Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Memory and Language 54 (2006) 161–184
Journal of Memory and Language
www.elsevier.com/locate/jml

Lexical access during the production of idiomatic phrases

Simone A. Sprenger [a,*], Willem J.M. Levelt [a], Gerard Kempen [a,b]

[a] *Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands*
[b] *Cognitive Psychology Unit, Department of Psychology, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands*

Received 23 July 2004; revision received 1 November 2005

Abstract

In three experiments we test the assumption that idioms have their own lexical entry, which is linked to its constituent lemmas (Cutting & Bock, 1997). Speakers produced idioms or literal phrases (Experiment 1), completed idioms (Experiment 2), or switched between idiom completion and naming (Experiment 3). The results of Experiment 1 show that identity priming speeds up idiom production more effectively than literal phrase production, indicating a hybrid representation of idioms. In Experiment 2, we find effects of both phonological and semantic priming. Thus, elements of an idiom can not only be primed via their wordform, but also via the conceptual level. The results of Experiment 3 show that preparing the last word of an idiom primes naming of both phonologically and semantically related targets, indicating that literal word meanings become active during idiom production. The results are discussed within the framework of the hybrid model of idiom representation.

Keywords: Idioms; Fixed expressions; Sentence production; Priming

Introduction

In everyday conversations speakers rely heavily on preformatted utterances. They talk about the skeletons in their neighbour's closet, about the new position they are looking forward to, and they bet their shirt that their colleague's new car cost an arm and a leg. Such utterances are not new creations of the speakers themselves. Instead, they are Fixed Expressions (FEs) that belong to the conventional repertoire of the native speaker of a language. Both meaning and form of these utterances are standardized, often allowing for only minimal variation.

Fixed Expressions are *phrasal units*, and they exist in many varieties (e.g., phrasal verbs, restricted collocations, idiomatic expressions, and sayings and proverbs). Idiomatic expressions or *idioms* are a particularly interesting variant of FEs, because their meaning is partly or completely non-compositional. That is, the relationship between the meanings of the words that make up the idiom and the idiom as a whole is at best indirect, if there is any relation at all. This is most obvious in idioms that are opaque, like, for example, *kick the bucket*. The literal meaning of this phrase does not suggest its figurative meaning *to die*. Still, native speakers of English *know* that *last night Jim kicked the bucket* means that Jim is dead. Of course, a literal reading is not excluded; in a context where there has been a discussion about people kicking buckets, the literal reading will be preferred.

Typically, idioms allow only few variations. Their words cannot generally be replaced or modified. For example, replacing *road* by *path* in *hit the road* yields a phrase that only has a literal interpretation and, at best,

* Corresponding author. Fax: +31 24 3521213.
E-mail address: simone.sprenger@mpi.nl (S.A. Sprenger).

0749-596X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2005.11.001

can be understood as a creative modification of the original idiom. The same holds for the insertion of a modifier as in *they hit the icy road* and for manipulations of the syntactic structure (*the road was hit by them*).

Still, idioms (or FEs in general) have hardly been addressed in standard accounts of language production, despite the fact that, from an empirical point of view, they are anything but exceptions. Jackendoff (1995) suggests that the number of FEs that speakers know (including names, titles, poems, and the like) and the number of single words in their vocabulary are at least of the same order of magnitude. He also argues that given their linguistic properties, the natural place to store FEs is the mental lexicon. This implies that estimates of the size of the (passive) mental lexicon (about 60,000 words; Miller, 1991) may have to be doubled. Even if only a portion of the FEs is actually part of the average speaker's *active* lexicon, clearly they are far from special: speakers use them quite frequently, which makes them an inherent feature of "native-like" language use (Pawley & Syder, 1983).

Incorporating idioms into the mental lexicon requires a theory of how they are stored, accessed, and processed. Much work has been done about the comprehension of idioms, but only few studies have been devoted to their *production*.

Both the non-literalness and the syntactic constraints of idioms show that we are dealing with special units of linguistic processing. Unlike literal phrases, idioms are not constructed on-line during speaking, suggesting that they must be retrieved from long-term memory. Their mental representation must comprise at least the set of words, their syntactic idiosyncrasies, and their figurative interpretation. The present study has been designed to further develop our understanding of how idioms are stored and produced.

After discussing the literature on idiom comprehension and production, we will present the findings from three experiments that explore the production of Dutch idiomatic expressions. We focus on the mental representation of idioms in the speaker's lexicon and the relationship between the idiom as a whole and the words it contains. We will argue that despite their special linguistic features, idioms are not exceptional from the point of view of the speaker and that they can be incorporated into standard models of language production.

Idiom comprehension

Psycholinguistic studies of idiom comprehension have addressed the questions of how idiomatic expressions are identified as such, how listeners derive the meaning of an idiomatic expression, and what role literal word meanings play in that process. Though the results of these studies cannot tell us much about the processes that come into play when idioms are *produced*, they can clarify how idioms are stored and represented in the mental lexicon, given the assumption that the same network of abstract concepts and linguistic representations is used for both language comprehension and production (e.g., Kempen & Harbusch, 2002; Roelofs, 2003).

Early accounts of idiom comprehension proposed a word-like representation of idioms in the mental lexicon (e.g., Bobrow & Bell, 1973; Swinney & Cutler, 1979), suggesting that the single words that make up the phrase and the semantic and syntactic information they contain do not play a role for the idiom as a unit. However, several observations argue against such a representation: There is correct stress assignment in idioms and many of them show (restricted) syntactic flexibility (Katz, 1973). Moreover, Peterson, Burgess, Dell, and Eberhard (2001) demonstrated a syntactic priming effect for idiomatic phrases, independent of the degree of the structural flexibility of a given idiom. Furthermore, the word-like account precludes the possibility of parts of an idiom carrying part of the idiomatic meaning. However, idioms can have components that refer separately to the components of their figurative referents. Such idioms are defined as semantically decomposable (Nunberg, Sag, & Wasow, 1994). For example, in *break the ice*, *ice* refers to a "cold" social atmosphere and *break* to the process of changing it. Thus, in semantically decomposable idioms certain roles and relationships between the entities addressed in the idiom can be mapped onto their figurative counterparts. Based on this observation, Gibbs and Nayak (1989) point out that, in decompositional idioms, internal modifications only change part of the idiom's meaning. They assume that each component makes its own contribution to the figurative interpretation of the idiom as a whole.

A related question concerns the role of the literal meanings of the words of an idiom. For literal language, processing has been shown to be non-optional, that is, we cannot decide *not* to process linguistic information (e.g., Miller & Johnson-Laird, 1976). This suggests that the literal meanings of the words of an idiom also become active during idiom comprehension. However, some additional process must be involved that can discover the non-literal nature of the utterance and that precludes noticeable disturbance by the utterance's literal meaning.

Cacciari and colleagues have focused on this question. Cacciari and Tabossi (1988) showed that, in the absence of contextual cues to the idiomatic meaning of a phrase, the activation of the literal meaning of its last word (that had been ambiguous between a literal and an idiomatic interpretation up to this position) precedes the activation of the idiomatic word meaning by about 300 ms. In contrast, given an *idiomatic* context, both the literal and the idiomatic word meanings are available immediately upon presentation. Cacciari and Glucksberg (1991) acknowledge that active literal meanings do not have to play a *functional* role in idiom understanding. Nevertheless, their activity can be measured, that is, the comprehension system does not seem to switch to a completely different manner of processing when running into idioms.

With their *Configuration Hypothesis*, Cacciari and Tabossi (1988) propose a theoretical framework accounting for their findings. An idiomatic phrase is assumed to activate the same lexical items that would otherwise be involved in the comprehension of literal discourse. This process immediately yields the words' literal interpretation. Access to the idiomatic meaning of a phrase requires recognizing the phrase as a special configuration. This configuration emerges after some information that uniquely identifies the idiom as such (the idiom's *key*) has been processed. The interpretation of an idiomatic phrase is therefore literal until the configuration has been recognized. This theory clearly differs from the unitary approach referred to earlier, because each word is represented in the lexicon only in one form and need not be marked as literal or idiomatic (Cacciari & Tabossi, 1988). Thus, the Configuration Hypothesis stresses the compositional aspect of idioms. In addition, the theory accounts for the syntactic parsing of idioms. However, the authors do not specify how the syntactic constraints that are typical of idiomatic expressions are represented within the framework. Moreover, the definition of idiom key is unsatisfactory in that it does not enable its unambiguous identification in arbitrary idioms (but see also Tabossi & Zardon, 1993).

In sum, idiom comprehension suggests that a theory of idiom representation has to solve a paradox: how to account for the unitary nature of idioms, given the literal interpretation of the single words involved.

Idiom production

One must be cautious when generalizing from idiom comprehension theories to a theory of idiom *production*. It should be kept in mind that the speaker's situation is quite different from that of the listener. The process of speaking starts with the conceptual message and ends with an utterance that can be taken either literally or not. While the listener makes a decision about one or the other interpretation, there is no doubt on the part of the speaker about the message to be conveyed. Still, in the case of idioms, the compositional meaning of the words produced does not match that message (see Nooteboom, 1999, for a discussion of speech errors and monitoring in idioms). The message that underlies an idiom often cannot even be paraphrased satisfactorily. Idioms have their own characteristic conceptual conditions and it seems therefore perfectly straightforward to assume, with Levelt (1989), that idioms have their own entry on the level of lexical concepts (see also Flavell & Flavell, 1992).

Accordingly, the first question that arises is how the speaker handles this seeming contradiction. On the one hand, we must investigate what role the individual words of an idiom play in production and how they are activated. On the other hand, we must assume some unitary conceptual representation of idioms.

To our knowledge, Cutting and Bock (1997) conducted the first experimental study answering some of the questions about the storage of idiomatic expressions in the mental lexicon and their retrieval during production. They studied semantic and syntactic influences on experimentally elicited idiom blends. Participants read two simultaneously presented (idiomatic) phrases (e.g., *meet your maker* and *kick the bucket*) and then, after a delay of 2 s, produced one of them in response to a cue. This procedure was expected to give rise to competition between the phrases, thereby setting the stage for the production of spontaneous phrase blends.

In their first experiment, Cutting and Bock (1997) investigated the sensitivity of idiom blends to both the internal structure and the figurative meaning of the idioms involved. They found that identical figurative meanings of two competing idioms resulted in significantly longer production latencies. Moreover, idioms with the same syntactic structure were more likely to blend than idioms with different structures. When examining intra-idiom errors in more detail, they found that these errors follow a grammatical class constraint (see also Stemberger, 1982). The authors conclude that idioms are not produced as "frozen phrases," but instead are syntactically analyzed.

In their second experiment, Cutting and Bock (1997) showed that phrase pairs with the same meaning produced more blends than phrase pairs with different meanings, irrespective of whether they were idiomatic or not. Moreover, the grammatical class constraint held for both conditions, that is, it was blind to the (non-)idiomaticity of the blending phrases. The results are interpreted as evidence for the activity of literal word meanings during the production of idiomatic phrases.

In a third experiment, Cutting and Bock (1997) investigated the hypothesis proposed by Gibbs and Nayak (1989) that the lexical representation of semantically decomposable idioms is less rigidly specified and more susceptible to change than that of non-decomposable idioms. All idiom pairs presented shared both their syntactic structure and their figurative meaning, but differed in decompositionality (e.g., *shoot the breeze* and *chew the fat* as non-decomposable pair and *hold your tongue* and *button your lip* as decomposable pair). The error rates were the same for both kinds of pairs, that is, the (non-)decompositionality of an idiom was not mirrored in the production process. The authors conclude that the lexical representations of decomposable and non-decomposable idioms are the same when they enter into the production process.

Based on these findings, Cutting and Bock suggest a way of integrating idiom production into current models of language production (Dell, 1986; Levelt, 1989). They assume that each idiom has its own *lexical concept node*. Thus, idioms are represented as *unitary entities* on at least one processing level. The authors assume furthermore that one concept can activate multiple lexical concept nodes (including other idioms), as is the case in non-idiomatic phrase production. For example, the concept that activates the lexical-conceptual representation of *kick the bucket* is assumed to activate *meet your maker* as well. This may lead to competition and to semantic blends, as in *meet the bucket maker*.

In contrast, semantic decomposition is modeled by multiple concepts activating *one* lexical concept node. Thus, for example, the lexical concept *pop the question* (to propose marriage) is linked to both the concepts for *suddenly* and *to propose*. However, in contrast to Gibbs and Nayak's (1989) hypothesis, this representational difference has no effect on the syntactic flexibility of compositional and non-compositional idioms. Once the level of (lexical) concepts has been passed, processing decompositional and non-decompositional idioms does not differ anymore. Decomposition of idiomatic expressions is thus relegated to the conceptual, not the syntactic domain (see Fig. 1).

When an idiomatic lexical concept node has been activated, activation spreads in *two* directions: first, the lemmas that together constitute the idiom get activated. Second, activation spreads to *syntactic* information in the form of prefabricated phrasal frames. Accordingly, the model explains blending errors in syntactically similar idioms by means of shared phrasal frames. Cutting and Bock conclude that

"Idioms may be special in their relationships to nonlinguistic concepts, but they are not special in the way they are produced in normal language use." (p. 69)

In sum, Cutting and Bock (1997) subscribe to the view that, although idioms are stored as a whole on some level of processing, they cannot be word-like entries without internal structure. Thus, Cacciari and Tabossi's (1988) view on idiom comprehension (*Configuration Hypothesis*) is mirrored in speech production.

The common factor of these theories is their solution of the above-mentioned paradox: idioms can be both unitary in that they require their own lexical entry, and compositional, in that they make use of simple lemmas in the mental lexicon. These simple lemmas *can* be used within an idiomatic context, but they are not restricted to it.

The present evidence for such a "hybrid" account of idiom production is largely based on elicited speech error data (Cutting & Bock, 1997). Though speech errors are a valuable source for theories of language production, they cannot show that error-free production takes place along the same pathways. A theory of idiom representation therefore needs to be complemented with data that show the pathway of activation during normal speech production. We will present three experiments which tested the predictions of the hybrid account for error-free speech production with different reaction time paradigms. In addition to these experiments, we independently assessed the decomposability of the idioms

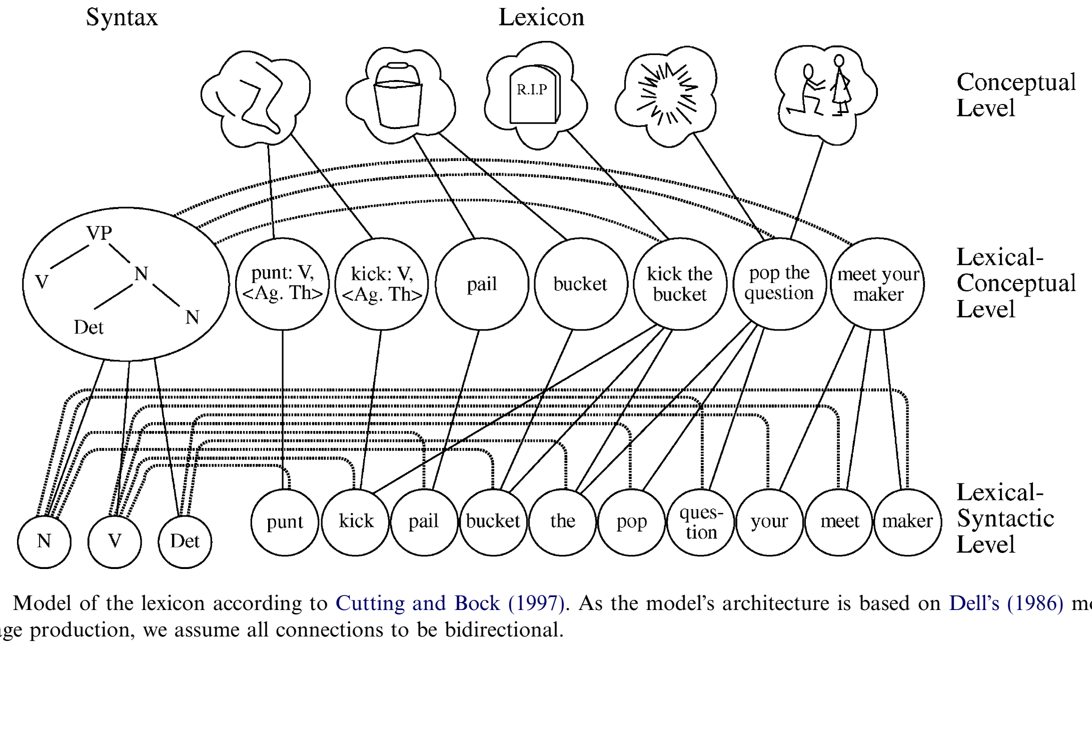

Fig. 1. Model of the lexicon according to Cutting and Bock (1997). As the model's architecture is based on Dell's (1986) model of language production, we assume all connections to be bidirectional.

that were used in this study. While Cutting and Bock (1997) did not find evidence for an influence of decomposability on speech error rates, we wanted to take into account the possibility that decomposability might affect speech onset latencies for idioms.

In the first experiment, we tested the two core assumptions of the hybrid account of idiom production: idioms are composed out of single words *and* they have their own representation in the mental lexicon that spreads activation to all its component parts. In the second and third experiments, we explored the consequences of this account for the network of semantic representations in the mental lexicon. Specifically, we investigated to what extent literal word meanings become active during idiom production.

The results of the decomposability rating, as well as a post hoc analysis of the data from all three experiments with Decomposability as a covariate, can be found in Appendix A.

Experiment 1

The first set of predictions that can be deduced from a hybrid account of idiom representation concerns the possibility of *priming* the simple lemmas that belong to a phrase. If simple lemmas involved in idiom production are indeed the same as those involved in compositional phrase production, it must be possible to prime these lemmas. Activating a lemma by means of an identity prime speeds up production (e.g., Glaser & Düngelhoff, 1984). Thus, priming *road* in *clean the road* by means of the word *road* itself can be expected to result in shorter production latencies, compared to a condition where the prime is unrelated to the target word. If our assumption that simple lemmas are involved in idiom production is correct, a similar effect of identity priming should be found for the production of *hit the road* as well. Therefore, we predict a significant main effect of prime type. In particular, we predict an effect of priming from identity primes (i.e., prime words that are identical to one of the words in the phrase), but not from control-primes that are phonologically and semantically unrelated to the to-be-produced utterance.

However, we do not expect the priming effect for idiomatic and literal phrases to be of the same magnitude. Instead, we predict a stronger effect from the identity prime in the case of idioms. Consider the case of *hit the road*. Hearing the word *road* should activate the lemma *road* and, if it is indeed connected with a common idiom representation, the lexical entry *hit the road* should be activated as well. Upon selection of this entry, further activation spreading will result in higher activation levels of all simple lemmas attached to the idiom, thus speeding up their selection. In our example, *hit* can be selected more easily, thereby affecting the production latencies for *hit the road*. A literal phrase like *clean the road* on the other hand cannot profit to the same amount from *road* being primed. Though priming of *road* should speed up production of the phrase involving that word to some extent, no benefit for the other lemmas belonging to the phrase is expected. The priming effect of *road* for *clean the road* should be smaller than that for *hit the road*, because no common lexical entry gets selected that binds the word *clean* to *road*. Their combination is transient and a consequence of conceptual decisions. In other words, we expect an interaction between the factors *Prime Type* (either related to one of the words of the phrase or unrelated) and *Idiomaticity* (literal vs. idiomatic phrases). If this interaction obtains, it argues for a connection in the mental lexicon between simple lemmas via a common idiom representation.

We tested these predictions in a cued-recall experiment. Participants produced idiomatic and literal phrases in response to a visually presented prompt word. Primes were presented auditorily and simultaneously with the prompt word presentation. They were either identical with the target or semantically and phonologically unrelated. Response time analyses were carried out in order to determine the effects of Priming and Idiomaticity on response latencies.

Method

Participants

Sixteen participants were tested, who were all undergraduate students of the University of Nijmegen and native speakers of Dutch. They were paid for their participation.

Materials

We constructed 16 item pairs on the basis of 16 idiomatic expressions, all of the same syntactic structure: [VP [PP Prep [NP art N]] V]. They were all judged by six native speakers to be well-known Dutch idiomatic phrases. All phrases were finite Dutch phrases, as for example

...viel buiten de boot (word-by-word translation: '...fell outside the boat,' i.e., the finite past tense form of an idiom meaning 'to be excluded from something').

Each idiomatic item was paired with a literal phrase that had the same syntactic form and the same noun as its idiomatic counterpart. A combination of an idiomatic and a literal phrase together yielded one item pair. Thus, for example, *...viel buiten de boot* and *...ging met de boot* (literally 'went with the boat,' i.e., took the boat) form two members (idiomatic and literal) of the same item pair.

Every item required a prompt word that could trigger the production of the phrase. For 13 item pairs, the prompt word was a common Dutch name (e.g., *Jan*). The same name was chosen for both items in an item pair (idiomatic and literal). For the remaining three item pairs, the prompts consisted of a short phrase (name [possessive s] [noun]), for example, *Jan's feestje*... 'Jan's party...' vs. *Jan's dochter*... 'Jan's daughter...,' which was different for the two versions of an item pair, but yet neutral with respect to the phrases' contents. Together, a name and a phrase always formed a complete sentence.

Some of the idioms required a direct object in their finite form (e.g., *Karin...hield* ***hem*** *op de hoogte*, 'Karin...kept **him** informed'). In these cases, the other item of an item pair was matched in length, whether or not it was necessary from a syntactic point of view. This was accomplished by means of the insertion of modifiers (*Karin...schrok* ***erg*** *van de hoogte*, 'Karin...was freightened by the height **a lot**'). The complete list of all items and the respective prompt words is given in Appendix B.

The identity prime for each item was its noun, which was therefore the same for the idiomatic and literal item of a pair. In addition, 16 unrelated prime words were retrieved from the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). They were frequency matched nouns that were semantically and phonologically unrelated to the phrases and their component words. A complete list of all primes is given in Appendix C.

All primes were spoken by the same female native speaker of Dutch and were recorded on DAT-tape in one session.

Procedure

Participants were tested individually and each session was recorded on DAT tape. The visual stimuli were presented on a computer screen, the acoustic stimuli via headphones. Responses were spoken into a microphone that was attached to a voice key, which in turn signaled the computer that a response had been initiated. The experiment was controlled by NESU (Nijmegen Experiment Set-Up). An experimental session included a preparatory learning phase and two experimental cued-recall blocks with a pause in between.

Learning

After reading the instructions, participants were presented with a list of eight names and associated phrases (half of them idiomatic and half of them literal). They were asked to memorize the phrases in such a way that they could produce the phrase fluently whenever they were presented with the names. When the participants indicated that they knew all phrases by heart, they were presented a list of names alone (in random order) and had to produce the appropriate phrases as quickly and fluently as possible. The production was judged by the experimenter. Any errors, dysfluencies or pauses led to a repetition of the learning phase and the rehearsal. Only when the participants succeeded in fluently producing all phrases, the first experimental block was started.

Cued recall

An experimental block consisted of the repeated presentation of eight previously learned prompt words and the production of their associated phrases by the speakers. After a fixation cross had appeared in the center of the screen, participants saw one of the prompt words. At the same time, a prime word was presented via the headphones. This prime was either identical to the noun of the to-be-produced phrase or unrelated. The participants' task was to react to the visually presented prompt word by producing the appropriate phrase as quickly as possible.

The responses triggered a voice key, signalling the production onset latency of the response. If the speaker failed to respond within 4200 ms, the computer automatically registered a missing response and a new trial was presented.

Each block consisted of 128 trials that were presented in pseudo-random order: there were at most two consecutive trials in the same condition (with condition defined as one of the four possible combinations of the variables *Priming*, two levels, and *Idiomaticity*, two levels). The minimum distance between two appearances of an item was three trials. Every first presentation of an item counted as a practice trial, thus serving to refresh the participant's memory of the items within the context of the experimental situation. The participants were instructed to react as quickly as possible, but in a fluent fashion and without making mistakes. They were also asked to reduce coughing, etc., as far as possible and to avoid unnecessary noises that would set off the voice key. They were informed that they would be recorded on audio tape.

Participants could pause between experimental blocks. A second learning set was presented after participants indicated that they were ready to continue. The procedure was identical to the first part of the experiment, except that new names and phrases had to be learned and produced.

Design

The design included two within-subject factors (Idiomaticity and Prime Type, with two levels each), yielding four experimental conditions. Every speaker saw one item out of each of the 16 item pairs, one half being idiomatic items, the other half literal. Both idiomatic and literal items were equally distributed over two experimental blocks. Every item was presented equally often with an identity prime as with an unrelated prime. The combination of an item with one of its primes was repeated eight times within a block, yielding 128 trials per block and 256 trials per subject. Since each participant only saw either the literal or the idiomatic item of an item pair, two different item lists were created. In addition, the order of block presentation was counterbalanced, yielding four different experimental lists. Every list was tested on four participants, who each received a different randomized version.

Analyses

For every participant every first measurement of an item was excluded from the data set. The DAT-tape recordings were checked for erroneous or missing responses and dysfluencies. A response was scored as erroneous if either the word order had been changed or if one or more words had been replaced. However, this rule did not hold for preposition or verb exchanges in idiomatic phrases if they reflected variants of the same idiom. For example, *brengt hem op de hoogte* (informs him) was considered equivalent with *houdt hem op de hoogte* (keeps him informed), if used consistently over the trials.

Reaction time data and error percentages were entered into separate analyses of variance with Idiomaticity and Prime Type as within-subject factors. Separate analyses were carried out with either Subjects or Items as random factors, yielding F_1 and F_2 statistics, respectively.

Results

Three percent of all data points were erroneous or dysfluent. An analysis of error percentages revealed no significant difference between the idiomatic and the literal phrases (on average 3.6% errors ($SD = 2.2$) for the idiomatic phrases vs. 2.6% errors ($SD = 1.6$) for the literal phrases; $t = -1.270$, $p = .223$, one-tailed test). The mean production latencies are shown in Table 1.

The results confirm our hypotheses. There is no main effect of Idiomaticity (F_1 and $F_2 < 0$). The main effect of Priming (average speech onset latencies of 906 ms in the unrelated condition vs. 820 ms in the related condition) is significant ($F_1(1,15) = 42.67$, $MS_e = 119{,}076$, $p < .001$, $F_2(1,15) = 46.58$, $MS_e = 117{,}242$, $p < .001$), and so is the interaction (57 ms difference for the literal phrases vs. 115 ms for the idiomatic phrases, $F_1(1,15) = 5.89$, $MS_e = 13{,}617$, $p < .05$; $F_2(1,15) = 7.29$, $MS_e = 15{,}155$, $p < .05$).

Table 1
Mean production latencies and standard deviations in Experiment 1

Idiomaticity	Prime type	
	Unrelated	Identity
Literal	890 (155)	833 (156)
Idiomatic	922 (167)	807 (145)

Though the main effect of Idiomaticity has proven to be non-significant, there is still a 32 ms difference in mean speech onset latencies in the unrelated condition. However, paired comparisons show that this difference is not significant ($t_1 = -1.235$, $SD = 103$, $p = .118$, $t_2 = -1.3804$, $SD = 94$, $p = .094$, one-tailed test). The difference between the mean speech onset latencies in the related condition (26 ms), with idiomatic phrases being faster than literal phrases, is significant in the analysis by subjects only ($t_1 = 2.020$, $SD = 53$, $p < .05$, $t_2 = 1.3201$, $SD = 89$, $p = .103$).

Discussion

The results support the hypothesis that during the planning of an idiomatic phrase the single words that make up the utterance are accessed separately. Both idiomatic and literal phrases can be primed successfully by means of priming one of their content words. This effect supports the compositional nature of idiomatic expressions. Moreover, the effect of Priming is stronger in the case of idioms. This is in favor of our hypothesis that the different components of an idiom are bound together by one common entry in the mental lexicon. Priming *one* of an idiom's elements results in spreading activation from the element to all the remaining elements via a common idiom representation, resulting in faster availability of these elements. For literal items, no such common representation exists. Priming speeds up the availability of the primed element, but cannot help preparing the remaining elements of the phrase. Taken together, the results of Experiment 1 confirm the idea of a hybrid account of idiom representation.

An important assumption of this model is that during the production of an idiomatic expression the same word representations that are used in literal language production are involved. Thus, for example, the production of the idiom *She was skating on thin ice* involves the same representations of *thin* and *ice* that are involved in literal utterances like, for example, *Their tent was covered with a thin layer of ice*. Fixed expressions and literal language only differ with respect to the *source* of word activation: while the words of a literal phrase are activated by their own lexical concepts, the words of a fixed expression will benefit from a common idiom node. Nevertheless, spreading activation from the word level to the concept level will lead to active literal word meanings in both cases.

This view of a direct link between fixed expressions and the semantic network of lexical concepts in the mental lexicon is mirrored in various taxonomies of idiomatic expressions (e.g., Cacciari & Glucksberg, 1991; Nunberg et al., 1994; Gibbs, Nayak, Bolton, & Keppel, 1988; Glucksberg & Keysar, 1993) that assume a contribution of literal word meanings to the idiomatic interpretation, be it to different degrees. Experimental evidence for the activity of literal word meanings during idiom comprehension comes from, for example, Cacciari and Tabossi (1988) and Titone and Connine (1999, but see also Peterson et al., 2001). On the production side, Cacciari and Glucksberg (1991) argue that literal word meanings contribute to an idiom's productive use in discourse, and Cutting and Bock (1997, Experiment 2) show that literal meaning similarity between an idiom and a non-idiomatic phrase enhances the probability of blending errors between the two.

We designed two experiments to test the assumption that literal word meanings become active during idiom production: Experiment 2 addresses the question whether the production of idioms involves the same lemmas that are otherwise part of non-idiomatic language production and have their own meaning and lexical concept. Experiment 3 tests if these literal word meanings are active when an idiom is produced.

Experiment 2

If idioms make use of word representations that are unique to the idiom, idiom production should not be affected by the presentation of a prime word that is semantically related to one of its words. If, however, the building blocks of an idiom are simple lemmas that are not unique to the idiom, they can either be activated by the idiom representation (and thus function as parts of an idiom), or by their own lexical concept as parts of a literal utterance. In these cases, the production of a simple lemma should be sensitive to the presentation of a semantically related prime word (compared to an unrelated condition). Such a sensitivity has been shown outside the domain of idiom production (e.g., Levelt et al., 1991; Peterson & Savoy, 1998). We tested this prediction in an idiom completion task that required the production of the last word of an idiom in response to a visually presented idiom fragment. Completing well-known idioms allows to study idiom production without an initial learning phase (as in Experiment 1), as reading the first part of the idiom provides easy access to its remaining parts. In our case, participants were asked to produce the last word of an idiom. This procedure allowed us to measure speech onset latencies for single words instead of phrases.

Acoustic prime words were presented at different stimulus onset asynchronies (SOAs), relative to the presentation of the visual stimulus. These prime words were either semantically related, phonologically related, or unrelated to the to-be-produced target word. Phonological priming manipulates the preparation of a word's phonological form and is therefore expected to be independent from meaning related factors like Idiomaticity. Given the possibility of a null effect for the semantic condition, the phonological effect can function as a general indicator of the paradigm's sensitivity to priming.

Method

Participants

Seventy-one participants were tested, all being undergraduate students of the University of Nijmegen and native speakers of Dutch. They were paid for their participation.

Materials

Sixteen Dutch idiomatic expressions were chosen as experimental items. They all were presented as finite phrases in past tense form, and they all shared the same syntactic structure:
Jan [VP [V [PP Prep [NP art/pro N]]]]. For example:

Jan viel door de mand.
Jan fell through the basket.
Jan was exposed/failed.

That is, their word order was 'Jan [verb, past tense singular] [preposition] [determiner] [noun]' in all cases.

The first part of the sentence (up to the determiner) functioned as stimulus for producing the last word of the idiom. Thus, the presentation of *Jan viel door de* ... was the stimulus for producing *mand*. In addition to the written stimulus, participants were presented with acoustic primes. These primes were either unrelated, phonologically related, or semantically related to the last word of the idiom. All prime words were short Dutch nouns, and they were all spoken by the same female native speaker of Dutch. All acoustic primes were recorded in one session. The semantic primes were chosen such that they belonged to the same semantic field as the noun of a particular item. The phonological primes were chosen such that they shared the noun's onset. For example, the prime words for the word *mand* 'basket' that belongs to *Jan viel door de*... were

map ('file'; phonological prime) and *korf* ('basket'; semantic prime).

Primes that are related to one item functioned as unrelated primes for the other items. The complete materials are listed in Appendix D.

Procedure

Participants were tested individually, and each session was recorded on DAT tape. The visual stimuli were presented on a computer screen, the acoustic stimuli via headphones. Responses were spoken into a microphone that was attached to a voice key, which in turn signaled the computer that a response had been initiated. The experiment was controlled by NESU.

The production of the correct nouns required the participants to be familiar with the idiomatic expressions. This was tested in a paper-and-pencil cloze task at the beginning of each experiment. Participants were presented a list of the 16 idioms without their respective nouns that they were asked to fill in. Each idiom was preceded by a short literal paraphrase, for example *"Jan was quite ashamed" or "Jan sank through the* .... Participants were asked to indicate on a scale from 1 to 5 how difficult it was to fill in the blank (with 1 = "very easy" and 5 = "very difficult"). The list was then checked by the experimenter and difficult items were clarified. Up to this point, the participants were not aware of the fact that they would have to produce these idioms in the remainder of the experiment. For the on-line task, participants were instructed to produce the missing nouns, just as they had done in the cloze task, in response to the sentence fragments on the computer screen. They were instructed to react as quickly as possible, but in a fluent fashion and without making mistakes. They were asked to reduce coughing, etc., as far as possible and to avoid unnecessary noises that would trigger the voice key. Participants were told that in addition to the visual stimuli they were going to be presented acoustic stimuli (that had to be ignored) via the headphones, and that their responses would be recorded on audio tape. The experiment started with a short practice session of 15 trials, in which participants could get acquainted with the different tasks and the experimental setting. They were then presented 512 experimental trials in four blocks. The blocks were separated by a short pause. A new block was started when the participant indicated that he or she was ready to proceed.

After a fixation cross had appeared in the center of the screen, participants saw one of the sentence fragments that they had seen earlier in the cloze task. Dependent on the experimental condition, an acoustic prime word or distractor was presented via the headphones. Prime onset was varied in relation to sentence presentation (stimulus onset asynchrony, SOA). The participants' task was to react to the visually presented sentence fragment by producing the appropriate noun as quickly as possible. Four different SOAs were tested (−150, 0, 100, and 200 ms, between Subjects). Response times were measured from visual stimulus presentation on. When the voice key triggered, the visual stimulus got removed from screen and a new trial was initiated. If, however, participants failed to respond within 2000 ms, the trial was stopped automatically and counted as a timeout error.

Design

Within each of the four SOAs, each of the 16 items was presented in 32 trials. Each item appeared in the following conditions: (1) with a semantic prime, (2) with a phonological prime, (3) with an unrelated prime, and (4) without prime.

Each item was presented eight times under each of the four conditions. In the unrelated condition, semantic and phonological primes were rotated over items such that they functioned as unrelated primes. Four unrelated primes/distractors stemmed from the group of phonological primes, and the four remaining ones stemmed from the group of semantic primes (yielding the two subconditions Phon-unrel and Sem-unrel). Together, this design resulted in 512 trials per experimental session.

Analyses

DAT-tape recordings of 71 participants were checked for erroneous or missing responses and dysfluencies. Data from seven participants were removed from the data set, because of more than 10% errors. Item 16 was removed from the data set, because of more than 22% errors (compared to 6% errors on average). For the remaining data, an analysis of errors was conducted.

Error percentages per Subject per condition were analyzed in a series of planned comparisons between different levels of the factor Priming, which has four levels: Phon-rel (phonologically related), Sem-rel (semantically related), No prime or distractor, and Unrel (unrelated). The factor level Unrel can further be divided into Phon-unrel (unrelated primes from the set of phonologically related primes) and Sem-unrel (unrelated primes from the set of semantically related primes).

Reaction times exceeding twice the standard deviation from the Subject means (per priming condition) counted as outliers and were excluded from the set of valid responses (2.7% of the valid responses). The reaction time data of the remaining set of correct responses were analyzed in a series of planned comparisons.

Results and discussion

The cloze task showed that participants were highly familiar with the idioms that were presented in the experiment. The average cloze probability was 84%, and the average difficulty score was 1.8. Ninety-five percent of all items that were not completed correctly in the cloze task were subsequently identified by the participants as well-known idioms. In the on-line experiment, five percent of all responses were errors. Table 2 shows the mean error percentages per level of prime type and SOA. In general, participants made fewer errors when a related distractor word was presented than when an unrelated distractor word was presented. Planned comparisons show a significant difference between Phon-rel and Phon-unrel and between Sem-rel and Sem-unrel for SOAs 0, 100, and 200. For SOA −150, the direction of the difference follows that of the other SOAs. When no prime word was presented, participants made fewer errors than when an unrelated prime was presented, and more errors than when a related prime was presented. *T* statistics for planned comparisons of the error percentages in the related and unrelated conditions are provided in Table 3.

Table 2
Mean error percentages per level of prime type and SOA in Experiment 2

SOA	Phon-rel	Phon-unrel	Sem-rel	Sem-unrel	Unrel	No prime
−150	4.9	5.5	4.6	5.9	5.7	5.3
0	4.0	8.4	4.7	8.3	8.4	5.4
100	3.3	7.6	4.1	7.1	7.4	4.4
200	3.1	5.8	3.6	5.8	5.8	5.4

Table 3
t statistics for planned comparisons of the error percentages in the related and unrelated conditions in Experiment 2

SOA	Comparison	t_1	*SD*	*p*	t_2	*SD*	*p*
−150	*Phon-rel–Phon-unrel*	−.592	0.04	.282	−.669	0.04	.257
	Sem-rel–Sem-unrel	−1.141	0.04	.089	−1.270	0.04	.113
0	*Phon-rel–Phon-unrel*	−3.345	0.05	<.05	−3.424	0.05	<.01
	Sem-rel–Sem-unrel	−2.709	0.05	<.05	−3.909	0.04	<.01
100	*Phon-rel–Phon-unrel*	−5.409	0.03	<.001	−4.273	0.04	<.01
	Sem-rel–Sem-unrel	−3.719	0.03	<.01	−2.240	0.05	<.05
200	*Phon-rel–Phon-unrel*	−2.230	0.05	<.05	−2.290	0.05	<.05
	Sem-rel–Sem-unrel	−2.956	0.03	<.01	−2.281	0.04	<.05
All	*No prime–Unrel*	−3.45	3.90	<.01	−2.912	2.23	<.05

p values are given for the one-tailed test. Values for t_1 refer to the analysis with Subjects as random factor (per SOA $df = 15$, across all SOAs $df = 63$), values for t_2 refer to the analysis with items as random factor ($df = 14$).

The mean reaction times per level of Priming per SOA are shown in Table 4. A comparison of the unrelated condition to the no-prime condition across SOAs shows that completing the phrase when hearing an unrelated word took on average 74 ms longer than completing it while no acoustic stimulus was presented. Similar to the difference in errors, this difference mirrors the processing of an extra stimulus. Therefore, we used the unrelated condition as a baseline for all effects of related priming. The resulting relative effects per SOA are illustrated in Fig. 2.

Planned comparisons of the mean reaction times in the related and unrelated conditions reveal significant priming effects (two-tailed test) for both phonologically related and semantically related primes, for SOAs −150, 0, and 100. t statistics are provided in Table 5.

For SOA 200, the phonological effect is significant, but in the subject analysis the semantic effect is only marginally significant. However, it should be kept in mind that when comparing the related conditions (Phon-rel and Sem-rel) to the unrelated conditions (Phon-unrel and Sem-unrel), the number of observations in the unrelated conditions is only half the number of observations in the related conditions. Thus, in terms of sample size, one might rather want to compare the related conditions to the overall unrelated condition (unrel), which summarizes the RTs of trials with unrelated distractors from both the set of phonologically related and the set of semantically related distractors. In that case, the effect of the semantically related condition for SOA 200 is significant as well ($t_1(1,15) = -2.891$, $SD = 36.1$, $p = .011$; $t_2(1,14) = -3.468$, $SD = 28.4$, $p = .004$).

The results show that the production of nouns as parts of idiomatic expressions can be speeded up by both phonologically related and semantically related acoustic distractors. The priming effect of phonologically related words confirms the sensitivity of the paradigm for the effects of acoustic priming. The effects of phonological and semantic priming have been found to be significant

Table 4
Mean reaction times (in ms) per level of prime type and SOA in Experiment 2

SOA	Phon-rel	Phon-unrel	Sem-rel	Sem-unrel	Unrel	No prime	Overall
−150	840	905	819	895	900	826	864
0	816	925	843	897	911	827	896
100	791	891	847	879	885	806	849
200	753	823	792	814	818	758	793

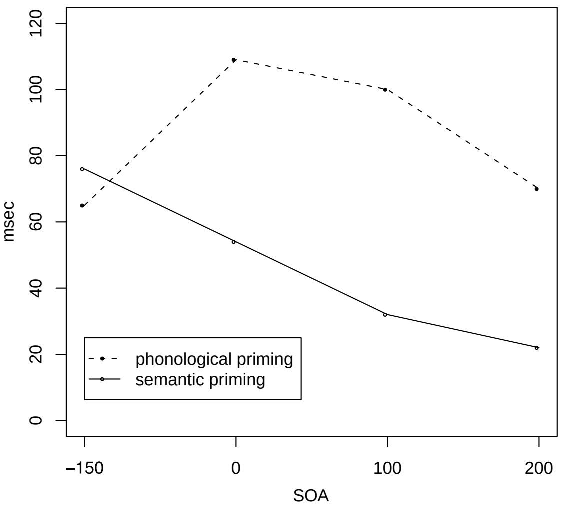

Fig. 2. Effects of phonologically related and semantically related primes in Experiment 2. Values on the vertical axis refer to mean differences between Phon-unrel and Phon-rel, and between Sem-unrel and Sem-rel, respectively.

across all four SOAs tested. The graph in Fig. 2 shows the relative effect of Priming across SOAs, illustrating that the effect of phonological priming is strongest when the prime is presented in parallel with or shortly after the idiom fragment. In contrast, the effect of semantic priming is strongest if the prime is presented 150 ms before the sentence fragment. Thus, we find a general pattern of early semantic and later phonological effects.

The fact that speech onset latencies for the production of an idiom noun can be influenced successfully with a semantically related distractor confirms the prediction made by the hybrid account of idiom production. Specifically, the results indicate that the very same lexical entry can be activated in fundamentally different ways. It can either be selected because of its semantics (as in normal language production), or because it has a fixed link to the representation of an idiomatic expression. It is important to note that, in both cases, we are dealing with the same lexical entry. If the representation of *ice* as part of *skate on thin ice* were different from *ice* as frozen water, then no effect of the semantically related distractor should have been established at all. Thus, the effects found support the assumption that the representation of an idiom activates simple lemmas that are lexical entries on their own. These simple lemmas are not special to the idiom, but are natural elements of the speaker's lexicon.

Experiment 3

Experiment 2 showed that idioms can be primed with a prime word that is semantically related to one of its content words. The priming effect strongly suggests a semantic link between these two words, but it does not prove that this link is bidirectional, that is, that activation actually spreads from an idiom word to its conceptual representation when an idiom is produced. In other words, it still remains to be shown that literal word meanings become active during idiom production.

Experiment 3 was designed to exploit a preparation effect that should arise when speakers who are planning to complete an idiomatic expression have to switch task and read out loud a visually presented word that is *semantically related* to the literal meaning of the target word. The preparation of the idiom's target lemma should co-activate words that are semantically related. For example, the preparation of *ice* as part of *skate on thin ice* should result in the co-activation of *freeze*. This co-activation is expected to affect the speech onset latencies in a reading task. Specifically, the semantically related target *freeze* is expected to be available faster than a semantically unrelated target like, e.g., *tree*. The prepa-

Table 5
t statistics for planned comparisons of the mean reaction times in the related and unrelated conditions in Experiment 2

SOA	Comparison	t_1	*SD*	*p*	t_2	*SD*	*p*
−150	*Phon-rel–Phon-unrel*	−7.616	33.9	<.001	−5.657	44.1	<.001
	Sem-rel–Sem-unrel	−7.449	40.7	<.001	−6.972	43.1	<.001
0	*Phon-rel–Phon-unrel*	−7.047	62.2	<.001	−6.626	63.9	<.001
	Sem-rel–Sem-unrel	−4.538	47.9	<.001	−4.563	46.2	<.001
100	*Phon-rel–Phon-unrel*	−9.137	44.0	<.001	−8.061	48.8	<.001
	Sem-rel–Sem-unrel	−2.961	63.6	<.01	−2.866	44.7	<.01
200	*Phon-rel–Phon-unrel*	−5.621	49.7	<.001	−5.608	47.8	<.001
	Sem-rel–Sem-unrel	−1.876	46.2	<.05	−2.159	37.1	<.05
All	*No prime–Unrel*	−10.641	55.6	<.001	−12.473	22.8	<.001

p values are given for the one-tailed test. Values for t_1 refer to the analyses with Subjects as random factor ($df = 15$, $df = 63$), values for t_2 refer to the analysis with items as random factor ($df = 14$).

ration effect thus should be mirrored in shorter speech onset latencies for *freeze* than for *tree*.

Method

A variant of a task used by Peterson and Savoy (1998) was used to explore the activation of literal word meanings during idiom production. In this task, the preparation of the last word of an idiomatic expression (completion task) was used to prime the production of visually presented target words (reading task). Again, target words could be either phonologically related, semantically related, or unrelated. The phonologically related condition was included in order to measure the sensitivity of the paradigm for the influence of idiom word preparation on word reading.

Participants

Seventy-two participants were tested, all being undergraduate students of the University of Nijmegen and native speakers of Dutch. They were paid for their participation.

Materials

The materials were identical to those used in Experiment 2, with two exceptions. Item 16 *Jan viel in de smaak* ('Jan fell into the taste,' meaning 'Jan was very popular') was replaced by *Jan viste achter het net* ('Jan fished behind the net,' meaning 'Jan did not get what he wanted'), due to the large number of errors for this item in Experiment 2. The words that had been presented as auditory primes in Experiment 2 now functioned as visual targets in the naming task.

Procedure

The experimental set-up was identical to the one described for Experiment 2. Again, participants were presented with a paper-and-pencil cloze task that tested their familiarity with the items before the actual experiment started.

For the on-line task, participants were told that their main task was the fast completion of visually presented idiom fragments in response to a question mark that would appear on the screen below the idiom fragment. They were also told that instead of a question mark, in some cases a word could be presented. In this case, participants would have to switch task and read out loud the word stimulus. Although in the instruction the latter task was presented as a secondary task, the actual ratio of completion trials and reading trials was 50:50. Both kinds of trials started with the presentation of a fixation cross, followed by the presentation of the idiom fragment. In the completion trials, a red question mark appeared after 100, 200, 300 or 400 ms (SOA), in a center position right beneath the idiom fragment. Response latencies were measured from the presentation of the question mark onward. The screen was cleared as soon as the voice key was triggered. If no response was given within 1200 ms, the screen was cleared automatically and the response was coded as timeout error. In the reading trials, a word appeared in red letters in the same position as the question mark would have appeared in the completion trials. The interval between the presentation of the idiom fragment and the word varied according to the SOA. The word could be either phonologically related, semantically related, or unrelated to the target word. Participants were instructed to read the word out loud instead of completing the idiom. Response latencies were measured from the presentation of the target word onward. As in the procedure used by Peterson and Savoy (1998), trial length was kept short in order to avoid strategic behavior and to encourage the preparation of idiom completion immediately after the beginning of idiom fragment presentation.

Design

Within each of the four SOAs, each of the 16 items was presented in 32 trials. Half of the trials were completion trials, the other half were reading trials. In the reading trials, each item appeared in three different conditions: (1) with a phonologically related target (Phon-rel), (2) with a semantically related target (Sem-rel), and (3) with an unrelated target (Unrel).

Each item was presented four times in condition one, four times in condition two, and eight times in condition three. In the unrelated condition, semantically and phonologically related targets were rotated over items such that they functioned as unrelated targets. Four unrelated targets stemmed from the group of phonologically related targets, and the four remaining ones stemmed from the group of semantically related targets (yielding the two subconditions Phon-unrel and Sem-unrel). With 16 different items, the design resulted in a total of 512 trials per experiment.

Analyses

DAT-tape recordings of 72 participants were checked for erroneous or missing responses and dysfluencies. Data from eight participants were removed from the data set, because of more than 20% errors in the idiom completion trials. For the remaining data, an analysis of errors was conducted.

Error percentages per subject per condition were analyzed in a series of planned comparisons between different levels of the factor Relatedness, which has three levels: Phon-rel (phonologically related), Sem-rel (semantically related), and Unrel (unrelated). The factor level Unrel can further be divided into Phon-unrel (unrelated primes from the set of phonologically related primes) and Sem-unrel (unrelated primes from the set of semantically related primes).

Table 6
Mean error percentages per task, SOA, and level of relatedness in Experiment 3

SOA	Naming				Completion
	Phon-rel	Phon-unrel	Sem-rel	Sem-unrel	
100	4.7	3.2	3.4	3.3	10.3
200	3.3	3.3	2.6	2.5	9.7
300	4.1	2.8	2.3	3.1	5.9
400	4.1	3.1	3.5	2.9	6.9

For the reading trials, reaction times that exceeded twice the standard deviation from the Subject × condition means counted as outliers and were excluded from the set of valid responses (2.7% of the responses). The reaction time data of the remaining set of correct responses were analyzed in a series of planned comparisons. Separate error and reaction time analyses were conducted for the four different SOAs.

Results and discussion

The average cloze probability was 89%, and the average difficulty score was 1.5. Ninety-five percent of all items that were not completed correctly in the cloze task were subsequently identified as well-known idioms.

In the on-line experiment, six percent of all responses were errors. As expected, most errors were made in the completion trials (8.3%). In the reading trials, error percentages are relatively low (3.6%).

Table 6 shows the mean error percentages per level of prime type and SOA. Planned comparisons show no significant difference between Phon-rel and Phon-unrel and between Sem-rel and Sem-unrel for any of the SOAs. The only significant difference is the one between tasks. Participants make more errors in the idiom completion task than in the word reading task. *T* statistics for the planned comparisons of the mean error percentages in the related and the unrelated conditions are provided in Table 7.

The mean reaction times per level of relatedness and SOA are shown in Table 8. The relative effects of the related primes per SOA are illustrated in Fig. 3. The last column in Table 8 shows the reaction times for the completion task. With longer SOAs, reaction times decrease in this task. This can be seen as evidence for the preparation of the utterance in response to the idiom fragment when subjects do not know yet what kind of task they have to perform. If subjects applied a strategy (e.g., wait until either the question mark or a word appears before preparing the response), no such decrease would have been observed.

Planned comparisons of the related and unrelated conditions reveal significant effects of idiom preparation for phonologically related and semantically related targets at different SOAs. At a SOA of 100 ms, both effects become significant in the subject analysis, but not in the item analysis. At SOA 200, the semantic effect is established in both subject and item analysis. The opposite holds for SOA 300. Here, only the phonological effect

Table 8
Mean reaction times (in ms) per task, SOA, and level of relatedness in Experiment 3

SOA	Naming				Completion
	Phon-unrel	Phon-rel	Sem-unrel	Sem-rel	
100	576	566	574	563	668
200	600	594	596	585	644
300	570	558	568	569	532
400	558	551	560	559	516

Table 7
t statistics for planned comparisons of the mean error percentages in Experiment 3

SOA	Comparison	t_1	*SD*	*p*	t_2	*SD*	*p*
100	*Phon-rel–Phon-unrel*	−1.99	0.03	<.05	−1.29	0.05	.107
	Sem-rel–Sem-unrel	−0.15	0.03	.442	−0.14	0.03	.447
	Completion–Naming	−7.45	0.04	<.001	−5.45	0.05	<.001
200	*Phon-rel–Phon-unrel*	0	0.03	1	0	0.03	1
	Sem-rel–Sem-unrel	−0.16	0.02	.438	−0.12	0.03	.453
	Completion–Naming	−5.06	0.05	<.001	−4.62	0.06	<.001
300	*Phon-rel–Phon-unrel*	−1.48	0.03	.080	−1.23	0.04	.137
	Sem-rel–Sem-unrel	1.41	0.02	.089	1.05	0.03	.155
	Completion–Naming	−2.10	0.05	<.05	−2.83	0.04	<.01
400	*Phon-rel–Phon-unrel*	−0.99	0.04	.169	−1.37	0.03	.095
	Sem-rel–Sem-unrel	−0.70	0.03	.249	−0.89	0.03	.192
	Completion–Naming	−4.07	0.03	<.01	−3.53	0.04	<.01

p values are given for the one-tailed test. Values for t_1 refer to the analysis with Subjects as random factor ($df = 15$), values for t_2 refer to the analysis with items as random factor ($df = 15$).

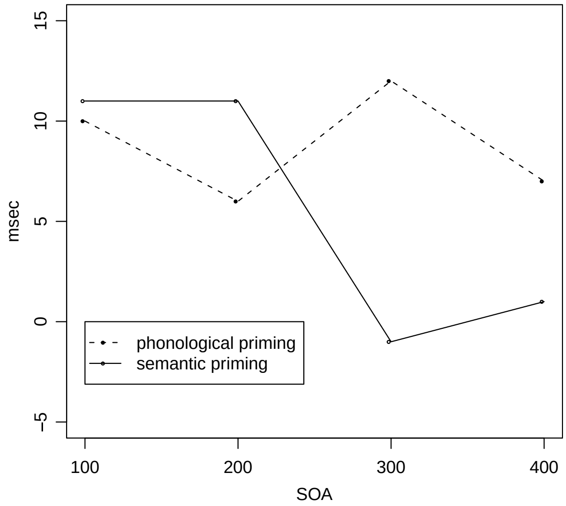

Fig. 3. Effects of idiom preparation on phonologically related and semantically related targets in Experiment 3. Values refer to mean differences between Phon-unrel and Phon-rel, and between Sem-unrel and Sem-rel, respectively.

is established. SOA 400 shows a significant effect of phonology, but again only in the subject analysis. *T* statistics are provided in Table 9. A more detailed item analysis was conducted in order to identify possible subgroups of items. An interaction of item group and condition might have explained the rather weak effects in the item analyses. However, no such subgroups were found. Fig. 3 illustrates the time course of the phonological and the semantic effect across the different SOAs. Although only the strongest effects reach significance in both item and subject analyses, the figure shows clear trends in the predicted direction. Like in Experiment 2, a pattern of early semantic effects and later phonological effects is established.

The results indicate that the preparation of a word as part of an idiom can affect the production latencies of words that are phonologically or semantically related to this word. Again, the presence of a phonological effect confirms the sensitivity of the paradigm to measure these effects. The presence of a semantic effect can be seen as evidence for the activation of literal word meanings during the production of idioms. The effect has been predicted by the hybrid account of idiom production, because it assumes the activation of the lexical concept nodes of the lemmas that have been selected as parts of the idiom representation.

General discussion

Three experiments were conducted to test a hybrid account of idiom representation. Such an account assumes that idioms are both unitary and compositional, although at different levels of their cognitive representation. They have a unitary idiomatic concept that points to individual lemmas. These lemmas together constitute the idiom, but they are not bound exclusively to an idiomatic meaning. For example, the idiom *he hit the road* 'he left' will involve the same lemma "road" that is active during the production of "he cleaned the road" (i.e., a literal phrase). It is the *source* of activation for "road" that differs in the two cases.

In Experiment 1, we observed a significant main effect of Priming during phrase production. Hearing a word identical to the noun of the phrase that is being planned significantly reduces production latencies for that phrase, relative to an unrelated condition. This holds for both idiomatic and literal phrases, suggesting that the underlying representation of idioms is a word-based representation. This idea is further supported by a significant interaction between Priming and Idiomaticity. The relative reduction in planning time accomplished by identity priming is larger in the case of idioms than in the case of literal phrases.

Thus, part-whole priming of idioms is not only possible, it is even more effective than part-whole priming of literal phrases. This is exactly what a hybrid account

Table 9
t statistics for planned comparisons of the mean reaction times in the related and unrelated conditions in Experiment 3

SOA	Comparison	t_1	*SD*	*p*	t_2	*SD*	*p*
100	*Phon-unrel–Phon-rel*	2.801	13.8	<.01	1.487	25.1	.079
	Sem-unrel–Sem-rel	2.103	21.0	<.05	1.681	24.1	.057
200	*Phon-unrel–Phon-rel*	.819	28.5	.213	.832	31.4	.210
	Sem-unrel–Sem-rel	1.824	23.8	<.05	2.626	16.6	<.05
300	*Phon-unrel–Phon-rel*	1.987	24.3	<.05	2.716	18.7	<.01
	Sem-unrel–Sem-rel	−.291	13.7	.388	−.022	28.6	.492
400	*Phon-unrel–Phon-rel*	2.780	10.7	<.05	1.217	25.1	.121
	Sem-unrel–Sem-rel	.377	16.5	.356	.317	25.8	.383

p values are given for the one-tailed test. Values for t_1 refer to the analysis with Subjects as random factor, values for t_2 refer to the analysis with items as random factor ($df = 15$ in both cases).

of idiom representation predicts, assuming that the individual words that constitute an idiom are separately addressable processing units, linked together in a common representation. Boosting the activation of one element of the idiom that is being planned affects all the remaining elements. The observed latency reduction by an identity prime must be ascribed to the enhanced availability of *all* the words that make up the idiom. In contrast, the reduction in literal phrase planning time represents the maximum gain that phrase planning can get from the enhanced availability of only *one* of its elements.

Experiment 2 showed that production latencies for a simple lemma that is produced as part of an idiomatic phrase are considerably shorter when primed with a semantically related word than when primed with an unrelated one. Spreading activation within the semantic network enhances the availability of the target lemma. Experiment 3 shows that the same effect arises when the preparation of an idiom's simple lemma functions as a prime for the production of a semantically related word. Again, spreading activation within the semantic network is assumed to be responsible for the effect. These findings imply that the literal word meanings become active during the production of idiomatic expressions. They are in line with results found by Cutting and Bock (1997), who found an increase in blending errors when there was literal meaning overlap between an idiom and a phrase. They also fit well with the idea that literal word meanings play a role in the productive use of idioms in discourse, as proposed by Cacciari and Glucksberg (1991).

A post hoc analysis of all three experiments that included Decomposability as additional factor showed that the priming effects that we established are not dependent on the extent to which the separate words of an idiom contribute to its overall meaning. However, our analysis also shows that there is at least *some* sensitivity of speech onset latencies for Decomposability. For Experiments 2 and 3 we find that high Decomposability is sometimes related to longer speech onset latencies (and in one case to shorter speech onset latencies). This is an interesting finding, because Cutting and Bock (1997) did not find an effect of Decomposability on speech error rates for idioms. However, the set-up of our experiments and the relative inconsistency of the effects does not allow further conclusions as to the origin of these effects.

In sum, all three experiments support a hybrid account of idiom production as formulated by Cutting and Bock (1997). Together with their speech error data, the results of this study constitute a firm empirical basis for the assumption of idioms as being compositional and non-compositional at the same time.

As was outlined in the introduction, the hybrid account of idiom production is generally compatible with Cacciari and Tabossi's (1988) model of idiom comprehension. According to their configuration hypothesis, all elements of an idiom contribute to the activation of an idiom representation. Special elements of an idiom (the idiom *key*) will lead to direct activation of the idiom representation in the mental lexicon, even before its last element has been perceived. In terms of an activation model, the necessity of an idiom key can easily be reformulated as the presence of an activation threshold. Accordingly, an idiom is recognized once its activation rises above a certain critical threshold. As some elements might activate an idiom more than others, idiom recognition speed will vary with the order of elements perceived.

The fact that the available evidence in the field of idiom processing supports two highly compatible theories of idiom comprehension and production is by no means trivial. It supports a view in which both the lexical-conceptual and the lexical-syntactic processing level are shared between comprehension and production (e.g., Kempen & Harbusch, 2002; Roelofs, 2003). According to this view, it should be possible to read a language processing model in two directions: top-down (from concepts to lemmas) as a production model, and bottom-up as a comprehension model. However, as we will discuss below, this mental exercise reveals an inconsistency in Cutting and Bock's account with respect to the connections between processing levels. We therefore propose a modification of the hybrid model enabling reading it in both directions. Moreover, this modification will offer an alternative for Cutting and Bock's suggestion to represent the syntactic format of idioms by means of *phrasal frames*.

In Cutting and Bock's (1997) model, idioms are represented by their own lexical concepts that spread activation to the lexical-syntactic (lemma) level. Unlike simple concepts that represent single words, idiom concepts spread activation to more than one word at the same time. Still, all connections that run top-down from the level of lexical concepts to the lemma level express a meaning relationship. Accordingly, these connections can be labelled "is (partly) expressed by": $hit_{(concept)}$ is expressed by $hit_{(lemma)}$, and $hit-the-road_{(concept)}$ is partly expressed by $hit_{(lemma)}$, $the_{(lemma)}$, and $road_{(lemma)}$.

Consequently, connections in the opposite direction (i.e., from lemma to concept) should express the relationship "has the meaning of": $hit_{(lemma)}$ has the meaning of $hit_{(concept)}$. However, this does not work in the case of idioms: $hit_{(lemma)}$ does not have the meaning of $hit–the–road_{(concept)}$. Instead, $hit_{(lemma)}$ is simply one of the elements of the idiom. Thus, the bottom-up connection between these two nodes expresses a hierarchical "element-of" relationship. In other words, the hybrid model requires two distinct types of connections from the lemma level to the lexical-conceptual level: one that expresses a meaning relationship, and one that expresses a simple allocation of one element to another. The different types of connections are illustrated in Fig. 4.

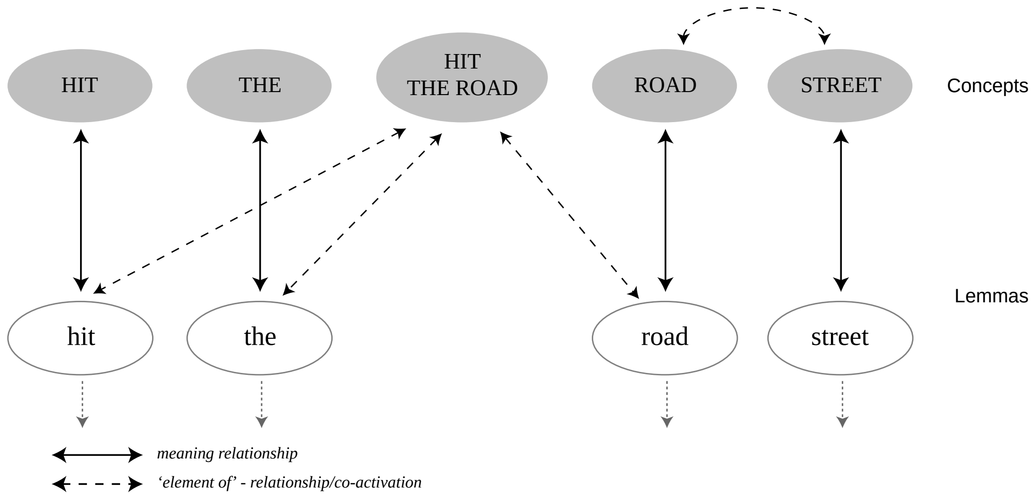

Fig. 4. Representation of the idiom *hit the road* in terms of the hybrid model, with two types of connections between the concept and the lemma level. The phrasal frame is not depicted.

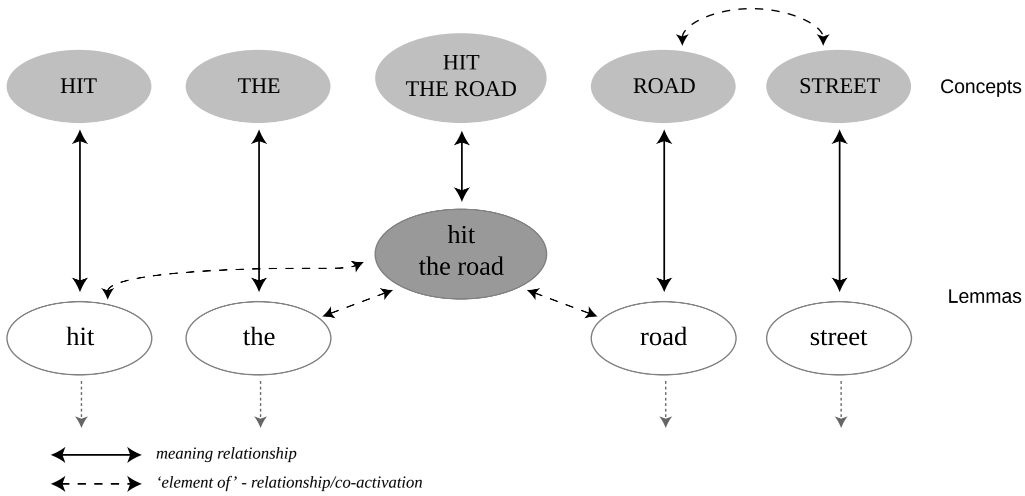

Fig. 5. Representation of the idiom *hit the road* according to the adapted hybrid model. The idiom is represented both at the concept level and the lemma level. All connections between processing levels denote the same relationship.

The assumption of qualitatively different connections implies that there are distinct processing mechanisms for idioms and for literal language. However, that stands in contrast to the notion that idioms are not special (e.g., Cutting & Bock, 1997).

In our view, this contradiction can be solved by introducing a *superlemma*, a separate representation of the idiomatic phrase on the lexical-syntactic processing level. The superlemma is a representation of the syntactic properties of the idiom that is connected to its building blocks, the simple lemmas (see Fig. 5). Thus, the hybrid character of the model is kept in good order. In addition, the connections between processing levels are now identical for both idioms and simple words. Specifically, all connections between the lemma level and the lexical concept level express a meaning relationship. The "element-of" connections between the simple lemmas and the idiom representation are now restricted to connections within the lemma level. With respect to these connections, we assume a spreading activation process that is similar to the the activation flow between lexical concepts.

In this approach, the selection and processing of an idiom representation is highly similar to the selection and processing of a single word. By representing idioms with their own lemma, idiom production follows the same rules of lexical competition and lexical selection as single words do. Thus, we assume that the production of the idiom *hit the road* requires the selection of its superlemma. Since lexical selection involves competition among co-activated lemmas, the superlemma *hit the road* competes, for example, with *leave* and will only be selected if it is the most highly activated node in the system. The probability of selecting the target superlemma from the mental lexicon is the ratio of the superlemma's degree of activation and the total activation of all lemmas (superlemmas and simple lemmas) in the lexicon (known as Luce's ratio). Upon selection of the superlemma, the syntactic constraints that come along with the idiom become available to the production system. They delimit or modify the syntactic properties of the simple lemmas involved. Moreover, the selection of the target superlemma fixates the set of simple lemmas that are to be selected in subsequent processing steps. We assume that superlemma selection is a *condition* on activation spreading towards the "dependent" simple lemmas. The process of simple lemma selection is, again, based on Luce's ratio. The target lemma is always in competition with all other active superlemmas and simple lemmas.

Thus, our modified version of the hybrid model simplifies the assumptions about the connections between processing levels at the expense of the introduction of a new node. However, these costs are balanced by the fact that the model renders the notion of *phrasal frames* unnecessary. By representing an idiomatic expression on the lexical-syntactic level, all syntactic constraints that come along with the idiom can be accommodated with one type of representation (the lemma). Moreover, the superlemma approach differs from the phrasal frame approach in that it specifies the grammatical relations between the actual lemmas involved in the idiom. A superlemma is best characterized as a (phrasal) function over some set of simple lemmas. It specifies the syntactic relationships between the individual lemmas, sometimes modifying the pre-existing syntactic options of the simple lemmas it dominates. For instance, the passive option of the simple lemma "hit" is disabled by the idiom's superlemma ("hit the road"). In terms of the Performance Grammar formalism (Kempen & Harbusch, 2002) one might say that the task of the superlemma is to finetune the "lexical frames" that are associated with individual lemmas. In this view, the syntactic features of the superlemma's elements form the building blocks of its structure. This structure is reduced in its syntactic potential, making the idiom syntactically less flexible. For example, the syntactic information for the idiom *hit the road* might be represented as shown in Fig. 6.

In our view, the constraint-based representation of the syntactic features of an idiom can easily explain the syntactic idiosyncrasies of individual idioms. For the shared phrasal frames, however, this is more difficult to accommodate, as illustrated by the following examples.

Phrasal frames provide a phrase structure with open slots that can be filled with the simple lemmas that are activated by the idiom's lexical concept in parallel. Thus, the sentence *John hit the road* is assumed to have activated a phrasal VP frame with open slots for a verb and a direct object. Filling in *hit* and *the road* in these slots seems rather straightforward. However, the situation is more complicated in the case of a VP with an additional NP: *be a wolf in sheep's clothing*. The phrasal VP frame provides two slots for two nouns that are possible fillers. In the phrasal frame approach it has not been specified how the system knows where the nouns *sheep* and *wolf* should be inserted. The nouns' lemmas themselves are not assigned to specific grammatical roles, and the phrasal frame is an abstract syntactic structure that is blind to the relationship between the concepts and the active lemmas. Therefore, the system cannot know whether it is *a wolf in sheep's clothing* or rather *a sheep in wolf's clothing* that the speaker intended. However, this order is not arbitrary, and changing it means loosing the idiomatic meaning of the phrase. Additional syntactic constraints must be assumed to solve the problem within the phrasal frame approach. Within the superlemma approach, no such problem arises. The syntactic relationships and constraints that come with the idiom are directly applied to the set of simple lemmas involved in the idiom. No additional operation is required that merges syntax and lemmas. When the simple lemmas get activated, they will already be provided with their

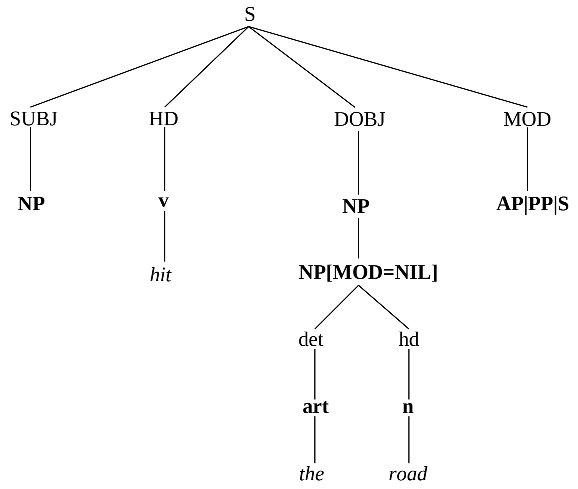

Fig. 6. Information represented in the superlemma for *hit the road*. The lemmas *hit*, *the*, and *road* are obligatory components. Modification of *road* (e.g., **hit the icy road*) is not allowed. Legends: 1. Syntactic functions. S, sentence; SUBJ, subject; HD/hd, head; DOBJ, direct object; MOD, modifier; det, determiner. 2. Syntactic categories. NP, noun phrase; AP, adverbial or adjectival phrase; PP, prepositional phrase. 3. Parts of speech. art, article; n, noun; v, verb. Syntactic notation according to Performance Grammar developed by Kempen and associates (Kempen, 1996; Kempen & Harbusch, 2002; Vosse & Kempen, 2000).

exact position in the idiom's syntactic structure. At the same time, the superlemma account can still explain idiom blends like those observed by Cutting and Bock (1997, Experiment 1). These blends result from the simultaneous activation of two superlemmas with similar meanings and/or syntactic structures.

In sum, the superlemma model offers a theoretical extension of the hybrid model of Cutting and Bock (1997). In addition to explaining the available empirical data on idiom production, it spells out the syntactic nature of idiom representations in more detail and makes clear-cut assumptions about how the syntactic constraints that come along with an idiom are realized during idiom production. Also, the superlemma approach is equally capable of solving the paradox that idioms seem to entail: the holistic nature of idiomatic expressions is not at variance with their generation out of single words in grammatical encoding. Most important, the theory is still fully compatible with our knowledge about *non*-idiomatic phrase production. As we mentioned in the introduction, the production of fixed expressions (and of idioms as prototypical FEs) is the rule rather than the exception. Therefore, both idiomatic and literal production should be captured by a single, consistent production theory. Consequently, we propose the hybrid model as a general model of the representation and processing of fixed expressions, which exist in various degrees of syntactic fixedness, semantic transparency, and decomposability. Although theoretically attractive, the validity of the model for these different types of FEs is not self-evident. For example, one might argue that the degree of decomposability has influence on the extent to which single word representations are involved in the processing of an idiom. Clearly, further research will be needed in order to explore the validity of the model for the various types of FEs. In addition, future research will be needed in order to distinguish empirically between Cutting and Bock's model of idiom representation and the superlemma model.

Acknowledgments

The authors thank Harald Baayen, Koen Kuiper, Hedderik van Rijn, Herbert Schriefers, Mandana Seyfeddinipur, Wietske Vonk, and three anonymous reviewers for their comments on earlier versions of this paper.

Appendix A. Effects of Decomposability

To control for possible interactions of our experimental manipulations with the decomposability of the items involved, we subjected our data to a post hoc analysis. To this aim, we collected Decomposability scores for all idioms that were used in Experiments 1–3. Following the procedure described in Gibbs and Nayak (1989), we asked 40 participants to assess the contribution of the individual words to the idioms' meaning. However, since we intended to include Decomposability into our analysis as a post hoc covariate, we treated it as a numerical variable rather than as a categorical one. Thus, instead of asking participants to categorize the items as either semantically decomposable, abnormally decomposable, or non-decomposable, we asked participants to rate each item's decomposability on a scale from 1 to 6 (the higher the score, the more decomposable the item). The item order was randomized for each participant. Average scores per item set, as well as their median and range are shown in Table A.1. Average decomposability scores for each individual item are listed together with the items in Appendices B and C.

We re-analyzed the data sets of Experiments 1–3 with linear mixed-effect (multilevel) analyses of covariance, with Decomposability as additional numerical predictor, and Subject and Item as crossed random effects (Bates, 2005; Bates & Sarkar, 2005; Pinheiro & Bates, 2000). We discuss the results of these analyses per experiment.

In Experiment 1 we examined the effect of identity primes on the speech onset latencies for literal and idiomatic phrases. Our earlier analyses show a main effect of Priming and an interaction between Priming and Idiomaticity. This interaction is due to a stronger priming effect for idioms than for literal phrases. To examine a possible contribution of Decomposability to the priming effect for idioms, we re-analyzed the speech onset latencies for these items. Decomposability was added as a fixed factor.

An ANOVA on the parameter estimates of the regression model reveals a significant main effect of Priming ($F(1,1744) = 55.65$, $MS = 2{,}252{,}803$, $p < .001$). Neither the main effect of Decomposability nor the interaction between Priming and Decomposability are significant ($F < 0$). Thus, the extent to which the idioms that we used in Experiment 1 were decomposable did not affect the size of the priming effect.

In Experiment 2 all items were idiomatic phrases. The experiment tested the effect of Phonological and Semantic Priming on the speech onset latencies for the last word of an idiom. Adding Decomposability as a numeric predictor variable in the analysis confirmed the earlier found pattern of effects. An ANOVA on the parameter estimates of the regression model of the complete data set shows a significant main effect of Priming ($F(4,27388) = 152.82$, $MS = 4{,}023{,}970$, $p < .001$), as well as a significant interaction of Priming and SOA ($F(12,27388) = 10.18$, $MS = 268{,}171$, $p < .001$). In addition, we find a significant interaction between Priming and Decomposability ($F(4,27388) = 6.26$, $MS = 164{,}780$, $p < .001$), as well as between SOA and Decomposability ($F(3,27388) = 6.15$, $MS = 161{,}930$, $p < .001$). Also, the three-way interaction between Priming, SOA, and Decomposability is significant ($F(12,27388) = 1197838$,

Table A.1
Descriptives of the average decomposability scores for the two idioms sets

Idiom set	Mean	Median	Range
1	3.9	4.0	2.9–5.0
2	3.8	3.9	2.7–5.0

Set 1 was used in Experiment 1, set 2 was used in Experiments 2 and 3.

$MS = 99{,}820$, $p < .001$), while the main effects of Decomposability and SOA are not ($F < 0$ and $F(3, 27388) = 1.12$, $MS = 30{,}439$, $p = .325$, respectively).

These results show that at least some of the variance in our data can be explained by the factor Decomposability. We conducted separate analyses per SOA and per type of Priming (Phonological or Semantic) to find out whether the effects that we established in our earlier analyses still hold in light of this effect. The results are shown in Tables A.2 and A.3. A comparison of these results with the those in Table 3 reveals that all but one effect (Semantic Priming at the longest SOA) remain significant if we allow Decomposability to explain part of the variance in the data. Thus, the overall pattern of Phonological and Semantic Priming that we found in this experiment cannot be ascribed to effects of the decomposability of the items involved. However, the post hoc analysis does reveal some influence of Decomposability on speech onset latencies. For SOA 200, we find a significant interaction between Phonological Priming and Decomposability. Inspection of the relative effect sizes for the fixed effects in the regression model reveals that speech onset latencies in the phonologically related priming condition are longer for idioms that score high on Decomposability ($\beta = 40$). In other words, the effect of Phonological Priming is weaker for highly decomposable idioms when SOA = 200.

For SOA = −150 and SOA = 0 ms we find a significant interaction between Semantic Priming and Decomposability. Effect sizes for the fixed effects reveal that these interactions are due to longer speech onset latencies for more decomposable items in the semantically related condition ($\beta = 34$ and 47, respectively). That is, the effect of Semantic Priming is weaker for highly decomposable idioms in SOA = −150 and SOA = 0 ms.

Taken together, re-analysis of the data of Experiment 2 with Decomposability as extra factor supports our earlier analyses. The effects we obtained for Phonological and Semantic Priming cannot be ascribed to effects of Decomposability. In addition to our earlier results, the analysis also shows that there is at least some effect of idiom Decomposability on the speech onset latencies for words that belong to an idiomatic phrase.

In Experiment 3, we tested the effect of idiom preparation on production latencies for words that are phonologically or semantically related to the idiom's noun. We found effects of Semantic Priming at early SOAs and effects of Phonological Priming which were stronger at late SOAs. We re-analyzed the data in order to uncover a possible influence from the decomposability of the idioms (i.e., the primes in the study) on the speech onset latencies for the target words.

An ANOVA on the parameter estimates of the regression model for the complete data set shows significant main effects of Priming ($F(3, 14959) = 9.05$, $MS = 66{,}420$, $p < .001$) and Decomposability ($F(1, 14959) = 4.97$, $MS = 36{,}465$, $p < .05$), as well as a significant interaction between Priming and SOA ($F(3, 86452) = 3.93$, $MS = 28{,}817$, $p < .01$). The results show that, also in this data set, the decomposability of the items involved can explain part of the variance. We conducted separate analyses of Phonological and Semantic Priming (per SOA) to see if our earlier pattern of results still holds if we take Decomposability into account.

Tables A.4 and A.5 show the effects of idiom preparation per SOA and per type of Priming. There are significant effects of Phonological Priming for three out of four SOAs. Also, the effect of Semantic Priming is significant for the two earliest SOAs. This pattern confirms our earlier results.

Table A.2
Effects of phonological priming and decomposability (per SOA) in Experiment 2

SOA	Effect	*df*	*MS*	*F*	*p*
−150	Priming	1,2560	2,153,962	60.43	<.001
	Decomp	1,2560	10,397	<1	.531
	Priming × Decomp	1,2560	67,916	2.57	.109
0	Priming	1,2530	4,449,026	162.42	<.001
	Decomp	1,2530	39,671	1.45	.223
	Priming × Decomp	1,2530	39,289	1.43	.231
100	Priming	1,2572	2,965,626	107.49	<.001
	Decomp	1,2572	1,390	<1	.822
	Priming × Decomp	1,2572	12,784	<1	.496
200	Priming	1,2631	1,157,970	50.02	<.001
	Decomp	1,2631	10,659	<1	.497
	Priming × Decomp	1,2631	325,760	14.07	<.001

Table A.3
Effects of semantic priming and decomposability (per SOA) in Experiment 3

SOA	Effect	*df*	*MS*	*F*	*p*
−150	Priming	1,2599	2,153,962	82.72	<.001
	Decomp	1,2599	2,110	<1	.776
	Priming × Decomp	1,2599	229,718	8.82	<.01
0	Priming	1,2542	695,699	25.71	<.001
	Decomp	1,2542	6,413	<1	.626
	Priming × Decomp	1,2542	436,895	16.15	<.001
100	Priming	1,2555	262,146	8.54	<.01
	Decomp	1,2555	978	<1	.858
	Priming × Decomp	1,2555	48,217	1.57	.210
200	Priming	1,2572	389	<1	.902
	Decomp	1,2572	8,318	<1	.569
	Priming × Decomp	1,2572	21,670	<1	.358

Table A.4
Effects of phonological priming and decomposability (per SOA) in Experiment 3

SOA	Effect	*df*	*MS*	*F*	*p*
100	Priming	1,1859	42,471	4.53	<.05
	Decomp	1,1859	36,206	3.87	<.05
	Priming × Decomp	1,1859	376	<1	.841
200	Priming	1,1866	14,045	1.52	.217
	Decomp	1,1866	12,806	1.39	.239
	Priming × Decomp	1,1866	424	<1	.830
300	Priming	1,1865	67,978	10.85	<.01
	Decomp	1,1865	25,056	4.00	<.05
	Priming × Decomp	1,1865	3,887	<1	.432
400	Priming	1,1863	27,150	5.79	<.05
	Decomp	1,1863	2,895	<1	.432
	Priming × Decomp	1,1863	50,361	10.74	<.01

Table A.5
Effects of semantic priming and decomposability (per SOA) in Experiment 3

SOA	Effect	*df*	*MS*	*F*	*p*
100	Priming	1862	53,215	6.04	<.05
	Decomp	1862	220	<1	.875
	Priming × Decomp	1862	971	<1	.739
200	Priming	1886	57,681	7.09	<.01
	Decomp	1886	34,869	4.28	<.05
	Priming × Decomp	1886	1,522	<1	.666
300	Priming	1878	437	<1	.802
	Decomp	1878	29,471	4.24	<.05
	Priming × Decomp	1878	14,325	2.06	.151
400	Priming	1864	937	<1	.664
	Decomp	1864	11,298	2.27	.132
	Priming × Decomp	1864	156	<	.859

In addition to these results, we also find an influence of decomposability on speech onset latencies for some combinations of SOA and type of Priming. The analysis of the phonological priming data shows a main effect of Decomposability for SOA = 100 and SOA = 300 ms ($\beta = 9$ and 11, respectively). Speech onset latencies increase with higher scores of decomposability. For SOA = 400 ms we find a significant interaction between Phonological Priming and Decomposability ($\beta = -17$). That is, the effect of Phonological Priming is stronger if subjects prepare highly decomposable idioms. The analysis of the semantic priming data shows a main effect of Decomposability for SOAs = 200 and SOA = 300 ms ($\beta = 10$ and 30, respectively). That is, speech onset latencies are longer if subjects prepare highly decomposable idioms.

Taken together, the results of the re-analysis show that idiom decomposability explains some of the variance in our data. However, their inconsistency and the set-up of our experiment do not allow further conclusions about their origin. Most importantly, the effects do not weaken our interpretation of the priming results that we established in Experiments 1–3.

Appendix B. Materials Experiment 1

Word order is [noun (name)][verb][preposition][noun]

Pair	Item	Prompt word(s)	Idiomatic phrase
01	01	Laura…	viel buiten de boot (3.8) *fell out of the boat*, was excluded from the group
	17	Laura…	ging met de boot *went with the boat*, took the boat
02	02	Mark…	hielp hem uit de brand (5.0) *helped him out of the fire*, helped him out of problems
	18	Mark…	waarschuwde hem voor de brand *warned him of the fire*, warned him of the fire
03	03	Paulien…	kroop uit het dal (4.5) *crawled out of the valley*, got herself together
	19	Paulien…	woonde in het dal *lived in the valley*, lived in the valley
04	04	Jan's feestje… Jan's party…	liep uit de hand (3.3) *went out of the hand*, got out of hand
	20	Jan's dochter… Jan's daughter…	zwaaide met de hand *waved with the hand*, waved her hand
05	05	Karin…	hield hem op de hoogte (3.6) *held him on the height*, kept him informed
	21	Karin…	schrok erg van de hoogte *was startled a lot by the height*, was startled by the height
06	06	Anna's vervalsing… Anna's forgery	kwam aan het licht (4.5) *came into the light*, was discovered
	22	Anna's ketting… Anna's necklace…	fonkelde in het licht *was glittering in the light*, was glittering in the light
07	07	Petra…	bracht alles aan de man (4.1) *brought everything to the man*, could sell everything,
	23	Petra…	sprak vaak met de man *often talked to the man*, often talked to the man
08	08	Kees…	viel vreselijk door de mand (3.0) *fell miserably through the basket*, failed miserably

Appendix B (*continued*)

Pair	Item	Prompt word(s)	Idiomatic phrase
	24	Kees...	legde het hondje in de mand *put the puppy in the basket*, put the puppy in the basket
09	09	Corien...	viste altijd achter het net (4.2) *fished always behind the net*, always came too late
	25	Corien...	ving de vlinder met het net *caught the butterfly with the net*, caught the butterfly with the net
10	10	Erik's schoenen... Erik's shoes...	sprongen in het oog (4.4) *jumped into the eye*, were eye-catching
	26	Erik's vader... Erik's father...	sloeg hem op het oog *hit him on the eye*, hit him on the eye
11	11	Suzan...	zette alles op het spel (3.6) *put everything on the game*, put everything on one card
	27	Suzan...	sjoemelde lelijk bij het spel *was faking awfully during the game*, was faking terribly during the game
12	12	John...	leidde haar om de tuin (2.9) *led her around the garden*, led her up the garden path
	28	John...	zat graag in de tuin *sat with pleasure in the garden*, enjoyed sitting in the garden
13	13	Ton...	sloeg de waarschuwing in de wind (3.4) *hit the warning into the wind*, ignored the warning
	29	Ton...	waarschuwde iedereen voor de wind *warned everybody about the wind*, warned everybody about the wind
14	14	Marieke...	riep de stichting in het leven (4.6) *called the foundation into life*, founded the foundation
	30	Marieke...	peinsde eindeloos over het leven *thought endlessly about life*, endlessly pondered on life
15	15	Henk...	trok vergeefs aan de bel (4.3) *pulled in vain at the bell*, warned without success
	31	Henk...	wachtte vergeefs op de bel *waited in vain for the bell*, waited in vain for the bell (to ring)
16	16	Sara...	zakte door de grond (3.6) *sank through the ground*, cringed with embarrassment
	32	Sara...	groef in de grond *dug in the ground*, dug in the ground

Average decomposability scores (on a scale from 1 to 6) are enclosed in parentheses.

Appendix C. Acoustic prime words Experiment 1

List of acoustic prime words used in Experiment 1

Pair	Prime type	
	Identity	Unrelated
1	boot *boat*	kat *cat*
2	brand *fire*	lijst *list*

(*continued on next page*)

Appendix C (*continued*)

Pair	Prime type	
	Identity	Unrelated
3	dal *valley*	jurk *dress*
4	hand *hand*	tijd *time*
5	hoogte *height*	muziek *music*
6	licht *light*	schroef *screw*
7	man *man*	huis *house*
8	mand *basket*	riem *belt*
9	net *net*	breuk *break*
10	oog *eye*	taak *task*
11	spel *game*	fax *fax*
12	tuin *garden*	heer *lord*
13	wind *wind*	spons *sponge*
14	leven *life*	feit *fact*
15	bel *bell*	koek *cake*
16	grond *ground*	pen *pen*

Appendix D. Materials Experiments 2 and 3

Average decomposability scores (on a scale from 1 to 6) for the complete idioms are enclosed in parentheses in the Stimulus column

Item	Stimulus	Idiom noun	Phon-rel	Sem-rel
1	Jan beet in het (3.6) *Jan bit into the* Jan bit the dust	stof *dust*	stok *stick*	vuil *dirt*
2	Jan liep op zijn (3.5) *Jan walked on his* Jan pushed himself to the limit	tenen *toes*	thee *tea*	vingers *fingers*
3	Jan stond aan het (4.9) *Jan stood at the* Jan was in control	roer *helm*	roem *glory*	mast *mast*
4	Jan zakte door de (3.6) *Jan sank through the* Jan cringed with embarrassment	grond *ground*	grot *cave*	aarde *ground*
5	Jan schoot in de (4.4) *Jan shot into the* Jan hit the nail on the head	roos *rose*	roof *rubbery*	tulp *tulip*
6	Jan leefde bij de (4.3) *Jan lived by the* Jan lived from day to day	dag *day*	dam *dam*	week *week*
7	Jan liep tegen de (2.7)	lamp	land	kaars

Appendix D (*continued*)

Item	Stimulus	Idiom noun	Phon-rel	Sem-rel
	Jan walked against the	*lamp*	*land*	*candle*
	Jan got caught			
8	Jan viel buiten de (3.8)	boot	boon	auto
	Jan fell outside the	*boat*	*bean*	*car*
	Jan was excluded from something			
9	Jan bleef op de (4.4)	been	beer	arm
	Jan stayed on the	*leg*	*bear*	*arm*
	Jan remained on his feet			
10	Jan ging voor de (3.2)	bijl	beits	hamer
	Jan went in front of the	*axe*	*stain*	*hammer*
	Jan gave in			
11	Jan zat in de (3.9)	put	punt	emmer
	Jan sat in the	*well*	*dot*	*bucket*
	Jan was depressed			
12	Jan viel door de (3.5)	mand	map	korf
	Jan fell through the	*basket*	*file*	*basket*(syn.)
	Jan failed			
13	Jan trok aan de (4.3)	bel	bed	gong
	Jan pulled at the	*bell*	*bed*	*gong*
	Jan raised the alarm			
14	Jan ging uit zijn (3.3)	dak	das	vloer
	Jan went out of his	*roof*	*tie*	*ground*
	Jan went extremely excited			
15	Jan stond in zijn (4.2)	hemd	held	broek
	Jan stood in his	*undershirt*	*hero*	*pants*
	Jan was ridiculed			
16 (Experiment 2)	Jan viel in de (4.4)	smaak	smaad	geur
	Jan fell into the	*taste*	*dafamation*	*smell*
	Jan was popular			
16 (Experiment 3)	Jan viste achter het (3.8)	net	nek	hengel
	Jan fished behind the	*net*	*neck*	*fishing rod*
	Jan missed the boat			

References

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database (on CD-ROM).

Bates, D. M. (2005). Fitting linear mixed models in R. *R News, 5*, 27–30.

Bates, D. M., & Sarkar, D. (2005). The lme4 library. <http://lib/stat/cmu/du/R/CRAN/>.

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition, 1*, 343–346.

Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 217–240). North Holland: Elsevier.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*, 668–683.

Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition, 25*(1), 57–71.

Dell, G. (1986). A spreading activation theory of retrieval in language production. *Psychological Review, 93*, 283–321.

Flavell, L., & Flavell, R. (1992). *Dictionary of idioms and their origins.* London: Kyle Cathie.

Gibbs, W. R., & Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology, 21*, 100–138.

Gibbs, W. R., Nayak, N. P., Bolton, J. L., & Keppel, M. (1988). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition, 17*(1), 58–68.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 640–654.

Glucksberg, S., & Keysar, B. (1993). How metaphors work. In A. Ortony (Ed.), *Metaphor and thought* (pp. 401–424). Cambridge, New York: Cambridge University Press.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E.-J. VanderLinden, & R. Schreuder (Eds.), *Idioms. Structural and psychological perspectives* (pp. 133–165). Hillsdale, NJ: Erlbaum.

Katz, J. J. (1973). Compositionality, idiomaticity, and lexical substitution. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 357–376). New York: Holt, Rinehart & Winston.

Kempen, G. (1996). Computational models of syntactic processing in human language comprehension. In T. Dijkstra & K. DeSmedt (Eds.), *Computational linguistics: Symbolic and subsymbolic models of language processing* (pp. 192–220). London: Taylor and Francis.

Kempen, G., & Harbusch, K. (2002). Performance grammar: A declarative definition. In M. Theune, A. Nijholt, & H. Hondorp (Eds.), *Computational linguistics in the Netherlands 2001* (pp. 148–162). Amsterdam: Rodopi.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*, 122–142.

Miller, G. (1991). *The science of words*. New York: Scientific American Liberary.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Nooteboom, S. G. (1999). Sloppiness in uttering stock phrases. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Bailey (Eds.), *Proceedings of the XIVth congress of phonetic sciences* (pp. 683–687). San Francisco.

Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language, 70*(3), 491–538.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191–226). London: Longman.

Peterson, R. R., Burgess, C., Dell, G., & Eberhard, K. M. (2001). Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1223–1237.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 539–557.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken word forms. In A. S. Meyer & N. O. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 115–158). Berlin: Mouton de Gruyter.

Stemberger, J. P. (1982). Syntactic errors in speech. *Journal of Psycholinguistic Research, 11*, 313–333.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18*, 523–534.

Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure and interpretation* (pp. 145–162). Hillsdale, NJ: Erlbaum.

Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics Special Issue: Literal and figurative language, 31*(12), 1655–1674.

Vosse, T., & Kempen, G. (2000). Syntactic structures assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75*, 105–143.