Contribution of Word Meaning and Components of Familiarity to Lexical Decision: A Study With Pseudowords Constructed From Words With Known or Unknown Meaning

Jean-François Le Ny and Françoise Cordier

Electronic version
URL: http://cpl.revues.org/416
ISSN: 1379-6100

Electronic reference

This text was automatically generated on 7 octobre 2016.

© All rights reserved
Lexical decision tasks and word pronunciation tasks are viewed as involving word representations that have two main constituents: the representation of the word form, which is orthographic in the case of written words, and the word meaning. There is now fairly general agreement that these two types of lexical knowledge contribute to lexical decision and word recognition. Balota, Ferraro, and Connor (1991), and Balota (1994) have reviewed the literature on the contribution of meaning to recognition performances, in particular in lexical decision tasks. They reported a number of studies, confirmed by subsequent research, indicating that several properties of word meaning influence word recognition: concreteness (Kroll & Merves, 1986), polysemy (Borowsky & Masson, 1996; Durkin & Manning, 1989; Hino, Lupker, & Pexman, 2002), associative links (Balota & Chumbley, 1984) and added meaning (Forster, 1985; Whittlesea & Cantwell, 1987). Thus, the idea that meaning contributes to lexical decision finds support in a wide range of reported data.

Another type of study has been suggested by Le Ny (1989). This suggestion is based on the observation that memory contains a particular type of word that has never been used in experimental studies: namely, known words devoid of meaning. We all have intuitive knowledge in our long-term memory of word forms that we know to be actual words in our language, although we are unable to say what they mean. Although these words of unknown meaning (UM-W) that are similar to words of known meaning (KM-W) are a
marginal part of everyone's lexicon, they permit a natural dissociation of the knowledge of word form representations from that of word meaning. This UM vs. KM factor could help us to distinguish better between the respective roles of lexical-unit constituents in the processing of KM words.

3 Word familiarity is also an important factor in word processing. Several authors (Azuma & Van Orden, 1997; Balota & Chumbley, 1984; Durkin & Manning, 1989; Gernsbacher, 1984) have stressed its role as a complement to that of meaning. Balota and Chumbley (1984) viewed it as forming composite information, and Balota and Spieler (1999) wrote that lexical decision is possible "because words are more familiar than nonwords (also more meaningful)."

4 However, meaning as knowledge can also be assumed to involve familiarity, as it results from word experience in the area of semantic processing. Gernsbacher (1984) has pointed out that word experience must be considered across all the modalities that relate to a word. The same observation applies to meaning. Speakers acquire familiarity with a word not only from encountering and perceiving it, but also from the semantic operations involved in its processing, particularly in comprehension, which requires the activation of both the word form representation and its meaning. Familiarity with these two representations may therefore be simultaneously increased every time the word is processed. If this is true, the familiarity values usually considered for KM words do not express a single global state of familiarity associated with the word, but rather a summation of at least two distinct familiarity values, each associated with one of the two main parts of the word representation: form and meaning.

5 On the other hand, the use of various kinds of nonwords in lexical decision experiments has also shown that wordlikeness, i.e. conformity with the regularities of the language in question, plays a role in lexical decision. This suggests that familiarity with form might, in turn, be assumed to consist of two sub-components, familiarity with a general form, which is a function of the regularity of the word across the language, and familiarity with the specific form, which is dependent on each individual word.

6 The results of an earlier experiment using familiarity ratings support these assumptions (Cordier & Le Ny, submitted). The material mainly comprised French KM and UM words. Their linguistic and experiential frequencies were equivalent. However, their rated familiarity was different, with the KM-W being more familiar than the UM-W. Unknown words (U-W) were also used.

7 This initial data tended to confirm the assumption that word familiarity involves two components that are related to the form (general and specific) of the word in question, and the meaning of that word. The goal of this study was to test the contribution of the postulated components of familiarity in a lexical decision task by using pseudowords constructed from KM words (KM-Pwords), UM words (UM-Pwords) and U words (U-Pwords).

8 The use of pseudowords makes it easier to dissociate the effects of familiarity and linguistic frequency, since the frequency of pseudowords is equal to zero. The pseudowords in our study were derived from the words we had used previously by replacing a single letter in them. The familiarity ratings for the pseudowords were collected first: the KM-Pwords were rated as being more familiar than the UM-Pwords. These pseudowords were used in the lexical decision task, in which they were contrasted...
with normal KM words. Frequent KM words were chosen because we expected them to produce greater word/pseudoword discriminability.

9 The data analysis examined the time it took to reject the pseudowords in the lexical decision task. We assumed that each of the postulated components of familiarity would be involved in the sequence of processes that leads to the decision. They should generate a transient state of familiarity, in relation to the rated familiarity of the original words, that should then intervene in the processing of the pseudowords and produce different degrees of interference.

10 The existence of an orthographic neighbour has been shown to produce interference in lexical decision on words (Andrews, 1997; Forster & Hector, 2002; Grainger & Jacobs, 1996). More specifically, the properties of words that are neighbours of pseudowords have an effect on the processing of these pseudowords: O’Connor & Forster (1981) showed that misspelled versions of high-frequency words produce more errors than misspelled versions of low-frequency words. We assumed that different degrees of interference would similarly be produced by orthographic neighbours of our pseudowords, since they differ in terms of participants’ knowledge of their meaning and their familiarity.

11 More precisely, we postulated that full familiarity with the meaning of earlier KM words will be lost when these words are changed into KM-Pwords, but that these might still be able to activate the meaning of their original neighbour KM word while being processed for a speeded decision. Although full familiarity with the specific form would also be lost for KM-Pwords, they might still be able to activate the familiar specific word form of the original KM stimulus during processing. However, in the case of these KM-Pwords, familiarity with a general word form would be preserved because the replacement of a single letter is not sufficient to eliminate their wordlikeness. We similarly postulated that UM-Pwords would continue to generate no familiarity with any meaning and cease to generate full familiarity with a specific form, but would still be capable of activating the familiar specific word form of the corresponding original UM stimulus during processing. Familiarity with a general word form would persist. Finally, U-Pwords should result in no loss of familiarity with a meaning or a specific word form, while familiarity with a general word form should be conserved. Given these assumptions, we predicted that more interference would be generated by KM word neighbours than by UM word neighbours and by UM word neighbours than by U word neighbours. Consequently, we predicted that it would be more difficult for participants to reject the KM-Pwords than the UM-Pwords, and that they would also find it more difficult to reject these than U-Pwords. The no decision times were expected to differ accordingly.

12 25 undergraduate psychology students from the University of Poitiers, native French speakers with normal or corrected-to-normal vision, volunteered to participate in this experiment.

13 The 3x15 original KM, UM and U words were nouns. They had a linguistic frequency of less than 6.55 occurrences per million, as specified in Lexique (New, Pallier, Ferrand, & Matos, 2001) and did not differ in frequency: F(2,28) = 2.62, p<.08. Their French first letter frequency (Andrews, 1997) was identical. They comprised no suffixed words and had no orthographic neighbours. As linguistic frequency is uncertain for low-frequency words, experiential frequency was determined in a rating experiment (Table 1).
45 frequent usual words were also selected. They were also nouns, with a relative frequency of over 27 occurrences per million, Lexique (New et al., 2001). Their experiential frequency was also determined by rating.

Familiarity ratings were collected for these different categories of words. They are presented in Table 1.

<table>
<thead>
<tr>
<th>Categories of original words</th>
<th>Experiential frequency</th>
<th>Rated familiarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Known meaning words</td>
<td>1.81 (.37)</td>
<td>5.94 (.36)</td>
</tr>
<tr>
<td>Unknown meaning words</td>
<td>1.63 (.63)</td>
<td>1.93 (.57)</td>
</tr>
<tr>
<td>Unknown words</td>
<td>1.34 (.02)</td>
<td>1.12 (.10)</td>
</tr>
<tr>
<td>Frequent usual words</td>
<td>5.85 (.43)</td>
<td>6.64 (.19)</td>
</tr>
</tbody>
</table>

An ANOVA was performed on the experiential frequency (means collapsed across items). This revealed that the critical difference, i.e. that between the UM words and KM words, was not significant: t (28) = -0.18, p = .166. There was a significant difference between the frequent usual words and the other categories. The differences between the unknown words and the other categories were also significant: U words vs. UM words: t (28) = -4.45, p = .0257; U words vs. KM words: t (28) = 5.314, p = .0004.

Another ANOVA was performed on the familiarity ratings (means collapsed across items). It showed that the category effect was reliable: F(5, 245) = 2020.1, p < .0001. The difference between UM words and KM words was significant: F(1,245) = 2461, p < .0001, as was the difference between U words and UM words: F(1, 245) = 98.6, p < .0001. A comparison of the scores for the U words with the fixed value 1 proved to be significant: t = 10.9, p < .01. The difference between the high-frequency words and the low-frequency KM words was significant: F(1, 245) = 64.9, p < .0001.

Three sets of pseudowords were derived from these words, 15 from the KM words, 15 from the UM words and 15 from the U words. The pseudowords were constructed by randomly changing either an initial, middle or final letter in the original words. All the resulting pseudowords were pronounceable. None had orthographic neighbours other than the original word.

Familiarity ratings were also collected for these different categories of pseudowords, and for the frequent normal words. They are presented in Table 2. Familiarity with the frequent normal words was thus rated twice: in the context of words (Table 1) and in the context of pseudo-words (Table 2). The pseudowords were contrasted in the lexical decision task with the normal words and there was a 50/50 balance between expected yes and no responses.

The participants were tested individually. They sat in a naturally lit room at a comfortable reading distance, about 50 cm, in front of a Medibus 17-inch monitor. The
words were displayed on the screen in black 12-point Times typeface. The material was presented using a Pentium PCI-ISA computer running a specially developed program that also recorded the response latencies and unexpected responses.

The items were presented in a random order, which was different for each participant. They appeared one after the other in the centre of the screen. Eye fixation was obtained before each presentation through the use of a star which was displayed for 100 milliseconds in the place where the next item was to appear. It was replaced by the item 100 milliseconds after it has disappeared. The participants were instructed to respond by pressing one of the two buttons of a mouse on which their preferred hand was positioned, positive responses being made with the forefinger and negative ones with the middle finger. The inter-trial intervals were controlled by participants: they pressed one of the response buttons when they wanted to see the next item.

The experiment began with a training session of 16 preparatory items, similar in nature and proportions to the experimental ones. The participants then rested for five seconds (this period being indicated on the screen) and then accessed the experimental items by pressing one of the response keys.

The participants were asked to decide, as quickly and accurately as possible, whether or not each sequence of letters was a French word.

An ANOVA was conducted on the familiarity ratings with items as the random factor: it showed that the main effect was significant: $F(3,42) = 30.28, p < .0001$. Pairwise comparisons revealed a significant difference between the 3 PW types: KM-Pwords vs UM-Pwords: $F(1,42) = 4.16, p < .001$; UM-Pwords vs U-Pwords: $F(1, 42) = 3.07, p < .008$.

The overall difference for yes responses to Pwords in the lexical decision task (Table 2) was significant: $F(2, 24) = 4.74, p < .0132$. But only the pairwise comparison for KM-Pwords vs. U-Pwords revealed a significant difference: $t(24) = 3.30, p < .003$. The other comparisons (KM-Pwords vs UM-Pwords and UM-Pwords vs U-Pwords) did not. The corresponding decision times were removed from the final data.

All decision times longer or shorter than two standard deviations from the means were removed from the data, i.e. 2.0%, 3.0% and 1.0% respectively. Final data per condition are summarized in Table 2.

Table 2: Mean Rejection Times for the Three Categories of Pseudowords and Mean Acceptation Times for Usual Words, in Milliseconds, Standard Deviations, Mean Percentages of Errors and Values of Rated Familiarity (with SD).

<table>
<thead>
<tr>
<th>Categories of pseudowords</th>
<th>Mean RT</th>
<th>SD</th>
<th>Errors</th>
<th>Rated familiarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>KM-Pwords</td>
<td>945</td>
<td>238</td>
<td>8.5%</td>
<td>2.02 (.56)</td>
</tr>
<tr>
<td>UM-Pwords</td>
<td>894</td>
<td>226</td>
<td>6%</td>
<td>1.43 (.24)</td>
</tr>
<tr>
<td>U-Pwords</td>
<td>838</td>
<td>208</td>
<td>3%</td>
<td>1.24 (.14)</td>
</tr>
<tr>
<td>Usual words</td>
<td>662</td>
<td>125</td>
<td>1.1%</td>
<td>6.21 (.35)</td>
</tr>
</tbody>
</table>

An Anova on the rejection times for pseudowords was conducted with subjects ($F_1$) and items ($F_2$) as random factors. The main effect was highly reliable: $F_1(2, 48) = 20.66, p < .0001$; $F_2(2, 42) = 6.61, p < .0032$. Pairwise comparisons revealed a significant difference between our main Pword types: $F_1(1, 48) = 9.14, p < .0004$ and $F_2(1, 42) = 5.18, p < .027$ for KM-Pwords vs. UM-Pwords. Only the comparison on participants revealed a significant
difference for UM-Pwords vs U-Pwords, but not the comparison on items: \( F(1, 48) = 11.57, p < .001; F(1, 42) = 1.73, \text{NS} \).

The means of the decision times pointed to order KM-Pwords > UM-Pwords > U-Pwords in compliance with the order KM-words, UM-words, U-words postulated for the original stimuli. This order also paralleled that of the rated familiarity values collected for the categories of pseudowords and original words. There was only one significant difference in the error rates on the U-Pwords. The main result was that the rejection times were longer for KM-Pwords than UM-Pwords. This result accords with others previously obtained for lexical decision on words, thus indicating that this task involves a contribution of meaning. Our results showed that word meaning also intervenes in the processing leading to a rejection decision.

Word meaning, or more precisely knowledge of meaning, expressed separately for each word in the word familiarity ratings, contributed to decision times in this situation by introducing different degrees of interference. This effect did not come from the meaning of the stimulus itself, but from that of its unique neighbour. It is consistent to view these differences in rejection times for the KM and UM-Pwords as indicating the respective contributions to pseudoword processing of the two initially postulated components of familiarity. This difference (51 ms) shows a greater interference for KM-Pwords, which is presumably due to a greater derived familiarity resulting from their neighbours. In this case, the participants' familiarity can be formulated as: "this stimulus resembles a word I know and whose meaning I know" for the KM-Pwords and simply "this stimulus resembles a word I know" for the UM-Pwords. Such implicit analog knowledge could have intervened during the lexical decision processing of these stimuli as a transient state of derived familiarity with them.

Although meaning was not involved in these two categories of stimuli, it can be argued that the time difference observed on the UM-Pwords and U-Pwords (56 ms) similarly indicates a difference in derived familiarity: with the general form of the words only (wordlikeness) in the case of the U-Pwords, and additionally with the specific form of each particular word in the case of the UM-Pwords.

These results can help us gain a better understanding of how meaning affects lexical decision in usual situations: through the effect of two distinct components of word familiarity, relating to word form and meaning respectively.

Our results showed that no complete identification of a particular meaning was necessary to modify decision times: derived familiarity with the meaning of a neighbour word was sufficient. This can be accounted for by assuming that a transient state of familiarity, involving the interaction of the two components of familiarity, took place in our situation, and generally takes place in normal lexical decision tasks during the processing sequence that leads to decision. This use of the notion of a transient familiarity is closely related to that used in many memory models that attribute an important role to familiarity in recognition (Balota & Chumbley, 1984; Gillund & Shiffrin, 1984; Goldinger, 1998; Hintzman, 1984; Joordens and Becker, 1997; Murdock, 1982; Ratcliff & McKoon, 1988). These models postulate that the recognition of a stimulus always involves a transient state of familiarity, which is generated and gradually increases as the stimulus is processed, and the level of which results in recognition if it attains a critical value. This process would also appear to underpin lexical decision.
If we consider our data on pseudoword decision in conjunction with the rated familiarity values, we can understand that the process leading to no decisions in this situation involved two phases: an early transient familiarity state was generated at the beginning of processing, due to activation resulting from one or both of the two memory sources, depending on the stimulus: representation of the neighbour word form and meaning. This activation first emerged and was accumulated as a function of its components of familiarity, as if to produce a yes decision, but was subsequently inhibited, thus producing a final state of unfamiliarity that resulted in the no decision and response.

The different levels of interference in the processing of pseudowords in this experiment can thus be interpreted as being due to different degrees of resistance-to-rejection-before-rejection for the three categories of pseudowords. This would involve different levels of inhibition being applied, during the second phase of processing, to the inappropriate activation generated in the first phase. Inhibition can be viewed as being similar, in this situation, to that involved in many cognitive activities, in particular lexical access and discourse comprehension (Dagenbach & Carr, 1994; Gernsbacher 1997; Gernsbacher & Faust, 1995; Simpson & Kang, 1994). This experiment illustrates a type of suppression that should be added to those identified by Gernsbacher (1997). It is to some extent similar to the process assumed to take place in the disambiguation of ambiguous words, i.e. initial processing activates two meanings and further processing suppresses the inappropriate one (Faust & Gernsbacher, 1996; Gernsbacher & Faust, 1995). Only one meaning, that of the neighbour word, could be activated for KM-Pwords, in addition to the neighbour word's word form representation. However, both were inappropriate, and had to be suppressed. A representation of the neighbour word form was also presumably activated in the case of the KM-Pwords, and this activation was also inappropriate. The decision times were then dependent on the rate of initial activation, and the subsequent inhibition required: our results are in agreement with the view that the period necessary to achieve the suppressions is a function of the strength of the inhibition that has to be produced during this processing.

This schema helps us understand how the components of familiarity, which contributed to no lexical decisions on our pseudowords, may similarly contribute to the yes responses in a normal lexical decision task. The use of words with an unknown meaning in a complex lexical decision task can thus help us gain a better understanding of the components and processes involved in regular lexical decision. However, no matter how these processes work, one general conclusion that can be drawn from this study is that distinct components of familiarity should be taken into consideration in any study involving lexical decision.

We thank Tim Pownall for editing this article.

BIBLIOGRAPHY


Current psychology letters, 12, Vol. 1, 2004 | 2004


REFERENCES


NOTES

1. A complete list of the stimuli can be obtained from the LIMSI Internet site at http://www.limsi.fr/Individu/leny/index.html

ABSTRACTS

45 French words were selected: 15 of them were known by the participants who were also aware of their meaning (KM-W), 15 were known although their meaning was unknown (UM-W), and 15 were unknown (U-W). Although linguistic and experiential frequencies were equivalent for the KM and UM words, familiarity ratings indicated that the former were more familiar. This suggests that usual word familiarity involves two components: familiarity with the word form, and familiarity with the meaning. Pseudowords were constructed from the original words: the familiarity ratings also indicated that KM-Pwords were more familiar than UM-Pwords. These pseudowords were used in a lexical decision task which also made use of normal words. No responses were longer for KM-Pwords than UM-Pwords. This finding supports the assumption
that two components are involved in the familiarity of normal words and that this fact should be taken into consideration in any experiment involving lexical decision.

On a sélectionné des mots français appartenant à 3 catégories: mots connus des participants et de signification connue d’eux (KM-W), mots connus mais de signification inconnue (UM-W), mots complètement inconnus (U-W). Bien que les mots KM et UM fussent de fréquences linguistique et subjective équivalentes, des jugements de familiarité montrèrent que les premiers étaient plus familiers. Cela suggère qu’il existe deux composants principaux de la familiarité des mots: avec leur forme et avec leur signification. On a construit à partir de là des pseudo-mots KM-Pwords, UM-Pwords et U-Pwords. Des jugements de familiarité montrèrent à nouveau que les KM-Pwords étaient plus familiers que les UM-Pwords. Ces pseudo-mots furent utilisés dans une tâche de décision lexicale en opposition à des mots usuels. Les temps de décision Non furent plus longs pour les KM-Pwords que pour les UM-Pwords. Cela corrobore l’hypothèse de deux composants de la familiarité des mots: ce fait devrait être pris en considération dans toute étude utilisant la décision lexicale.

INDEX

*Keywords*: meaning, lexical decision, pseudowords, word familiarity

AUTHORS

JEAN-FRANÇOIS LE NY
LIMSI-CNRS, University of Paris-Sud,
BP 133, 91403 Orsay Cedex, France.
E-mail: Jean-Francois.Le.Ny@limsi.fr

FRANÇOISE CORDIER
Laco, FRE CNRS 2725, University of Poitiers,
99 avenue du Recteur Pineau, 86022 Poitiers Cedex, France.
E-mail: francoise.cordier@mshs.univ-poitiers.fr