DISTINGUISHING LANGUAGE FROM THOUGHT: Experimental Evidence That Syntax Is Lexically Rather Than Conceptually Represented

Jeffrey S. Bowers,1 Gabriella Vigliocco,2 Hans Stadthagen-Gonzalez,3 and David Vinson2

1University of Bristol, Bristol, England; 2University of Wisconsin-Madison; and 3Rice University

Abstract—It is generally assumed that syntax is represented linguistically rather than conceptually, consistent with the more general view that language and thought are coded separately. This claim is widely defended on logical grounds, but it has received little experimental support. In the present study, we asked Spanish and English speakers to make semantic and syntactic categorizations for pictures and their corresponding names. Consistent with past results, latencies to semantically categorize pictures and words were similar. The new finding is that participants were faster to make syntactic decisions for words compared with pictures, suggesting that syntactic features such as grammatical gender and the count-mass distinction are more closely linked to lexical than conceptual representations.

There is a widespread consensus regarding the main architectural joints of the language production and comprehension systems, as outlined in Figure 1. According to this general framework, language production begins with the construction of a preverbal message that specifies the basic concepts the speaker wishes to communicate. The message triggers language-specific processes that make up the formulation stage. This stage is itself divided into two main components: grammatical and phonological encoding. During grammatical encoding, the system retrieves lexical-semantic representations that encode (or are associated with) syntactic information required for computing the hierarchically organized syntactic structures for a complete sentence (cf. Garrett, 1992; Levelt, 1989). Syntactic features include the grammatical category of words (e.g., noun, verb, adjective), their grammatical function (e.g., subject, object), the kinds of syntactic structures they can be part of (e.g., noun phrase, verb phrase), and more specific lexical features, such as count versus mass (Vigliocco, Vinson, Martin, & Garrett, in press) and grammatical gender (for Romance languages such as Spanish) for nouns. These lexical-semantic codes and their associated syntactic features are called lemmas (Kempen & Huijbers 1983). In the following phonological encoding stage, the phonological codes associated with the lemmas are retrieved, and they are combined in the order specified by the syntactic structure of the sentence computed in the previous stage. According to one view, phonological codes are themselves coded lexically; these codes are called lexemes and, in turn, they map onto phonemes (e.g., Levelt, 1989). According to another view, lemmas directly connect to phonemes (see Caramazza & Miozzo, 1997). In either case, the output of the phonological stage then serves as input to the articulators, which convert phonological plans into overt speech.

The main processing stages involved in production are also found in comprehension. According to Figure 1, the same set of representations is involved in production and comprehension, but it is possible that different sets of representations mediate these two processes (for review of this and related debates, see Bock, 1995).

The assumption that word retrieval during formulation is divided into two main stages—the selection of lemmas followed by the retrieval of phonological knowledge—is supported by a variety of evidence, including observations from slips of the tongue (cf. Dell, 1986; Fay & Cutler, 1977; Garrett, 1976), experimental evidence indicating that the semantic activation of words precedes phonological activation (e.g., Schriefers, Meyer, & Levelt, 1990), and the finding that speakers in a tip-of-the-tongue state can have access to semantic and syntactic (e.g., grammatical gender and the count-mass distinction) features of words in the complete absence of any knowledge regarding their phonological form (e.g., Vigliocco, Antonini, & Garrett, 1997; Vigliocco et al., in press).

However, the distinction between conceptual-level information and lemma-level information is less well supported by experimental evidence. The acceptance of this distinction is largely based on logical arguments for the separation of language and thought (e.g., Fodor, Bever, & Garrett, 1974) and the related commonsense assumption that speakers of different languages think the same way, despite the different ways in which these thoughts are expressed, contra the view of Whorf (1956). But there is little experimental evidence in support of this view, and the separation of language and thought is disputed in some camps (e.g., Slobin, 1996).

The distinction between conceptual- and lemma-level information concerns both whether conceptual knowledge can be distinguished from lexico-semantic information and whether conceptual knowledge can be distinguished from syntactic information. The reader is referred to Schriefers et al. (1990) for some evidence compatible with a distinction between conceptual and lexico-semantic information. The focus of the present study is the distinction between conceptual and syntactic information. In particular, we assess whether syntactic features that do not bear much semantic force nonetheless become part of conceptual structures for adult native speakers and therefore show processing effects similar to those of other conceptual features.

To this end, we contrast speakers’ latencies to make semantic versus syntactic decisions for a set of words and corresponding pictures. Figure 1 schematically outlines how written words and objects are generally assumed to interface with semantic and syntactic representations. In the case of written words, lexical-orthographic codes have direct access to lemma representations that encode syntactic information, whereas conceptual representations that encode various semantic relations are contacted only via the lemma codes. In the case of objects, however, the perceptual outputs feed directly into the concep-
Schematic diagram of the language comprehension and production systems and associated perceptual systems.

**Fig. 1.**

**EXPERIMENT 1**

In Spanish, all nouns are assigned either a masculine or a feminine gender. For nouns that refer to entities with an intrinsic sex, such as chico ("boy") or chica ("girl"), there is a systematic correspondence between the gender of the noun and the sex of the referent. In these cases of so-called natural gender, the gender of a noun can be determined by the underlying conceptual representations. However, for nouns referring to objects and abstract entities, gender is arbitrarily assigned to the noun and is referred to as grammatical gender. For nouns in this class, being feminine or masculine is strictly a linguistic property, and has nothing to do with the conceptual properties of the referent; different languages can assign different grammatical gender to words referring to the same object. For example, milk is feminine in Spanish (la leche) but masculine in Italian (il latte).

It is important to note that the arbitrariness of the relation does not rule out the possibility that syntactic information is coded within the conceptual system. In fact, there are accounts in the literature that make such an assumption. For example, Konishi (1993) asked speakers of Spanish and speakers of German to rate words on the semantic differential scale (Osgood, Suci, & Tannenbaum, 1957) and found that grammatically masculine words were rated higher on semantic dimensions that have masculine connotations, such as power and strength, and grammatically feminine words were rated higher on semantic dimensions that have feminine connotations, such as nurture. Speakers of Spanish and German differed in their ratings for words that have different gender in the two languages (e.g., the word meaning "sun" is masculine in Spanish and feminine in German). On the basis of this correspondence between speakers' semantic ratings and the grammatical gender of nouns, Konishi argued that grammatical gender is intimately related to the conceptual representation for words.

A similar argument was presented by Tawmoski-de Ryck and Verluyten (1982), who noted that information about grammatical gender can be used together with information about the semantic content of nouns in the selection of appropriate referring expressions. For example, the gender of a noun determines the form of a pronoun when there is no pronominal referent in the sentence. For instance, a Spanish speaker who wants to talk about a car would use the masculine pronoun el even when the noun coche ("car") is not produced. If instead the speaker wants to refer to a motorcycle, moto, which is feminine, he or she would use the feminine pronoun la. Because the corresponding noun (coche or moto) is not retrieved in these cases, the authors argued that gender must be represented with the conceptual information about the object and not with the lexical form (see Garnham, Oakhill, Ehrlich, & Carreiras, 1995, for a similar view in comprehension).

Finally, Sera, Berge, and del Castillo Pintado (1994) presented developmental data concerning a task in which speakers of English and Spanish were asked to classify pictures as masculine or feminine. Grammatical gender was used in the classification task by Spanish children in second grade as well as by adult speakers of Spanish. By contrast, English-speaking children and adults tended to use other conceptual criteria, such as assigning feminine gender to natural things and masculine gender to artifacts. Younger children of the two language groups did not differ in their classification (i.e., Spanish children did not show an effect of the gender of the noun until they were in primary school). On the basis of these findings, Sera et al. concluded that although grammatical gender of the nouns does not come into play until relatively late in development, it becomes part of the speaker's conceptual representations for the referents.

However, other findings in the literature seem to indicate that arbitrary syntactic features such as gender are lexically represented (see, e.g., Cacciari, Carreiras, & Barbolini-Cionini, 1997; Meyer & Bock, 1998). For example, Cacciari et al. (1997) showed that Italian speakers were slower in understanding a sentence when the grammatical
Lexical Syntax

gender of an Italian word referring to a human (e.g., *vittima*, “victim,” which can refer to both men and women but has feminine grammatical gender) did not match conceptual information (e.g., when *vittima* was used to refer to a man) than when there was a match between the syntactic and conceptual information. These results indicate that the conceptual and syntactic connotations of gender can be separated, and therefore support the argument that there is a nonconceptual basis for grammatical gender.

In order to provide some additional evidence regarding the relation between grammatical gender and concepts, we asked Spanish speakers in Experiment 1 to complete semantic decision and gender decision tasks in which nouns were presented in written and pictorial formats. As noted earlier, previous research has found similar response latencies for words and pictures in semantic categorization tasks, with pictures sometimes showing a small advantage (e.g., Theios & Amrhein, 1989). We expected to replicate this finding, but the key question concerned the latencies to make gender decisions. In this task, participants indicated the appropriate determiner for each noun by pressing one of two keys on a keyboard; in the semantic decision task, participants indicated whether a noun was an artifact or a natural kind by pressing one of the same keys. We expected that if gender is conceptually represented, we would find no difference between pictures and words in the syntactic task. However, if gender and conceptual knowledge are represented separately, as assumed in Figure 1, we expected participants would be faster to make the gender decision for words than for pictures.

**Method**

**Participants**

Twenty-four native Spanish speakers from the Houston area participated in this experiment.

**Materials and design**

Fifteen masculine and 15 feminine concrete nouns with grammatical gender were selected (see Table 1), and a corresponding set of 30 pictures that clearly depicted the words was constructed. Nouns were selected so that their phonological and orthographic forms did not mark gender; that is, words ending with the vowel *e* or with a consonant were used. This was necessary because the majority of words ending in *a* are feminine (e.g., *la casa*, “the house”) and the majority of words ending in *o* are masculine (e.g., *el carro*, “the car”), and thus for these items, participants could correctly identify the gender on the basis of the word’s phonology (or orthography). The words were displayed in 24-point, standard IBM font. The pictures were taken from a variety of sources and were hand drawn. Sixteen of the pictures depicted artifacts, and 14 depicted natural kinds; this distinction formed the basis for participants’ semantic categorizations. The experiment was run on a Texas Instrument 486/50 4000M portable computer with an active matrix screen, and stimuli were presented with the DMASTER software program developed by K.I. Forster and J.C. Forster, at the University of Arizona.

Each participant completed both the gender and the semantic decision tasks. Half of the participants completed the gender task first and half completed the semantic task first, and there was a minimum of a 2-hr delay between completing the two tasks. In each task, the 30 pictures and 30 words were presented one at a time in a random order. Thus, the experiment included test type (gender vs. semantic categorization) and stimulus type (picture vs. word) as within-subjects factors.

**Procedure**

Participants were tested individually on two different occasions. At the beginning of each session, they were presented with the list of critical pictures and their names in order to ensure there was no ambiguity regarding the names of the pictures. Next, subjects completed one of the two tasks. In the gender task, participants were instructed to press the right shift key as quickly as possible if the word or picture was preceded by the masculine determiner *el* when spoken in a phrase (e.g., *el lapis*, “the pencil”) and to press the left shift key if the word was preceded by the feminine determiner *la* (e.g., *la nariz*, “the nose”). Each item was immediately preceded by a plus sign that was displayed for 500 ms and acted as a fixation point. The same display conditions were used for the semantic decision task, and participants were instructed to press the right shift key if the item referred to an artifact and the left shift key if it referred to a natural kind. Prior to the critical picture and word trials, a set of practice items was included in each test session in order to familiarize the participants with the task.

**Results and Discussion**

A trial was dropped from the analysis if an error was made or if the reaction time (RT) was two or more standard deviations from the sub-

<table>
<thead>
<tr>
<th>Table 1. Nouns used in Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masculine</strong></td>
</tr>
<tr>
<td>Natural kind</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td><em>elefante</em>  (“elephant”)</td>
</tr>
<tr>
<td><em>rinoceronte</em>  (“rhinoceros”)</td>
</tr>
<tr>
<td><em>pie</em> (“foot”)</td>
</tr>
<tr>
<td><em>tomate</em> (“tomato”)</td>
</tr>
<tr>
<td><em>cactus</em> (“cactus”)</td>
</tr>
<tr>
<td><em>maiz</em> (“corn”)</td>
</tr>
<tr>
<td><em>tigre</em> (“tiger”)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The second experiment was carried out in English and exploited the syntactic distinction between count and mass nouns. Count-mass marking affects the choice of phrasal frames for a noun phrase. For example, being count or mass determines the selection of the correct quantifier (*fewer vs. less*) and the correct indefinite determiner (*a vs. some*). In contrast with grammatical gender, the count-mass distinction for English nouns is largely conceptually motivated. Things that occur in single tokens are count nouns; substances and things that occur in bunches are mass nouns. However, there is some arbitrariness. For example, why is knowledge a mass noun and opinion a count noun? Nevertheless, the count-mass distinction contrasts with grammatical gender in that it has some semantic grounding, and thus provides a stronger test of the concept-lemma distinction.

Each participant completed a count/mass task and a semantic decision task in which nouns were presented in written and pictorial formats. In the count/mass task, participants indicated whether each noun was a count or mass noun. In the semantic decision task, participants indicated whether each noun referred to an artifact or a natural kind.

**Method**

**Participants**

Forty English speakers from the University of Wisconsin-Madison participated in return for course credit.

**Design and materials**

Twelve words that refer to natural objects and 12 words that refer to artifacts were selected (see Table 2), and a corresponding set of 24 pictures that clearly depicted the words was constructed. Half of the words in each set were count nouns and half were mass nouns. Words were displayed in 24-point, standard IBM font. The pictures were taken from a variety of sources and were hand drawn. The experiment was run on PC computers using the DMASTER software program.

Each participant completed both the count/mass and the semantic decision tasks. Half of the participants completed the count/mass task first and half completed the semantic task first, and there was a minimum of a 2-hr delay between completing the two tasks. In each task, 24 pictures and 24 words were presented one at a time in a random order. Thus, the experiment included test type (count/mass vs. semantic categorization) and stimulus type (picture vs. word) as within-subjects factors.

**Procedure**

Participants were tested individually. Before the testing began, they were presented with the list of the target pictures and their names in order to ensure there was no ambiguity regarding the names of the pictures. Next, they completed one of the two tasks. In the count/mass task, participants were instructed to press the right shift key as quickly as possible if the word or picture referred to a count noun and to press the left shift key if the word referred to a mass noun. Specifically, they were asked to press the right shift key if the word was an "a word" and the left shift key if the word was a "some word." Each item was immediately preceded by a plus sign that was displayed for 500 ms and acted as a fixation point. The same display conditions were used for the semantic decision task, and participants were instructed to press the right shift key if the item referred to an artifact and the left shift key if it referred to a natural kind. Prior to the critical picture and word trials, a set of practice items was included in each test session in order to familiarize the participants with the task.

**Results and Discussion**

A trial was dropped from the analysis if an error was made or if the RT was two or more standard deviations from the subject’s overall mean. As a result of this procedure, 10.8% of the trials were dropped. In the semantic decision task, participants were slightly faster to respond to words (711 ms) than to pictures (733 ms), $F_1(1, 39) = 4.21$, $MSE = 5,450, p < .05$, and $F_2(1, 23) = 5.11, MSE = 4,533, p < .05$; there were slightly more errors to words (7.3%) than to pictures (4.5%), $F_1(1, 39) = 5.22, MSE = 54.2, p < .05$, and $F_2(1, 23) = 5.43, MSE = 34.2, p < .05$. This pattern of results reflects a small speed-accuracy trade-off, with no evidence of an advantage for pictures or words when both dependent measures are considered. Although this finding contrasts with the results of Experiment 1, in which latencies
to make semantic decisions were slightly shorter for pictures than for words, we note that the literature on the semantic categorization task is quite mixed, with some studies showing a small advantage for pictures over words (e.g., Potter & Faulconer, 1975) and others not showing such an advantage (Theios & Amrhein, 1989).

The critical results, however, concern the count/mass task. Consistent with the gender decision results, the data showed that participants were much faster to make count/mass decisions for words (766 ms) than for pictures (882 ms), \( F(1, 39) = .05, MSE = 5,848, p < .001; F(1, 23) = 145.434, MSE = 1,567, p < .001. \) There also was an interaction between the syntactic and picture-word variables, \( F(1, 39) = 3.708, MSE = 5,725, p = .08, \) and \( F(1, 23) = 4.769, MSE = 3,439, p < .05. \) The difference between pictures and words was greater for mass nouns (150 ms) than for count nouns (76 ms). The basis for this difference is unclear. Errors for words (6.4%) and pictures (7.5%) were not significantly different (\( F_1 < 1, F_2 < 1, \) indicating that the advantage for words over pictures cannot be attributed to a speed-accuracy trade-off. Accordingly, the present results suggest that the count-mass syntactic feature, like grammatical gender, is more closely associated with lexical than with conceptual information.

**GENERAL DISCUSSION**

As noted in the introduction, the distinction between conceptual representations, on the one hand, and lemma-level knowledge, on the other, has largely been based on logical rather than experimental considerations. However, the present findings of (a) faster gender decisions for words than for pictures, (b) faster count/mass decisions for words than for pictures, and (c) faster or equivalent semantic categorizations for pictures than for words provide some experimental evidence in support of this distinction, and the framework depicted in Figure 1 in particular.

These findings challenge the views of Konishi (1993) and Tawmoski-de Ryck and Verluyten (1982), according to whom arbitrary syntactic features such as grammatical gender become part of the conceptual representation for words during the course of language learning. Indeed, our results indicate that a grammatical distinction that has some semantic grounding—namely, the count-mass marking in English—is not incorporated into conceptual representations. Although our conclusions are different from the conclusions of these other researchers, the various empirical results can be reconciled. As noted earlier, these authors based their conclusions on the observation that the grammatical gender of nouns can influence performance on semantically related tasks. We note that these findings can also be explained by assuming that lexically specified syntactic features (e.g., grammatical gender) that have an affinity with important conceptual distinctions (male-female) can be used by speakers in making semantic judgments. So, for example, when asked to rate a word on a semantic scale that includes power as one of the dimensions, a speaker might consider the grammatical gender of the word in order to constrain his or her response, even though syntax is coded separately from semantics.

Thus, our findings provide support for a distinction between features that need to be retrieved in a semantic task and features that need to be retrieved during a syntactic task. More strongly, our study indicates that syntactic features such as grammatical gender and whether a noun is count or mass are strictly linked to the lexical representation for a word rather than to the corresponding concept. In the general framework we described in the introduction, these findings support a distinction between a lemma (a lexical representation that specifies some semantic properties of the word as well as its syntax) and other conceptual correlates, and also challenge views according to which syntactic properties are indistinguishable from concepts (for an overview of various syntactic approaches of this latter type, see Langacker, 1998).

The findings we have reported here, however, are compatible with a number of different architectures for comprehension and production processes: They are compatible with models that assume decompositional representations (e.g., Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Zorzi & Vigliocco, 1999) and with models that assume nondecompositional conceptual representations (Levett, Roelofs, & Meyer, 1999). They are also compatible with models that assume lexical retrieval involves two stages (lemma and lexeme retrieval; Garrett, 1992; Levett, 1989; Levett et al., 1999), as well as with models that include only one step (lemma retrieval followed by sublexical phoneme retrieval; Caramazza & Miozzo, 1997; Dell et al., 1997). Whatever the resolution to these latter debates, the present findings provide some empirical support for the widespread assumption that concepts and lexical representations (thought and language) are represented separately.

**Acknowledgments**—Part of the research reported here was supported by a grant from the National Science Foundation (SBR9711829) to Gabriella Vigliocco.

**REFERENCES**


Vigliocco, G., Vinson, D., Martin, R., & Garrett, M.F. (in press). Is “count” and “mass” information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language.


(RECEIVED 9/3/97; REVISION ACCEPTED 1/13/99)