available at www.sciencedirect.com

SCIENCE DIRECT®

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

On understanding idiomatic language: The salience hypothesis assessed by ERPs

Jean-Paul Laurent[a,*], *Guy Denhières*[b], *Christine Passerieux*[c], *Galina Iakimova*[c], *Marie-Christine Hardy-Baylé*[c]

[a]*Equipe de Recherche en Psychologie Clinique, Université de Paris 8, UFR de Psychologie, Pratiques Cliniques et Sociales, 2, rue de la Liberté-93526 Saint-Denis Cedex 02, France*
[b]*Laboratoire de Psychologie Cognitive UMR 6041 CNRS Université de Provence, France*
[c]*Laboratoire Universitaire de Recherche, Hôpital Mignot, 78150-Le Chesnay, France*

ARTICLE INFO

Article history:
Accepted 28 October 2005
Available online 4 January 2006

Theme:
Neural basis of behavior
Topic:
Cognition

Keywords:
Salience
Figurative utterance
Idiom
Pragmatic
N400
P600
Psychophysiology

ABSTRACT

Giora's [Giora, R., 1997. Understanding figurative and literal language: the Graded Salience Hypothesis. Cogn. Linguist. 7 (1), 183–206; Giora, R., 2003. On Our Mind: Salience Context and Figurative Language. Oxford Univ. Press, New York] Graded Salience Hypothesis states that more salient meanings–coded meanings foremost on our mind due to conventionality, frequency, familiarity, or prototypicality–are accessed faster than and reach sufficient levels of activation before less salient ones. This research addresses predictions derived from this model by examining the salience of familiar and predictable idioms, presented out of context. ERPs recorded from 30 subjects involved in reading and lexical decision tasks to (strongly/weakly) salient idioms and (figurative/literal) targets indicate that N400 amplitude was smaller for the last word of the strongly salient idioms than for the weakly salient idioms. Moreover, N400 amplitude of probes related to the salient meaning of strongly salient idioms was smaller than those of the 3 other conditions. In addition, response times to salient interpretations (the idiomatic meanings of highly salient idioms and the literal interpretations of less salient idioms) were shorter compared to the other conditions. These findings support Giora's Graded Salience Hypothesis. They show that salient meanings are accessed automatically, regardless of figurativity.

1. Introduction

Two types of psychological models have been proposed to explain the cognitive processes involved in understanding idiomatic expressions: non-compositional and compositional models. The main differences between these two types of models are based on the importance assigned to exceptions to the standard theory of literal processing and to the modalities of meaning construction (Marquer, 1994; Denhière and Verstiggel, 1997; Glucksberg, 2003). Cacciari and Tabossi (1988) proposed a configurational model, which is compatible with the compositional view. This model postulates complex units or familiar configurations of words in an individual's memory, some of which are idiomatic while others are not. For example, "Allons enfants de la patrie" (first line of French National Anthem), "chercher la petite bête" (to split hairs), and "un,

* *Corresponding author*. Fax: +33 01 49 40 67 54.
E-mail address: Jean-Paul.Laurent@univ-paris8.fr (J.-P. Laurent).

0006-8993/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2005.10.076

deux, trois, soleil" (first words of a French counting rhyme) can be defined as fragments that are highly entrenched in the knowledge network of an individual and possess a certain degree of fixedness. Literal meaning related to elements of these configurations can thus be always activated, whereas, as soon as the subject has enough clues to recognize a configuration, it can be activated as is. However, once a subject realizes that a configuration is not fixed and may vary from one expression to another, what differentiates between them is what Cacciari and Tabossi call the "idiomatic key". Among other factors, this key depends on the predictability of idiomatic expressions and the incongruent context in which they are presented. Titone and Connine (1994a,b) extended the configurational model proposed by Cacciari and Tabossi. In particular, they studied the effects of four factors believed to influence the processing of idioms: familiarity, predictability (the probability that an idiom can be rapidly recognized as such), literality (the probability of the literal interpretation of an idiom), and "decomposability" (the contribution of the meanings of the words to the idiomatic meaning of the expression).

Giora's Graded Salience model (Giora, 1997) introduces another point of view which eliminates the distinction between processing literal and figurative meaning but suggests instead the salience–nonsalience continuum. This alternative should resolve the debate about sequential versus direct processing with regard to idiomatic expressions. The Graded Salience Hypothesis assumes two distinct mechanisms: one bottom–up, sensitive only to linguistic information, and another top–down, sensitive to both linguistic and extra-linguistic knowledge. However, unlike the traditional modular assumption (Fodor, 1983), it assumes that the bottom–up lexical access mechanism is ordered. According to this view, coded meanings foremost on our mind due to familiarity, frequency, conventionality, or prototypicality would be accessed when encountered, regardless of contextual information or authorial intent. Coded meanings of low salience may not reach sufficient levels of activation to be visible in a context biased toward the more salient meaning of the word. Consequently, it is the degree of salience rather than the figurative or literal nature of an expression that determines the type of processing involved. Salience is the property through which the meaning of an expression is computed directly on the basis of entries in the mental lexicon rather than from inferences made on the basis of the linguistic or extra-linguistic context. According to the Graded Salience Hypothesis, idioms may be processed directly when their meaning is highly salient, as in the case of conventional idioms. On the other hand, processing may be sequential when a less salient meaning is intended, as in the case of new metaphors and the literal meaning of conventional idioms.

To distinguish between these two models, our study used event-related potential (ERP) methods. The study of sentence processing has advanced considerably since the development of ERP methods. Various language-relevant ERP components have been reported over the years. A well-established language-relevant ERP component is the N400. The N400 has a negative polarity, and it peaks at about 400 ms after word onset. Kutas and Hillyard (1984) showed that the N400 is sensitive to semantic violations in a sentence. However, it is not the case that only semantic violations elicit N400 effects. For instance, several studies have shown that N400 is also sensitive to semantically correct but relatively unexpected words. Specifically, it has been shown that N400 amplitude is inversely related to subjective predictability (cloze probability) of a word in its context (Kutas and Hillyard, 1984; Kutas et al., 1984). On the other hand, syntactic violations elicit a different ERP pattern altogether. This component, commonly referred to as P600, is a late centro-parietally distributed potential with positive polarity, starting at about 500 ms and typically extending up to at least 800 ms (Harris et al., 2000). In the domain of language processing, a P600 effect occurs in response to sentences that (a) contain a syntactic violation, (b) have a non-preferred syntactic structure, or (c) have a complex syntactic structure (Gunter et al., 1997; Friederici et al., 1999; Osterhout et al., 2002; Hagoort, 2003; Kuperberg et al., 2003; Wicha et al., 2003). P600 has therefore been claimed to reflect various kinds of syntactic processing difficulties, such as the inability of the parser to assign the preferred structure to incoming words or difficulty with syntactic reanalysis or syntactic integration (Kotz and Friederici, 2003; Kaan et al., 2000; Matzke et al., 2002; West and Holcomb, 2002; Kim and Osterhout, 2005). Finding of P600 effects in response to abnormal semantic reversal in Dutch sentences (Kolk et al., 2003) was unexpected because these sentences did not possess the characteristics previously associated with P600 effects. van Herten et al. (2005) tested the hypothesis that the P600 response to semantic reversal anomalies could have been caused by mismatch between an observed and a predicted grammatical morpheme. Their results demonstrated the weakness of this hypothesis. The authors ended up proposing an alternative hypothesis suggesting that the mismatch between the expected and the observed but unexpected sentence meaning triggered the P600 effects.

A certain number of studies using evoked potentials have so far been published in the field of metaphor processing (Blasko and Connine, 1993; Pynte et al., 1996; Passerieux et al., 2000; Bonnaud et al., 2002; Coulson and Van Petten, 2002; Tartter et al., 2002; Kazmerski et al., 2003). These studies have evaluated the N400 component of ERP in fairly different paradigms. Most of them have reported an increase in the amplitude of this component induced by metaphors compared to literals. For instance, in a study by Pynte et al., results supported the literal–nonliteral dichotomy. They showed elevated N400 for metaphors compared to literals. These results were partly replicated by Coulson et al. Coulson et al. presented participants with expressions whose final word either had a plausible literal meaning, a plausible metaphoric meaning, or involved difficulties of integrating literal and metaphoric meanings (Fauconnier and Turner, 1998). Their findings in fact propose a literal–nonliteral continuum rather than a literal–nonliteral divide. Tartter et al. used literals and novel metaphors and showed that a difference between metaphors and literals appeared as early as the time window of N200. As the authors suggested, this effect might have reflected differences in the characteristics of the materials. In their study, the frequency of the last word of the metaphoric phrases was lower than that of literal phrases. Nonetheless,

these differences disappeared at the N400 window, even though N400 component has long been shown to be sensitive to a number of psycholinguistic parameters such as word frequency, cloze probability, incongruity, familiarity, and concreteness (Kutas and Hillyard, 1980, 1984; Holcomb, 1993; Kounios and Holcomb, 1994; Van Petten, 1995; Kutas and Iragui, 1998; Van Petten et al., 1999; West and Holcomb, 2000; Federmeier and Kutas, 2001). On the basis of all of these findings, it seems essential to check the nature of experimental materials if we wish to be certain that any change in N400 truly reflects a difference between metaphoric and literal processes.

In this study, we aimed to test the hypothesis that it is not the literality or the metaphoricalness of the items that is reflected in N400 amplitudes but the degree of salience of the items involved. To do that, we tested idiomatic expressions whose literal and figurative meanings make sense, and which, unlike metaphors, enjoy a high degree of entrenchment or fixedness. Indeed, to be considered idiomatic, an expression must be recognizable and identifiable as such and make up an entry or be listed in the mental lexicon (Rat, 1999). As shown for homophones and homographs (Mullet and Denhière, 1997; Thérouanne and Denhière, 2002), this type of expression allows to study the time course of meaning activation and the effect of context on such activation (Giora and Fein, 1999). More specifically, we aimed to compare between two types of models that examine the comprehension of idiomatic expressions. One such model derives from Giora's Graded Salience Hypothesis. The other one is the configurational model of Cacciari and Tabossi. To tease apart these two models, we manipulated the degree of salience of familiar and predictable idioms presented out of context. To tap the activation of the various plausible meanings of these idioms, we recorded evoked potentials. We focused on N400 and P600 component amplitudes because the first discloses the time course of semantic memory involved in language comprehension (Kutas and Federmeier, 2000) and the second is sensitive to domain-general factors of probability, task relevance, and syntactic or semantic expectations (Coulson et al., 1998). Research into syntactic ambiguity alludes to the role lexical salience plays in the resolution of syntactic ambiguity (Trueswell et al., 1994; Trueswell, 1996; Trueswell and Kim, 1998). Use of an associative priming paradigm should allow us to evaluate the strength of the activation of literal and/or figurative meanings of idiomatic expressions.

According to Giora's hypothesis, we predicted that the N400 and P600 amplitudes for the last word would be smaller for strongly salient items compared to weakly salient items and that N400 would be smaller for targets compatible with the strongly salient meaning of the idiom (i.e., figurative meaning) compared to targets in the other three conditions. According to the configurational model, we predicted that there would be no significant differences in N400 and P600 amplitudes for the last word for strongly and weakly salient idioms and no significant differences in N400 and P600 amplitudes for the four different targets.

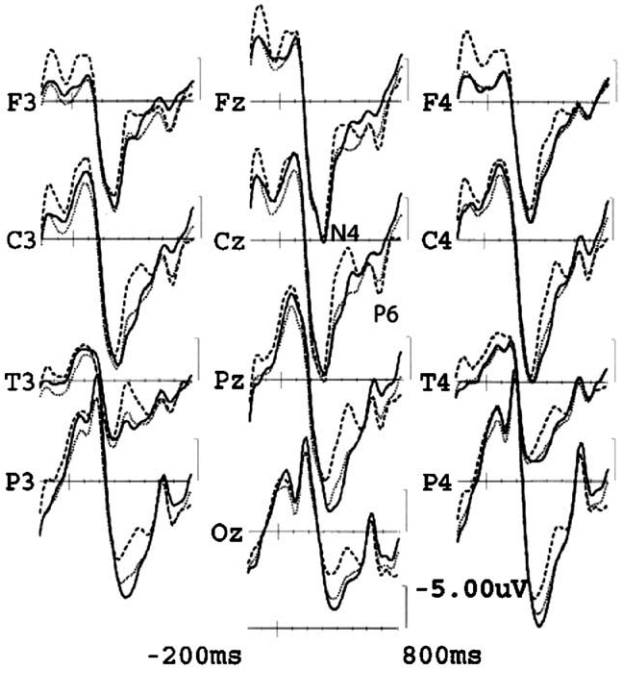

Fig. 1 – Evoked potentials in response to the last word of strongly salient (solid line) and weakly salient (dashed line) idiomatic expressions and control literal expressions (dotted line) from 12 electrodes between –200 ms and 800 ms after the beginning of the last word presentation.

Table 1 – Mean amplitude (μV) of LPC according to (*I+*) strong or (*I–*) weak saliency at Fz, F3, Cz and C4 sites

	Fz	F3	Cz	C4
I+	1.37	0.35	2.31	2.10
I–	3.92	1.73	4.01	3.28
P	0.00	0.02	0.00	0.04

I+: strongly salient idioms, *I–*: weakly salient idioms, *P*: probability.

2. Results

2.1. *Last word of experimental idiom*

Fig. 1 illustrates average waveforms of the last words for strongly and weakly salient idioms at twelve electrode sites. Visual inspection of recorded plots made it possible to identify two distinct components which were modulated in different ways by saliency factor. The earlier component, N400 (Kutas and Hillyard, 1984; Bentin et al., 1985; Kutas, 1985), was largest in the 320–500 ms window; and the later component P600 was largest in the 500–700 ms window after the last word of the expression.

MANOVA indicated a main effect of saliency (Rao $R_{(2,28)} = 511.64$ $P < 0.000$) of electrode site (Rao $R_{(22,8)} = 10.42$

Table 2 – Percentage of hits according to (*I+*) strong or (*I–*) weak saliency and type of target word

	Literal target (%)	Figurative target (%)
Strongly salient idiom	62.67	90.33
Weakly salient idiom	83.00	58.67

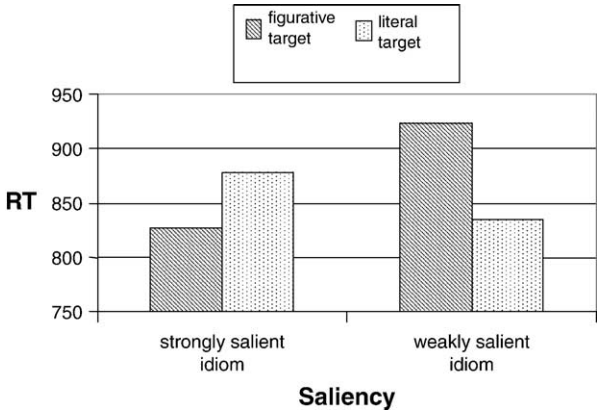

Fig. 2 – Response time according to saliency and target word.

P < 0.000) and an electrode × saliency interaction (Rao $R_{(22,8)}$ = 3.27 P = 0.044).

2.1.1. *N400 component*

ANOVA including an electrode factor ($F_{(11,319)}$ = 9.4 P < 0.000) revealed a maximum negative amplitude at anterior compared with posterior sites (F3 = 3.29 μV, T3 = 3.36 μV, Fz = 5.42 μV, Cz = 5.90 μV T4 = 5.9 5 μV compared with P4 = 10.89 μV, P3 = 9.22 μV, Pz = –9.16, C4 = 8.41 μV, respectively). Strongly salient utterances elicited significantly smaller N400 amplitudes than weakly salient utterances (8.30 μV vs. 4.92 μV respectively; $F_{(1,29)}$ = 17.20 P < 0.000).

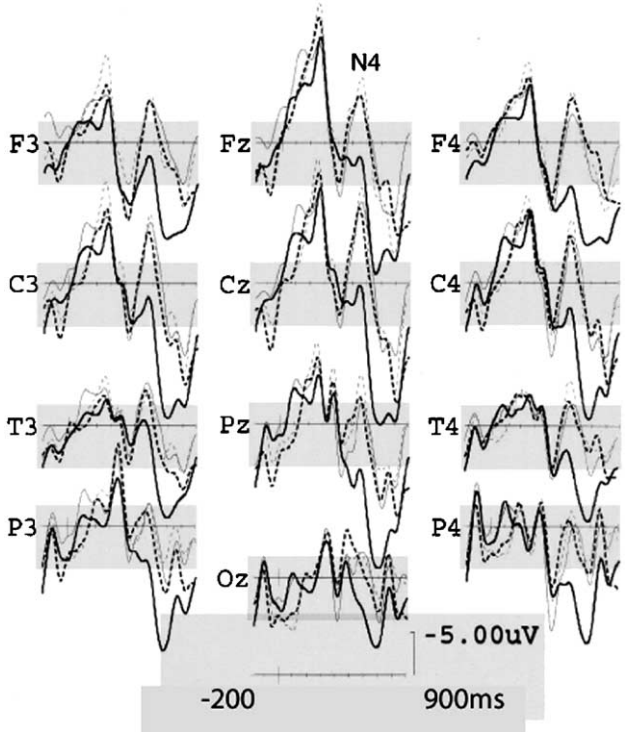

Fig. 3 – Evoked potentials in response to a target word related to figurative meaning preceded by a strongly salient idiomatic expression (solid thick line), a target word related to literal meaning preceded by a strongly salient idiomatic expression (solid thin line), a target word related to figurative meaning preceded by a weakly salient idiomatic expression (dashed thin line), and a target word related to literal meaning preceded by a weakly salient idiomatic expression (dashed thick line), with 12 electrodes between –200 ms and 900 ms after the beginning of the target presentation.

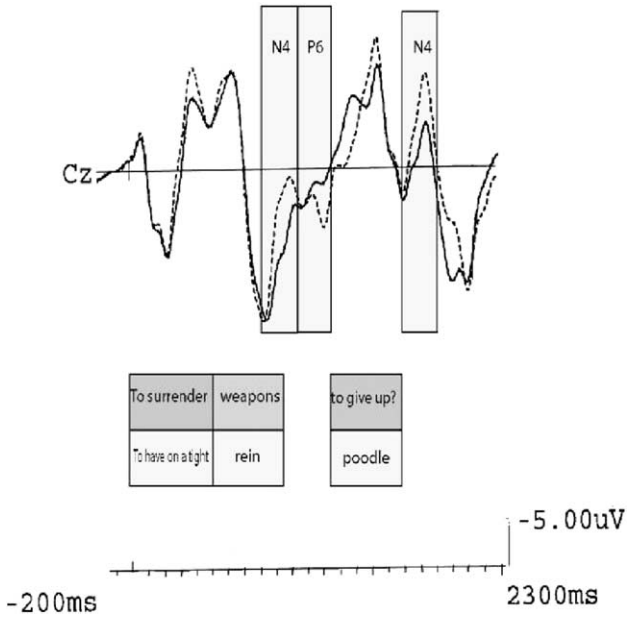

Fig. 4 – Pattern of evoked potentials over time recorded at Cz during presentation of strongly salient (solid line) and weakly salient (dashed line) idiomatic expressions.

2.1.2. *P600 component*

Analysis of electrode effect ($F_{(11,319)}$ = 3.02 P = 0.001) revealed that amplitude in the P600 epoch reached maximum in the left hemisphere and at centroparietal sites (Pz = 3.95 μV, P3 = 3.40 μV, Cz = 3.15 μV and C3 = 2.98 μV, respectively). Finally, the saliency effect indicated a significant decrease in P600 amplitude on strongly compared with weakly salient utterances ($F_{(11,319)}$ = 2.09, P = 0.021). Table 1 summarizes these results.

2.2. *Target words*

2.2.1. *Behavioral data*

The results indicated a significant (Rao $R_{(2,28)}$ = 79.3 P < 0.000) saliency × target type interaction for percentage of correct responses (CR%) and response times (RT). Table 2 summarizes the main results of data analyses. Post-hoc tests revealed that CR% was greater for figurative (90%) than literal targets (63%; P < 0.000) in strongly salient situations. CR% was greater for literal (83%) than figurative targets (59%; P < 0.000) in weakly salient situations.

Fig. 2 illustrates saliency by target type interactions according to RT. Post-hoc tests indicated that figurative targets were responded to faster (827 ms) than literal targets (879 ms; P = 0.019) in strongly salient condition. Literal targets were responded to faster (835 ms) than figurative targets (923 ms; P < 0.000) in weakly salient condition.

Table 3 – Population characteristics

	Mean	Standard deviation	Range
Age	30.53	8.41	21–50
Education level (years of study)	14.10	2.99	9–20
Binois–Pichot	27.97	3.89	19–35

Binois–Pichot is a scale of measurement of verbal IQ.

Moreover, the difference between the last two conditions was significant (RT = 923 ms vs. RT = 879 ms $P = 0.047$).

2.2.2. *ERP data*

As shown in Fig. 3, we also observed a significant saliency × target type interaction on N400. As for behavioral data, only figurative target/strongly salient idiom variables presented a decrease in N400 amplitude ($F_{(1,29)} = 7.76$ $P = 0.010$).

As shown in Fig. 4 waveforms, we observed that N400 amplitude was at its maximum for the T3 site (0.43 μV; $F_{(5,145)} = 5.40$ $P < 0.000$) compared with T4 (2.11 μV), Pz (3.80 μV), and P4 (3.87 μV) sites, respectively.

3. Discussion

The aim of this study was to compare results predicted by Giora's Graded Salience model with those predicted by the configurational model of Cacciari and Tabossi. We used differences in psychophysiological (amplitude of evoked potentials) and behavioral (% correct responses and reaction times) variables in a semantic judgment task, focusing on the existence of a semantic relationship between familiar and predictable idioms with two plausible meanings and a target word. According to Giora's model, only the salient meaning is activated by highly salient idioms, whereas, in the case of weakly salient idioms, both plausible meanings are activated. These predictions would be supported by smaller amplitudes of the N400 and P600 components for the last word and the target word that is compatible with the salient meaning of highly salient idioms compared with those of the weakly salient idioms. They would also be supported by longer RTs (reaction times) and a smaller number of correct responses to target words preceded by weakly salient idioms compared to those compatible with the salient meaning of highly salient idioms. In contrast, in the case of familiar, predictable, and compositional idioms, the configurational model postulates that the two meanings would be immediately available, as they are in the case of delayed selection of meaning for polysemous terms (see Introduction). In the latter case, we should not see any difference between the amplitude of the N400 for the last word of the idiom and the target word or the percentage of correct responses or RTs to the target words which follow highly or weakly salient idioms.

In our results, the amplitudes of N400 and P600 were smaller for the last word of highly salient idioms. Similarly, the amplitudes of N400 for target words compatible with salient meanings of salient idioms were smaller than those for the three other conditions. Finally, the percentage of correct responses was significantly greater in this experimental condition. All these findings support the predictions of Giora's model. Changes in RT in the semantic judgment tasks were also compatible with this model. The shortest RTs were for responses to figurative targets preceded by a highly salient idiom and for responses to literal targets preceded by a weakly salient idiom. In fact, there was no statistical difference between these two kinds of responses. These results show that the fastest RTs were to the salient meanings only—the idiomatic meaning of high-salience idioms and the literal meanings of low-salience idioms. This is entirely consistent with Giora's Graded Salience Hypothesis (GHS).

Do the ERPs findings corroborate these results? In the case of the idiomatic meaning of high-salience idioms, the N400 amplitude is statically smaller for the idiomatic than for the literal meaning. This result corroborates the behavioral results and supports the GHS. But, this statistical effect is not found for the comparison between literal and idiomatic meanings of low-salience idioms despite slight amplitude difference depicted in Fig. 3 (see difference between medium (literal) and thin (idiomatic) dashed line). Our results show that the behavioral data corresponded well to the ERP levels predicted by the GHS, but this was less obvious in the case of the psychophysiological results relating to weakly salient idiom. We suggest that the latter indicate only post-lexical phenomena of integration related to the ordered access of the two plausible meanings. The activation and integration of these meanings did not result in dissonance. Rather, these plausible meanings (the literal and the figurative) were perceived as closely related to each other and were therefore indistinguishable. Despite this disagreement, the major tenets of the GSH, namely, that it is not figurativeness or literalness that matters but salience, is supported by our data.

In a lexical decision study in which the trigger word might be polysemous, Chwilla et al. also obtained contradictory results

Table 4 – Linguistic characteristics of idiomatic expression

	Length of expression (*n* of words)	Length of first word (*n* of letters)	Length of last word (*n* of letters)	Frequency of last word	Latent semantic analysis between first and last word	Cloze probability
Strongly salient idiomatic expression	3.15 (0.37)	8.80 (1.20)	5.65 (1.66)	13,518 (19,994)	0.341 (0.143)	0.59 (0.29)
Weakly salient idiomatic expression	3.15 (0.37)	9.40 (1.76)	5.50 (1.67)	23,501 (28,137)	0.368 (0.162)	0.46 (0.26)
Idiom fillers	3.59 (0.94)	10.36 (3.09)	5.93 (1.81)	31,419 (23,207)	0.331 (0.158)	0.48 (0.36)
Literal fillers	3.34 (0.58)	10.54 (2.33)	6.12 (1.86)	15,991 (23,069)	0.396 (0.141)	0.57 (0.31)

Table 5 – Characteristics of experimental expression

	Familiarity	Compositionality	Literality	Cloze probability	% Idiomaticity	Absolute idiomaticity
Strongly salient idiomatic expression	90.35 (4.95)	64.55 (7.78)	67.15 (10.10)	58.80 (28.89)	78.20 (14.76)	89.85 (6.09)
Weakly salient idiomatic expression	88.85 (6.63)	70.75 (9.08)	75.35 (11.28)	45.55 (26.08)	15.70 (10.99)	17.10 (11.34)

between electrophysiological and behavioral findings. They proposed the following interpretation: "The opposite pattern, that is an N400 effect in the absence of an RT effect, will be interpreted as reflecting unique integration processes disconnected from specific decision and response-related processes…. This led to the proposal that N400 reflects a more global integration process that, in contrast to the integration process tapped by RT, can operate independently from decision and response-related processes" (Chwilla and Kolk, 2003).

Comparison of our results with those in the literature presented in the Introduction section is not easy because our aim was not (as in other studies) to compare different levels of N400 induced by literal and metaphoric items. However, as shown in Fig. 3, we had a control condition comprising only familiar, compositional, and highly predictable literal phrases and thus comparable with our materials in terms of these characteristics. This figure shows clearly that the curve for control literal expressions was hardly different from the curve of highly salient expressions (electrodes T3, T4, Cz). Increases in amplitude in the temporal window of N400 for the idiom condition were only present for weakly salient idioms. The salience of control literal expressions was 92.29%, attesting to the strength of the salience of these expressions. Our highly salient idioms and control literal phrases were comparable in terms of salience and, not surprisingly, in terms of amplitude of N400. As mentioned above, the findings in the literature generally involved new metaphoric expressions. In this case, showing that the amplitude of N400 is greater in the metaphor than in the literal condition might then be interpreted as an effect of weak salience of the metaphoric expression compared to the literal expression tested, rather than as an effect of figurativeness. Referring to Pynte's study, Giora remarked (Giora, 2002): "It is hard, however, to draw conclusions from their study as to how salience (or familiarity) affects processing since, in fact, they only used familiar metaphor vehicles (e.g. lions). Consequently, their 'familiar' (Those fighters are lions) and 'unfamiliar' (Those apprentices are lions) metaphoric utterances differed only in aptness, i.e., in how similar/relevant their vehicle is to their topic. Out of context, 'familiar' vs. 'unfamiliar' items did not differ (Experiment 2). However, when the targets (lions) were incompatible with prior context ("They are not idiotic: Those fighters are lions") they elicited larger N400 amplitude than when coherent with context ("They are not cowardly: Those apprentices are lions") (Experiments 3 and 4). At best, these endings can be taken to suggest that (relatively) familiar metaphors involve their salient metaphoric meaning upon encounter, regardless of either context or aptness". We too suspect that, in fact, all of Pynte's results can be interpreted along the lines of Giora's model, suggesting that the amplitudes of N400 were invoked by low aptness rather than by figurativeness.

As for the findings in Coulson et al.'s study, they might be a result of the items having different degrees of salience, with literals being highly salient, metaphors being of low salience, and literal mappings being of intermediate salience (see also Giora, 2003, chapter 5). In our own study, idiomatic controls ranged between weak salience (15.7%), intermediate salience (45%), and high salience (78.2%). We found larger P600 for weakly salient expressions (as those found for the metaphoric items in Coulson's study) compared with highly salient expressions (as those found for the literals in Coulson's study).

Tartter et al. experimental materials were also different from ours in two respects: we used very familiar idioms, and they used new metaphors; our materials were equivalent in terms of cloze probability and frequency of use for all the experimental conditions, and their metaphors were of lower cloze probability and frequency of use than those of their literal phrases. They attributed the differences that they found between the two conditions to the influence of the frequency of use in an early window (200–300 ms) generally interpreted as that of N200. On the other hand, they found no significant difference between these two conditions in the N400 window. We studied and statistically analyzed the amplitude of N200 for the last word of the idiomatic expressions in our results and did not find any difference between the two conditions. This agrees with the GSH but is not in accord with the results of Tartter et al. However, given that their metaphors were of lower cloze probability and frequency of use than their literal phrases, the N200 could reflect their lower salience compared to their literals. Hence, the differences. Still, how can we explain their lack of N400 effect? It might be the case that the difference in degree of salience between the two types of materials (metaphors and literals) was not that great and could be resolved in the N400 window.

Interesting in this respect is a metaphor interference task used by Kamezski (Giora, 2002). Kamezski asked his subjects to read a series of words and to decide on the reliability of

Table 6 – Formal linguistic characteristics of target words

	Word length		Word frequency	
	List 1	List 2	List 1	List 2
Related target	6.83* (1.96)	6.81 (1.74)	9022 (14,864)	7475 (11,384)
Non-related target	6.38* (1.62)		6565 (7575)	

* Significance $P < 0.01$.

Table 7 – Latent semantic analysis

Latent semantic analysis	First word of expression—figurative target word	First word of expression—literal target word	Final word of expression—figurative target word	Final word of expression—literal target word
Strongly salient idiomatic expression	0.38 (0.19)	0.30 (0.15)	0.16 (0.11)	0.20 (0.15)
Weakly salient idiomatic expression	0.38 (0.11)	0.30 (0.13)	0.15 (0.10)	0.22 (0.19)

the literal sense. He presented literal phrases, relatively new metaphoric phrases, and absurd phrases. His results revealed differences between the processing of metaphors in a group of participants of high and low IQ. The high IQ participants could directly access the metaphoric meanings of idiomatic expressions. However, the low IQ participants accessed the literal meanings of the same expressions before activating the metaphoric meaning. The authors interpreted their results within the framework of the constraints satisfaction model (Katz and Ferretti, 2001) and the predication model of Kintsch (2001). Earlier research into individual differences also supported an ordered access view for low-span individuals. Miyake et al. (1994) found that, for low-span readers, difficulty in ambiguity resolution varied with the degree of salience of the various interpretations. However, high-span individuals could retain multiple interpretations in neutral context, even though one of the interpretations was more frequent/salient. It is possible that high-span individuals can maintain multiple meanings of different salience simultaneously.

Our results and those published by others lead us to think that many of the findings in the literature regarding the amplitude of N400 might be effected by the salience of the different types of linguistic materials which the subjects had to process, independently of the figurative or literal nature of the material. Although the question has not been definitively answered, it could still be hypothesized that, in addition to its agreed upon sensitivities, N400 can reflect a late mechanism of semantic integration (Holcomb, 1993; Gunter et al., 1995; Salisbury, 2004) associated with the size of the verbal working memory and with frontal mechanisms (see also Kamezski et al.).

4. Experimental procedure

4.1. *Subjects*

Thirty subjects (including 17 women) were recruited by advertisement and were paid to take part in the study. All were right-handed except for two. All subjects were native speakers of French and had normal or corrected-to-normal sight. Table 3 summarizes their characteristics.

4.2. *Stimuli*

The materials consisted of 240 sentences followed by a target word which was sometimes semantically related and sometimes not. Of these sentences, 120 were French idioms including 40 experimental idioms extracted from Denhière and Verstiggel's (1997) idiomatic database. Half of them were characterized by strong idiomatic saliency. In an earlier study, saliency was evaluated by groups of 30 to 60 subjects who were asked to write down the first word that came to mind after having read each expression. The other 120 expressions were fillers of only literal significance. The strengths of the semantic links between the first and last words of each expression and between the target words and last words of the idioms were evaluated with "Latent Semantic Analysis" software (LSA) (Landauer et al., 1997) and with "Français-Livre" semantic space (Denhière and Pariollaud, 2000). Table 4 summarizes the formal characteristics of these materials and shows no significant difference between these features. Word lengths at the beginnings of sentences were

Table 8 – Example of stimulus utterance

	n	Utterance onset	Utterance final word	Target	
				Related	Non-related
Strongly salient idiom-related target word	20	Rendre les (to surrender)	Armes (weapons)	Abandonner (to give up)	Déposer (to put down)
Weakly salient idiom-related target word	20	Enfoncer le (to hammer it)	Clou (home)	Insister (to insist)	Fixer (to fix)
Idiomatic filler-related target word	40	Balancer (to squeal on)	Quelqu'un someone	Dénoncer (to denounce)	
Idiomatic filler-non related target word	40	Cirer les (To lick someone's)	Bottes (boots)		Éloigner (to remove)
Literal filler-related target word	40	Tondre la (to mow the)	Pelouse lawn	Jardiner (to garden)	
Literal filler-non related target word	80	Mettre la (to set the)	Table table		Vibrer (To vibrate)

n: number of expressions.

significantly shorter for the experimental materials compared to the fillers (9.1 vs. 10.46, $P = 0.03$). Moreover, word frequency was significantly greater (15991 vs. 31419, $P > 0.000$) and latent semantic similarities significantly smaller (331 vs. 396, $P = 0.003$) for the idioms compared to fillers.

Experimental idioms were comparable in terms of familiarity, predictability, and latent semantic similarity between strong and weak saliency. On the other hand, strongly salient idioms were low on literality and compositionality (these two variables being strongly correlated) than weakly salient idioms ($P < 0.03$). Our experimental idioms can thus be regarded as highly familiar and predictable, with low salient idioms high on literality and compositionality. Saliency was obviously ($P > 0.000$) greater for strongly compared to weakly salient idioms. Table 5 summarizes these characteristics.

4.2.1. *Target words*

Of 320 target words, 200 were related to the overall meaning of the preceding expression. Of 80 fillers, 40 targets were related to the meaning of the preceding idiom. Half of all target words preceding experimental idioms were related to the figurative meaning and the others to the literal meaning. To control for semantic priming between the last word of the expression and the target word, one-way ANOVA was conducted on latent semantic similarity with idioms (strong vs. weak) and targets (figurative vs. literal) as factors. No significant effect was found. Two experimental lists were constructed. Items from each list were presented in fixed order following randomization. All subjects saw the same set of sentences, but only half received the same figuratively and literally related targets (i.e., figuratively related targets that appeared in one list appeared in literal conditions in the other and vice versa (see Table 6)).

Table 7 presents the LSA (Latent Semantic Analysis) between the first and last words of the expression and analysis of relationship between last and target words.

In summary, of 120 idioms with two meanings (figurative and literal) and 120 literal sentences with only one meaning, 40 experimental idioms (20 strongly and 20 weakly salient) made up the experimental set for analysis. Moreover, 120 of 240 targets were related to the meaning of the previous bearing in mind that all experimental targets were related to the meanings of the idioms. Table 8 presents the set of expressions and target words.

4.3. *Procedure*

All subjects were informed of the general aims of the study and submitted a written consent. They were randomly assigned to one of two experimental lists. They were seated comfortably, about 80 cm in front of a computer screen. The task was a semantic priming paradigm in which subjects had to decide whether or not the target was semantically related to the meaning of the utterance. They responded by pressing a button on the mouse with their writing hand. The left button of the mouse corresponded to “yes” and the right to “no”. The participants were allowed to practice with two training lists in a training session that preceded the experiment. The entire recording period lasted approximately 20 min. Onset of the expression was presented in black lower-case letters on the white computer screen for 550 ms, and the last word was displayed for 450 ms. The screen then went blank and had remained totally white for 300 ms before the target word followed by a question mark was displayed for 450 ms. A window of 1850 ms was provided for the subject’s response recording from target onset. There was a 3600 ms interval between each item and the next (IIS).

4.3.1. *ERP recording and analysis*

EEGs were recorded using 12 electrodes arranged on each participant’s scalp in accordance with international convention: three electrodes in the frontal region (F3, Fz, and F4), three in the central region (C3, Cz, and C4), three in the parietal region (P3, Pz, and P4), two in the temporal region (T3 on the left and T4 on the right), and one in the occipital area (Oz). Electrodes were referred to linked earlobes. Four electrodes were used to record an electro-oculogram (EOG): two opposite the external canthi and one above and one below the eye. All impedances were kept below 2 kΩ. EEGs were recorded continuously using the InstEP system, digitalized at a frequency of 256 points per second, and eye movements were corrected off-line using an automatic program. Finally, data were digitally filtered at a bandwidth of 0.08 to 12 Hz.

Fig. 4 illustrates ERPs at Cz site from onset of the expression to target presentation. Strongly salient idioms are in solid line, and weakly salient are in dashed line.

All 30 subjects included in this study completed 20 trials without artifacts in each situation. Five subjects who did not satisfy this criterion were eliminated.

Based on visual inspection of the grand average waveform, mean evoked potential amplitude following the last word of the expression was computed for 2 time windows of 870 to 1050 ms and 1050 to 1250 ms after the 200-ms prestimulus baseline.

ERPs elicited by target words were analyzed within 1750–1950 ms windows.

Behavioral data (response times and error percentages) were analyzed separately by repeated-measure ANOVA. Significant interactions were analyzed using post-hoc tests.

We first carried out repeated-measure Multivariate Analyses of Variance (MANOVAs) on ERP amplitudes for the last word to compare saliency (strong, weak) and electrode effects (12 sites) in early and late windows. ANOVAs for the highest amplitude electrode sites were conducted to reveal saliency effects (strong and weak).

We then carried out repeated-measure MANOVAs on ERP amplitudes for the target word comparing saliency (strong, weak), target (figurative, literal) and electrode effects (12 sites) in early and late windows. Main effects and interactions were analyzed at the largest amplitude electrode sites by ANOVAs.

REFERENCES

Bentin, S., McCarthy, G., et al., 1985. Event-related potentials, lexical decision and semantic priming*1. Electroencephalogr. Clin. Neurophysiol. 60 (4), 343–355.

Blasko, D.G., Connine, C.M., 1993. Effects of familiarity and aptness on metaphor processing*1. J. Exper. Psychol., Learn., Mem., Cogn. 19 (2), 295–308.

Bonnaud, V., Gil, R., et al., 2002. Metaphorical and non-metaphorical links: a behavioral and ERP study in young and elderly adults. Neurophysiol. Clin./Clin. Neurophysiol. 32 (4), 258–268.

Cacciari, C., Tabossi, P., 1988. The comprehension of idioms. J. Mem. Lang. 27, 668–683.

Chwilla, D.J., Kolk, H.H.J., 2003. Event-related potential and reaction time evidence for inhibition between alternative meanings of ambiguous words. Brain Lang. 86 (2), 167–192.

Coulson, S., Van Petten, C., 2002. Conceptual integration and metaphor: an event-related potential study. Mem. Cogn. 30 (6), 958–968.

Coulson, S., King, J.W., et al., 1998. ERPs and domain specificity: beating a straw horse. Lang. Cogn. Processes 13 (6), 653–672.

Denhière, G., Pariollaud, F., 2000. Using LSA for scoring story recall. Latent Semantic Analysis and language processing Workshop. Aix-en-Provence 17–18.

Denhière, G., Verstiggel, J.-C., 1997. Le traitement cognitif des expressions idiomatiques: activités automatiques et délibérées. La locution: entre le lexique, syntaxe et

pragmatique—Identification en corpus, traitement, apprentissage. P. L. M. F. P. E. In P. Fiala. Paris, Klincksiek: 119–148.
Fauconnier, G., Turner, M., 1998. Conceptual integration networks. Cogn. Sci. 22, 133–187.
Federmeier, K.D., Kutas, M., 2001. Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. J. Exper. Psychol., Learn., Mem., Cogn. 27 (1), 202–224.
Fodor, J., 1983. The Modularity of Mind. MIT Press, Cambridge.
Friederici, A.D., Von Cramon, D.Y., et al., 1999. Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. Brain 122 (6), 1033–1047.
Giora, R., 1997. Understanding figurative and literal language: the Graded Salience Hypothesis. Cogn. Linguist. 7 (1), 183–206.
Giora, R., 2002. Literal vs. figurative language: different or equal? J. Pragmat. 34 (4), 487–506.
Giora, R., 2003. On Our Mind: Salience Context and Figurative Language. Oxford Univ. Press, New York.
Giora, R., Fein, O., 1999. On understanding familiar and less-familiar figurative language. J. Pragmat. 31 (12), 1601–1618.
Glucksberg, S., 2003. The psycholinguistics of metaphor. Trends Cogn. Sci. 7 (2), 92–96.
Gunter, T.C., Jackson, J.L., et al., 1995. Language, memory and aging: an electrophysiological exploration of the N400 during reading and memory-demanding sentences. Psychophysiology 32, 215–229.
Gunter, T.C., Stowe, L.A., et al., 1997. When syntax meets semantics. Psychophysiology 34 (6), 660–676.
Hagoort, P., 2003. How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. NeuroImage 20 (Supplement 1), S18–S29.
Harris, T., Wexler, K., et al., 2000. An ERP investigation of binding and coreference. Brain Lang. 75 (3), 313–346.
Holcomb, P.J., 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. Psychophysiology 30, 47–61.
Kaan, E., Harris, A., et al., 2000. The P600 as an index of syntactic integration difficulty. Lang. Cogn. Processes 15 (2), 159–201.
Katz, A., Ferretti, T., 2001. Moment-by-moment reading of proverbs in literal and nonliteral contexts. Metaphor Symb. 16, 193–221.
Kazmerski, V.A., Blasko, D.G., et al., 2003. ERP and behavioral evidence of individual differences in metaphor comprehension. Mem. Cogn. 31 (5), 673–689.
Kim, A., Osterhout, L., 2005. The independence of combinatory semantic processing: evidence from event-related potentials. J. Mem. Lang. 52 (2), 205–225.
Kintsch, W., 2001. Predication. Cogn. Sci. 25, 173–202.
Kolk, H.H.J., Chwilla, D.J., et al., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85 (1), 1–36.
Kotz, S.A., Friederici, A.D., 2003. Electrophysiology of normal and pathological language processing. J. Neurolinguist. 16 (1), 43–58.
Kounios, J., Holcomb, P.J., 1994. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. J. Exper. Psychol., Learn, Mem., Cogn. 20 (4), 804–823.
Kuperberg, G.R., Sitnikova, T., et al., 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cogn. Brain Res. 17 (1), 117–129.
Kutas, M., 1985. ERP comparisons of the effects of single word and sentence contexts on word processing. Psychophysiology 22, 575–576.
Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4 (12), 463–470.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.
Kutas, M., Iragui, V., 1998. The N400 in a semantic categorization task across 6 decades. Electroencephalogr. Clin. Neurophysiol. 108 (5), 456–471.
Kutas, M., Lindamood, S.A., et al., 1984. Word expectancy and event-related potentials during sentence processing. In: Kornblum, S., Requin, J. (Eds.), Preparatory States and Processes. Erlbaum Press, Hillsdale, USA, pp. 217–238.
Landauer, T.K., et al., 1997. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol. Rev. 104, 211–240.
Marquer, P., 1994. La compréhension des expressions idiomatiques. L'Année Psychol. 94, 625–656.
Matzke, M., Mai, H., et al., 2002. The costs of freedom: an ERP-study of non-canonical sentences. Clin. Neurophysiol. 113 (6), 844–852.
Miyake, A., Just, M.A., et al., 1994. Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. J. Mem. Lang. 33, 175–202.
Mullet, V., Denhière, G., 1997. Accès au lexique et ambiguïtés lexicales nominales: effet de la polarité des homographes et de la nature du contexte. In: François, J., Denhiere, G. (Eds.), Sémantique Linguistique et Psychologie Cognitive. Aspects Théoriques et Expérimentaux. Presses Universitaires de Grenoble, Grenoble, pp. 51–74.
Osterhout, L., Allen, M.D., et al., 2002. Brain potentials elicited by prose-embedded linguistic anomalies. Mem. Cogn. 30 (8), 1304–1312.
Passerieux, C., Iakimova, G., et al., 2000. Metaphor comprehension in patients with schizophrenia: an event-related potential study. Schizophr. Res. 41 (1), 151.
Pynte, J., Besson, M., et al., 1996. The time-course of metaphor comprehension: an event-related potential study. Brain Lang. 55 (3), 293–316.
Rat, M., 1999. Dictionnaire Des Expressions et Des Locutions Traditionnelles. Larousse, Paris.
Salisbury, D.F., 2004. Semantic memory and verbal working memory correlates of N400 to subordinate homographs. Brain Cogn. 55, 396–399.
Tartter, V.C., Gomes, H., et al., 2002. Novel metaphors appear anomalous at least momentarily: evidence from N400. Brain Lang. 80 (3), 488–509.
Thérouanne, P., Denhière, G., 2002. Effet d'un contexte lexical sur l'accés à la signification d'homographes. L'Année Psychol. 102, 31–63.
Titone, D.A., Connine, C.M., 1994a. Descriptive norms for 171 Idiomatic expressions: familiarity, compositionality, predictability, and literality. Metaphor Symb. Act. 9 (4), 247–270.
Titone, D.A., Connine, C.M., 1994b. Comprehension of idiomatic expressions effects of predictability and literality. J. Exper. Psychol., Learn., Mem., Cogn. 20 (5), 1126–1138.
Trueswell, J.C., 1996. The role of lexical frequency in syntactic ambiguity resolution. J. Mem. Lang. 35 (4), 566–585.
Trueswell, J.C., Kim, A.E., 1998. How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. J. Mem. Lang. 39 (1), 102–123.
Trueswell, J.C., Tanenhaus, M.K., et al., 1994. Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. J. Mem. Lang. 33 (3), 285–318.
van Herten, M., Kolk, H.H.J., et al., 2005. An ERP study of P600 effects elicited by semantic anomalies. Cogn. Brain Res. 22 (2), 241–255.
Van Petten, C., 1995. Words and sentences: event-related brain potential measures. Psychophysiology 32, 511–525.

Van Petten, C., Coulson, S., et al., 1999. Time course of word identification and semantic integration in spoken language. J. Exper. Psychol., Learn., Mem., Cogn. 25, 394–417.

West, W.C., Holcomb, P.J., 2000. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. J. Cogn. Neurosci. 12 (6), 1024–1037.

West, W.C., Holcomb, P.J., 2002. Event-related potentials during discourse-level semantic integration of complex pictures. Cogn. Brain Res. 13 (3), 363–375.

Wicha, N.Y.Y., Bates, E.A., et al., 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. Neurosci. Lett. l346 (3), 165–168.