Assessment (CRAMC) and Department of Psychology. Universitat Rovira i Virgili.

Crta. de Valls s/n, Campus Sescelades, 43007, Tarragona. Spain.

E-mail: juan.haro@urv.cat

Telephone Number: +34-977-558567

Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition?

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 & Tokowicz, 2015, and Simpson, 1984, for reviews). The results of these studies have been inconclusive concerning the role of semantic ambiguity 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 & Masson, 1996; Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; but see Forster & 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, Garfield, & Millikan, 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 & Masson, 1996) or by a large semantic-to-orthographic feedback (Balota, Ferraro, & Connor, 1991; Hino & 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 & Plaut, 2008, 2011; Beretta, Fiorentino, & Poeppel, 2005; Klepousniotou & Baum, 2007; Rodd, Gaskell, & Marslen-Wilson, 2002; Tamminen, Cleland, Quinlan, & Gaskell, 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 & Van Orden, 1997; Millis & 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, Gaskell, and Marslen-Wilson (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 & 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, Kusunose, & Lupker, 2010; Hino, Pexman, & Lupker, 2006; Pexman, Hino, & Lupker, 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 & Masson, 1996; Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Millis & 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 & Masson, 1996; Hino & 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 & 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 & 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, Demestre, Boada, & Ferré, 2017; Hino et al., 2010; Hino et al., 2006; Lin & Ahrens, 2010; Pexman, Hino, & Lupker, 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 & Ahrens, 2010), and Spanish (Haro et al., 2017). 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, Ray, & Bland, 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 & 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 2x2 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 *p*s > .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, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). On the other hand, familiarity, concreteness, subjective age of acquisition, arousal and emotional valence ratings were taken from the database of Haro, Ferré, Boada, and Demestre (2017).

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,	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
few senses	(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,	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
many senses	(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,	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
few senses	(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,	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
many senses	(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									
	-	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
Unambiguous		(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)
	-	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
Homonyms		(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)

Note. 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.

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 & 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 & Forster, 2003) was used for displaying the stimuli and recording 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.

Data analyses

Data on reaction times (RT) and error rates (%E) were analyzed separately using linear mixed-effect models (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008). For this we used the lme4 package of R (Bates, Maechler, Bolker, & Walker, 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 x 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, Brockhoff, & Christensen, 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 ($p < .05$).

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)

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 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; Hino et al., 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 & 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. (2017). 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; Hino et al., 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 ps $< .05$). 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; Hino et al., 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 & 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, Padrón, Perea, & Comesaña, 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. (2017)'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 *p*s > .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%). 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 ($p < .05$).

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)

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 & 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; Hino et al., 2006; Lin & 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., 2017). 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. (2017)'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 *p*s > .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; Hino et al., 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 & 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 & Masson, 1996; Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Millis & Button, 1989), and with the findings of more recent studies that have relied on subjective norms (e.g., Haro et al., 2017; Hino et al., 2010; Hino, Lupker, & Pexman, 2002; Hino et al., 2006; Lin & 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 & 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 & 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 & 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, Siakaluk, & Yap, 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, Pexman, Wellsby, Hargreaves, Huff, 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 & 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

1. We thank an anonymous reviewer for this suggestion.

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 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 & Joordens, 1995; Masson & 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, Morris, & Rayner, 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 & 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.

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., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445–459. doi: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. Available online at: 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*, 515–301.

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. doi: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. [ONLINE]. Available: 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.