Generalizable distributional regularities aid fluent language processing: The case of semantic valence tendencies

*Luca Onnis[1], Thomas A. Farmer[2], Marco Baroni[3], Morten H. Christiansen[2], and Michael J. Spivey[4]

[1]University of Hawaii, Honolulu, HI

[2]Cornell University, Ithaca, NY

[3]University of Trento, Italy

[4]University of California, Merced

Running Head: Fluent sentence comprehension

Word count: 8,220

*Corresponding Author:
University of Hawaii at Manoa
Department of Second Language Studies
Center for Second Language Research
493 Moore Hall
1890 East-West Road
Honolulu, HI 96822
email: lucao@hawaii.edu
phone: (808)-956-2782

Abstract

Sentence processing is an extraordinarily complex and speeded process, and yet proceeds, typically, in an effortless manner. What makes us so fluent in language? Incremental models of sentence processing propose that speakers continuously build expectations for upcoming linguistic material based on partial information available at each relevant time point. In addition, statistical analyses of corpora suggest that many words entail probabilistic semantic consequences. For instance, in English, the verb *provide* typically precedes positive words (e.g., 'to provide work') whereas *cause* typically precedes negative items (e.g., 'to cause trouble'; Sinclair, 1996). We hypothesized that these statistical patterns form units of meaning that imbue lexical items, and their argument structures, with *semantic valence tendencies* (SVTs), and that such knowledge assists fluent on-line sentence comprehension by facilitating the predictability of upcoming information. First, a sentence completion task elicited such tendencies in adults, suggesting that speakers constrain their free productions to conform to the connotative meaning of words. Second, fluent on-line reading was slowed down significantly in sentences that contained a violation of a valence tendency (e.g. *cause optimism*). Third, an automated computer algorithm assessed the pervasiveness of valence tendencies in large computerized samples of English, supporting the hypothesis that valence tendencies are a distributional phenomenon. We conclude that not only can aspects of meaning be modeled with word cooccurrence statistics, but that such statistics are likely to be computed by

the human brain during the processing of language. They thus simultaneously contribute to our understanding of the use of language and the psychology of language.

1. Introduction

In ordinary day-to-day human conversation, language comprehension and production under real-time circumstances is extremely fluent, i.e. it is very rapid and yet proceeds effortlessly. Achieving language fluency may appear a trivial feat to most language users, until we consider that it involves the rapid integration of several concurrent types of information cues (sublexical, lexical, semantic, syntactic, and pragmatic) in real time. In addition, given the open-ended nature of language, we all understand and produce novel sentences on a regular basis, such that our ability to pick up linguistic information on the fly must somehow be flexible enough to encompass fluent generativity in both comprehension and production. Given this state of affairs, it becomes relevant to understand the cognitive mechanisms underlying fluent language processing.

In this article, we consider the hypothesis that adult speakers possess implicit knowledge of distributional patterns of words accumulated during years of language usage. We also argue that this accumulated distributional knowledge may facilitate fluency in on-line human sentence comprehension. In particular, we advance the hypothesis that native speakers capitalize on distributional patterns that form units of meaning larger than the word (Sinclair, 1996) in the service of fast and fluent sentence comprehension. One example of the extended units of meaning which we shall consider here can be seen in the observation that the verb *cause* is usually associated with unpleasant words, such as ‘cause problems’, or 'cause trouble' (Sinclair, 1991). Importantly, these extended units of meaning arise from word combinations that are constrained and yet productive at the same time, thus going beyond knowledge of frozen expressions like idioms and collocations. Hence, adult speakers may (at least implicitly) be sensitive to a *generalized* pattern '*cause* + general expectation of an unpleasant word', which they bring to bear as they read or hear sentences in real time. Notice further that the 'core' denotational meaning of these words may not a priori involve a positive or negative reading. There is no reason to assume that *to cause* or *to encounter* anticipate negative words or events. Thus, another intriguing aspect is that the connotational meaning of these words may emerge as meaning distributed over the context of their occurrences in language.

The proposal that certain word distributional patterns may contribute to language fluency is consistent with recent suggestions that on-line sentence comprehension takes place incrementally, and can be driven by expectations made on the basis of the partial linguistic input available at each time step (Altmann & Kamide, 1999; Elman, 1995; 2004; Kutas & Hillyard, 1984; McRae, et al., 2005). For instance, upon hearing the sentence fragment "Yesterday's news caused …" a native speaker of English may have an implicit expectation for a noun phrase that is likely to have a negative connotation, although the specific word to follow is unknown. Therefore, the language processing system may be facilitated in processing the continuation of the sentence "Yesterday's news *caused pessimism* among the viewers" even though that specific sentence or the specific word combination (collocation) ‘cause pessimism’ may have never been encountered before, or has very low frequency in a large sample of English. In this proposal, we refer to this positive or negative character of an implicit linguistic expectation for the predicate of a verb as a *semantic valence tendency* (SVT). Importantly, this latter aspect preserves the generativity of language, while at the same time imposing probabilistic constraints in terms of what to expect for the continuation of a sentence. In the literature there is mounting evidence, discussed below, that humans use expectations as the sentence unfolds in order to reduce the set of possible competitors to a word or sentence continuation. In other words, at each time step the linguistic processor uses the currently available input and the lexical information associated with it to anticipate possible ways in which the input might continue.

It should be pointed out that the case for patterned and extended units of meaning in language is not entirely new. As we detail below, it has been fruitfully exploited in some linguistic circles—in particular, those adhering to usage-based accounts of language. Analyses of large databases of written and spoken language have started to show that most language is patterned, such that word combinations are constrained not only by syntactic but also by lexical factors in very subtle ways. Corpus analyses have also provided initial evidence for SVTs for a relatively small number of words. However, so far these facts have often been confined to linguistic enquiry with little effect on psycholinguistic research. Our first objective is thus to show that the valence tendencies suggested by linguists have a direct impact on sentence comprehension, by way of on-line reading experiments where reaction times are measured. We aim to show that if semantic valence tendencies are important semantic specifications of words and at the same time go beyond single words, then violations of them (for instance '*cause* + a new word with positive valence') should slow down response times significantly in self-paced reading experiments. In this spirit, we aim to help unify the tradition of usage-based linguistics with the tradition of constraint-based psycholinguistics, with the hope of fostering cross-fertilization of ideas between the two areas.

A second new contribution with respect to the original corpus studies is the use of an automated algorithm for evaluating the semantic valence tendency of a word in a psycholinguistic context. We thus explore the possibility that connotative aspects of lexical semantics can be extracted on a distributional basis with simple associative mechanisms, contributing to the growing work in computational linguistics on sentiment analysis (e.g., Pang and Lee, 2004), while at the same time providing evidence that SVTs can be interpreted as a distributional phenomenon.

Before documenting three experiments on semantic valence tendencies in English, we briefly discuss previous relevant work in the two camps of investigation (linguistics and psycholinguistics) that we aim to bring together.

2. The usage-based approach in linguistics

Several linguists have long discussed how native speakers of a language must somehow possess language-specific knowledge that goes well beyond knowledge of syntactic rules and words as single lexical entries in a mental dictionary. The language specificity of certain word-combinations is perhaps most apparent when the expressions for a given equivalent action in two different languages are compared. For instance, the equivalent of *brushing one's teeth* in Italian is *washing one's teeth* (lavarsi i denti). This fact is sometimes referred to as knowledge of native-like selection or "idiomaticity"— the notion that words develop language-specific combinatory potentials. Pawley and Syder (1983) pointed out that certain situations and phenomena recur within a community, thus producing, within that community, standard ways of describing these recurrent 'pieces of reality'. A native speaker of a language will have learned these standard ways of expression, which consist of more than one word or certain clausal constructions. Bolinger (1976) and Hopper (1998) objected to a purely generative approach that stresses the uniqueness of each utterance and thus treating independent utterances as if they were completely novel. Instead, they suggested that everyday language is built up, to a considerable extent, of combinations of prefabricated parts, which Jackendoff (1997) estimated to be comparable in nature to the number of single words.

In line with the claims above, Harris (1998) demonstrated the "linguistic unit" status of the words that comprise popular idioms in English. Participants were presented with either the first two words of popular idioms (*comparing apples*), or two words that are typically adjacent in an idiom but that are in the middle of it (*apples to*), and word recognition times on the final word of the idiom (*oranges*, in either condition) were measured in a lexical decision task. Harris found that in either condition, the priming effect occurred at approximately the same strength as it did for the target words in a series of control conditions where the priming of a target word from a very highly semantically associated prime word was investigated. Through these and other results, Harris argued that all four words of the idioms used in the study, together, comprised one linguistic unit. That is, the presence of two words in a frequently encountered idiom was enough to prime the final word of the idiom. These results suggest that the two-word combinations were entrenched as part of a larger linguistic unit, so much so that the presence of the bigram strongly entailed the other portions of the idiom.

More relevant to the central theme of this present paper, a particularly interesting case of language-specific lexical restrictions on word-combinations is that of extended generalized units of meaning, which we name semantic valence tendencies (related to ‘semantic prosodies’; Louw, 1993; Sinclair, 1991). The interesting aspect of semantic valence tendencies lies in their being potentially productive, and yet constrained at the same time. For example, Sinclair (1991) noted that *cause* and *happen* are associated with unpleasant words (e.g. *cause trouble, accidents happen*). Conversely, *provide* appears to be connoted positively (e.g. *provide work*, Stubbs, 1995). This creates patterns of ‘lexical item + valence tendency'. Table 1 presents a random sample of a query that was conducted for the verb *cause* in the British National Corpus (about 100 million words). Each line represents a fragment of a text in the corpus where the verb is found, and angled brackets indicate the verb + direct object.

------- insert Table 1 about here ------

Although corpus studies represent a very important means of locating patterns that might otherwise go undetected, one limitation is that they explore linguistic patterns in static sentences (already spoken or written) and cannot attest, directly, to the degree that semantic valence tendencies can exert any influence on the time course of on-line sentence processing. Although it has been suggested that stored low-level patterns incorporating particular lexical items 'do much, if not most of the work in speaking and understanding' (Langacker, 1988), this has largely remained a speculation with scant experimental evidence from human processing data (but see McDonald and Shillcock, 2003 for effects of collocational strength on reading).

Thus, one outstanding question that is left unanswered regarding semantic valence tendencies is their psychological status, and thus, their impact on on-line sentence comprehension. In addition, one important feature of SVTs is that they are not as lexically restrictive as idiomatic (or unitized) expressions such as *brush your teeth*. Semantic valence tendencies, instead, do not appear to restrict lexical choices too narrowly because they are not entirely fixed. For instance, '*to cause pessimism'* may be a relatively low frequency word combination even in extraordinarily large collections of language such as the World Wide Web, but its acceptability by native speaker standards may derive from its conforming to the general negative valence tendency of *cause*. This argument, however, is hard to support by simply examining corpus data, because corpora often contain counterexamples, and may be subject to sampling skewness. As we shall see, the generativity of SVTs can be better tested by on-line sentence processing methods that employ Reaction Times (RTs) as a measure of fluent and disfluent processing. For this reason, we now turn to introduce psycholinguistic literature relevant to our studies.

3. The constraint-based approach in psycholinguistics

Why should semantic valence tendencies be relevant for on-line sentence processing? In psycholinguistics, increasing interest has been directed to the way language is statistically patterned in order to explain how comprehenders construct an understanding of what they hear or read in real time. Possibly because of an educational bias toward the printed word, we tend to think of sentences as static and complete entities, like this page of text. In fact, both in speaking/hearing and in reading, language necessarily unfolds in real time as each word is heard or read. Sounds within a single word unfold in time and have their specific time course (Gaskell & Marslen-Wilson, 2002; Marslen-Wilson, 1987). Incremental models of language comprehension (e.g., Altmann and Kamide, 1999; Tanenhaus et al., 1989) propose that the hearer does not wait until the end of a clause or of a structural element in the sentence but instead makes predictions about what is most likely to come next at each time step. Using eye-tracking techniques, this work demonstrated that when processing a target item (e.g., hearing the word “candle”), comprehenders will often make brief eye movements not only to the correct referent object displayed in front of them (a candle) but also to another object displayed whose name bears phonological similarity to the target item (e.g., a candy. Allopenna et al., 1998; Spivey-Knowlton et al., 1998; Tanenhaus et al., 1995). Allopenna et al. also found that soon after its acoustic offset, looks to the candy decreased while looks to the candle continued to increase. This suggests that, as the target word unfolds in real time, both “candle” and “candy” are activated during language processing, but that as soon as information is available to eliminate the wrong competitor, the linguistic processor uses it readily.

Strong expectations about upcoming linguistic material exist not only for sublexical fragments but also for entire words of a sentence as the sentence unfolds in time. In Altmann & Kamide (1999), participants were shown a scene portraying a cake, a toy car, a boy, and a ball. They launched saccadic eye movements significantly more often at the cake when they heard *The boy will* *eat…* than when they heard *The boy will move…* These data suggest that the processor immediately applies the semantic constraints afforded by the verb's selectional restrictions to anticipate a forthcoming postverbal argument. Other results suggest that expectations are made on the basis of information such as the typicality of thematic roles (McRae, et al. 1997) and the degree to which the nouns associated with a verb's arguments are typical agents and/or instruments for the verb (McRae et al., 2005). In our example of the verb *cause*, the negative semantic valence tendency can be seen as another dimension of semantic selectional restrictions imposed on the verb, but these kinds of restrictions have never been tested before. In addition, what is not known at present is whether the verb has a dominant role in directing sentence interpretation. Semantic valence tendencies are a particularly interesting test case for incremental models because they seem to apply not only to a verb's argument structure, but also to all word categories (cf. Barker & Dowty, 1993). Many adjectives and adverbs whose definitions do not carry any evaluative component seem nonetheless to involve favorable or unfavorable semantic valence tendencies. For instance, from one preliminary corpus analysis we performed, the adverb *perfectly* exhibited a distinct tendency to co-occur with 'good things': *capable, correct, fit, good, happy, harmless, healthy, lovely, marvelous, natural*, etc. *Utterly*, on the other hand, has collocates such as *helpless, useless, unable, forgotten, changed, different, failed, ruined, destroyed, etc.* (Stubbs, 1995). Hence, one novel prediction is that the human processor will selectively anticipate different semantic groups of adjective continuations in the two sentence pairs below:

[1] Given her curriculum, it appears that our applicant is utterly…

[2] Given her curriculum, it appears that our applicant is perfectly…

where *utterly* and *perfectly* are the prime words. Given the current predominance attributed to the verb and its arguments in assigning structure and interpretation to sentences in psycholinguistics (Altmann & Kamide, 1999), it would be a significant contribution to show that the language processor can use any type of linguistic material to start interpretation, and this may occur as early as the first word, as in *Clearly…the cook was not at his best today*. Conversely, if results of semantic valence tendency sensitivity were found only for verbs (e.g. *cause*) and their arguments, and not for, say, adverbs (e.g. *perfectly*), this would lend support to current theories on the predominance of the verb, at least for English.

Sinclair (1996) has proposed that constructions like semantic valence tendencies may constitute 'units of meaning' in the sense that they constitute single lexical choices on the part of the speaker/hearer, despite the fact that they can be segmented into individual words and each word can be described in a separate entry in a dictionary. This opens up the possibility that lexical knowledge is not a list of single words in a mental dictionary, but instead a network of complex units of meaning that interact with the structure of the sentence in on-line processing (e.g., Elman, 2004).

4. Experiment 1: Elicitation of SVTs by sentence completion

We conducted an exploratory sentence completion experiment to determine the valence of a group of words proposed by corpus linguists to have clear semantic valence tendencies. That is, large corpus analyses that examined the collocates of these words suggested a strong connotative orientation for each one of them. Throughout, we shall call these words ‘primes’, because their role as primes for the next ensuing word was estimated. Priming is widely used in psychological research to explore the nature of underlying cognitive processes. The basic idea is that a preceding stimulus, for instance, a particular word or sentence, increases the likelihood that the hearer will access a related word or sentence. Alternatively, the prime word also reduces the time it takes to process the related word (for instance by facilitating its reading) as compared to an unrelated control word. In Experiment 1, we used priming in an elicited production task, while in Experiment 2, we examined RTs for a given word as measure of priming. Although the specific interpretation of the priming effect may depend on a particular theoretical stance, priming is widely accepted as a sign of fluent association or processing facilitation between two words or stimuli.

4.1 Method

4.1.1. Participants. 24 Cornell undergraduate students participated for course credit. All were native speakers of English and had no reported language disability. Nineteen students participated in a Sentence Completion Task and 5 students participated in a Fragment Rating Task (see below).

4.1.2. Materials and design. Twenty-two word primes were used as stimuli in the experiment, 5 with a proposed positive orientation (*to provide*, *perfectly*, *pure*, *profoundly*, and *known for*), and 16 with a proposed negative orientation (*to cause*, *to harbor*, *to incite*, *to encounter*, *to peddle*, *to be bent on*, *clearly*, *to commit*, *deeply*, *to express*, *to be involved in*, *markedly*, *to be notorious for*, *patently*, *to reveal*, *sheer*, and *utterly*). The primes were a combination of verbs, nouns, adjectives, and adverbs. In the Sentence Completion Task participants were asked to complete sentences where the prime appeared as the last word. For example, given the incomplete sentence "I believe that 20th Century philosophers have peddled..." participants were asked to write down a plausible ending to it, with no particular restrictions other than not to think too long about any given sentence. This allowed us to elicit semantic valence tendencies for the sentence continuations. In particular, since the context preceding the prime (*peddled* in the example above) was chosen to be as neutral as possible in terms of connotational value, the main influence on participants' choice of sentence continuations could be attributed to the prime words. A set of 18 filler sentences were also included, such that each participant completed a total of 40 sentences. The order of sentences was randomized for each participant.

At the end of the experiment, sentence continuations for the trial sentences were collected (filler sentence continuations were discarded) and the shortest number of words to the right of the prime that formed a self-contained phrase were included in a Fragments List of sentence continuations. For instance, one participant completed the sentence “I believe that 20th century philosophers have peddled…the same crap as other philosophers.” The phrase ‘…the same crap’ was retained and included in the Fragments list. Because they were elicited immediately after the prime words, these fragment phrases should capture something of the spontaneous semantic valence tendency of a prime. For each given prime, 19 fragment continuations were collected (corresponding to 19 participants), and the complete Fragment List consisted of 19 x 23 = 437 Fragments.

The five participants who had not participated in the Sentence Completion Task participated in the Fragment Rating Task. They were asked to rate each phrase in the Fragment list for their valence on a scale between –3 and + 3, where 0 was neutral on a 7-point-Likert-scale. For example, one participant rated -3 the phrase ‘the same crap’ as having a very negative valence. Since they were unaware of the beginning of the sentences containing the prime word, these ratings were taken as an independent evaluation of semantic valence tendency.

4.1.3. Procedure. Participants sat in front of a PC in a quiet room. In both tasks, sentence or fragment trials appeared one at a time on the screen and participants were asked to write down on a sheet of paper either a continuation (Sentence Completion Task) or a rating (Fragment Rating Task). The experiment lasted no longer than 40 minutes.

4.2. Results

In total, 2,185 separate ratings were collected (19 fragment continuations x 23 primes x 5 participant raters). Ratings were collapsed and averaged over the 23 primes, such that each prime had a mean value of its semantic valence tendency. It was hypothesized that if a given prime (e.g. *harbor*) displayed a consistent valence tendency, this would show up as a robust positive or negative mean rating.

A Mann-Whitney test performed on the 23 primes divided in two groups (positive or negative) revealed a significant difference between the two groups, $z(21)=3.29$, $p<0.001$. This result suggests that words in the positive group were judged consistently more positively than words in the negative group (see Table 2).

Overall, the results of Experiment 1 suggest that adult speakers spontaneously generated sentence continuations that were consistent with the semantic valence tendencies proposed by corpus studies for our list of 23 primes. Furthermore, in the Sentence Completion task there was considerable variation in the continuations of sentences, suggesting that the semantic valence tendency of a word manifests itself as a broadly *generalized* preference for positively or negatively oriented companion words.

----- insert Table 2 about here ----

5. Experiment 2: On-line sentence processing of semantic valence tendencies

Experiment 1 provided initial evidence that speakers possess some knowledge of what is the most natural continuation of a sentence given the semantic valence tendency of a word. Importantly, participants' choices were quite idiosyncratic, and in only a few cases did sentence continuations overlap substantially across participants for the same given sentence. This implies that the preceding contexts allowed considerable free choice, and that participants did not pick the most frequent frozen collocation to complete the sentence. And yet most continuations displayed a clear orientation toward a specific connotative valence. It is possible to conceive of semantic knowledge as a high-dimensional state space (Rogers & McClelland, 2004; Vigliocco et al. 2004) in which each word in a sentence contributes to creating a dynamic trajectory that preferentially directs sentence interpretation in certain regions of the space, and not others. Thus, the choice of an adverb, say *perfectly* (as opposed to, say, *utterly*) already contains a statistical hint to express a positively oriented predicate that applies to the object being predicated, as in this actual continuation from Experiment 1:

[3] It seemed like the firm was perfectly…prepared for the new case.

Furthermore, these results from elicited production (Experiment 1) lead to a new hypothesis. The presence of semantic valence tendencies may not only facilitate language production, but may also serve to facilitate language comprehension in real-time situations. If producing a given word in a sentence, say the verb *to encounter*, implies that the producer has already narrowed down to some extent the set of possible sentence continuations she may utter, then the receiver's sensitivity to this semantic valence tendency will help him anticipate the sentence continuation, with a measurable gain in fluent comprehension.

In Experiment 2, we thus set out to investigate whether the reading of words such as *cause* can prime their negative semantic valence tendency in the form of an implicit expectation for a range of upcoming words. Consider the following sentences:

[4] The mayor was surprised when he encountered refusal from his constituents regarding the new road improvement plan.

[5] The mayor was surprised when he encountered consent from his constituents regarding the new road improvement plan.

In [4], the prime *encountered* precedes a word that is consistent, in terms of its polar bias on the negativity-positivity dimension, with its predicted negative valence (refusal), while in [5] *encountered* precedes an inconsistent valence word (consent), yielding an inconsistent prime-target pairing. If it is the case that the semantic valence of a prime word aids in the creation of an expectation about the nature of the information to follow, then one would predict that RTs, as measured by the amount of time participants spent reading each word of a sentence, would increase significantly when the target is inconsistent with the semantic valence tendency of the prime than when it is consistent. In the study detailed below, we tested this prediction in the context of a controlled experimental design.

5.1 Method

5.1.1. Participants

Twenty-eight Cornell undergraduate students participated in a self-paced reading task for extra credit in a psychology course. All participants were native speakers of English and had no reported language disability.

5.1.2. Materials and design

A subset of six prime words from Experiment 1 were used here to generate the experimental sentences: *to cause*, *to incite*, *to peddle*, *perfectly*, *to harbor*, *to encounter*[1]. For each prime word, two sentences were constructed, yielding a total

[1] Five other primes were originally included in the materials but could not be used: *to provide* and *patently* had repetitions due to typing errors in the program that precluded a proper analysis of RTs. *To be known for*,

of 12 experimental sentences across the six experimental-sentence frames. One sentence contained a consistent prime-target pairing, and the other contained an inconsistent prime-target pairing, as in examples [4] and [5], respectively. The initial portion of each sentence, all the way up to the onset of the target word, was held constant across the consistent and inconsistent versions of each experimental-sentence frame in order to ensure that any observed processing-related differences could not be attributable to different sentence-initial contexts. Additionally, the beginnings of both sentences in each of the six sentence-frames were designed to be neutral, in terms of their valence, in order to avoid introducing a bias in the nature of the event depicted in each sentence that might favor a downstream positive or negative continuation of the sentence after the prime word.

----- insert Table 3 about here ----

We aimed to control for several concomitant factors that have been shown to influence the speed with which the words of sentences are read. At the sentential level, for example, we conducted a plausibility norming study in order to ensure that the sentences containing consistent prime-target pairings were not significantly more plausible than the sentences containing inconsistent prime-target pairings. Sixteen separate native English-speaking Cornell undergraduates rated sentences for plausibility on a seven-point Likert-type scale (7=Very

to be notorious for, *to be bent on*, and to *be involved in* are all multi-word fragments where the word prior to the target was not the prime (e.g., *known*) but a very common proposition (e.g. *for*). This again precluded a clear analysis of what to consider as prime.

Plausible). Two lists were constructed. One list contained six sentences with consistent prime-target pairings and six sentences with inconsistent prime-target pairings, but only one version of each sentence frame, and a second contained the opposite version of each sentence frame. That is, for each word prime embedded in an item-frame, raters saw only one of the two possible sentence continuations (beginning with, of course, the consistent or inconsistent target word). Additionally, 20 unrelated filler items were included, and participants were randomly assigned to receive one of the two lists. There were no significant differences in overall plausibility ratings between the sentences containing consistent and inconsistent prime-target pairings, $t(5) = .85$, $p = .434$ (the by-condition means and standard deviations on this and all other control variables can be found on Table 3).

At the word level, no significantly reliable differences existed between the consistent and inconsistent prime-target pairings (for each item) in the overall length, in characters, of the target words, $t(5) = .54$, $p = .61$, the frequency of the target words (as evident by frequency counts extracted from the BNC), $t(5) = .67$, $p = .53$, or the associated log-frequency of the targets, $t(5) = 1.10$, $p = 0.321$). Additionally, the frequency of the prime-target bigrams were very low, as estimated on a Google search over the World Wide Web[2]. This ensured that the prime-target pairs were a relatively new combination in both the consistent and

[2]Because the occurrence of specific word combinations is quite rare even in relatively large corpora (Zhu & Rosenfeld, 2001), such as the BNC, we used Google-based frequencies to overcome this data sparseness problem. Although web-based word co-occurrence frequencies incorporate a certain amount of noise, the resulting frequencies are not only highly correlated with BNC frequencies (when available), but provide even better correlations with human plausibility judgments than do BNC frequencies (Keller & Lapata, 2003).

inconsistent sentences, such that any differences in reading times could not be easily attributed to familiarity with specific word collocations. Notably, there was no reliable difference in log-frequency between consistent and inconsistent prime-target pairs, $t(5) = 1.277$, $p=0.230$.

The 12 sentences were counterbalanced across two different presentation lists in such a way that each participant saw six sentences in each of the two conditions, but saw only one version of each of the six sentence frames. The items were presented along with 40 unrelated filler items and eight practice items.

5.1.3. Procedure

Participants sat in front of a PC in a quiet room, and were randomly assigned to one of the two presentation lists. All sentences were presented randomly in a non-cumulative, word-by-word moving window format (Just et al. 1982) using Psyscope version 1.2.5 (Cohen et al. 1993).

Participants initially viewed a brief tutorial designed to acquaint them with the task. Participants were then instructed to press the ‘GO’ key to begin the task. The entire test item appeared on the center (left-justified) of the screen in such a way that dashes preserved the spatial layout of the sentence, but masked the actual characters of each word. As the participant pressed the ‘GO’ key, the word that was just read reverted to dashes and the next word appeared. The computer recorded RTs in milliseconds for each word presented. After each sentence had been read, participants responded to a Yes/No comprehension question, and upon another key press, the next trial began.

5.2. Results and Discussion

As illustrated in Figure 1, although RTs were relatively similar at the prime word of each prime-target pairing across each condition, RTs were substantially higher on the target word in the inconsistent prime-target pairing condition than they were in the consistent prime-target pairing condition. That is, as predicted, an increase in RTs occurred from prime to target when the bias of the target word (on the negativity-positivity dimension) was inconsistent with the semantic valence tendency of the preceding prime word, but not when a consistency was present in the word-pair. A 2 (consistent vs. inconsistent) x 2 (prime vs. target) repeated measures ANOVA yielded a significant two-way interaction, $F(1,27) = 4.679$, $p = .039$, indicating that the increase in RTs from the prime word to the target word was dependent upon the consistency status of the prime-target pairing. Indeed, follow-up paired sample *t*-tests revealed a statistically reliable increase in RTs from the prime to the target for the inconsistent prime-target pairing condition, $t(27) = 3.475$, $p = .002$, but not for the consistent pairing condition, $t(27) = 2.254$, $p > .05$.

These results show that, as predicted, participants exhibited sensitivity to the incongruence of semantic valence tendency between the prime and the target in the inconsistent condition. More specifically, they suggest that at the time of reading the prime, expectations about subsequent words are generated, and can encompass general biases toward an expected semantic valence tendency of a word. As noted in the introduction, such a result is consistent with expectation-based and constraint-based accounts of sentence processing, where information is taken up incrementally and continuously as a sentence unfolds in time.

----- insert Figure 1 about here ----

6. Experiment 3: Corpus analyses and algorithm

We have argued that SVTs are not the consequence of denotational factors (there is no intrinsic semantic reason why, say, *reveal* should tend to be associated with negative words while *provide* is associated with positive words). Therefore, semantic orientation may be the product of usage-based distributional generalizations: *reveal* is connotated negatively because it typically occurs with negative words, and language learners pick this statistical generalization. Our interpretation of SVTs leads to the prediction that it should be possible to model them in terms of corpus-based distributional patterns.

The pioneering studies on corpus linguistics deserve the merit of having highlighted the potential importance of word distributional patterns, such as the semantic valence tendency phenomena studied here, for language use. However, evidence for SVTs has been limited to a handful of examples, and it has typically rested on procedures of ‘eyeballing’ sample concordance lines from corpora (very similar to our sample Table 1). Little effort has been made in producing statistical analyses to support the robustness of the evidence, or to empirically assess the direction and strength of the SVT associated with a word (but see Hoey, 2005 for tighter empirical analyses). Accordingly, in order to further assess the potential utility of SVTs, we also tested simple computational procedures, based on word distributions, for the automatic extraction of the strength and direction of a word's semantic valence tendency. Thus, we looked to the literature on computational linguistics and information retrieval. Sentiment Analysis has recently been a very active area of research in these fields (e.g., Pang and Lee, 2004), and various algorithms to discover the semantic orientation of words have been proposed.

In Experiment 3, we piloted a semi-automated algorithm for the extraction of semantic valence tendencies based on Turney & Littman (2003), who introduced a method for automatically inferring the direction and intensity of the semantic orientation of a word from its statistical association with a set of positive and negative paradigm words. We asked whether the algorithm could assign a semantic orientation to the primes used in Experiment 1, thus supporting our hypothesis that SVTs are a distributional phenomenon to which learners become sensitive by being exposed to language.

6.1. Method

The algorithm was tested on 21 word primes. The semantic valence tendency S*VT* of a prime word (e.g. *to harbor*) was calculated from the strength of its association

A (see Equation [a]) with a set of positive words (Pwords) minus the strength of its association with a set of negative words (Nwords) (Turney and Littman, 2003):

[a] $$SVTA(prime) = \sum_{pword \in Pwords} A(prime, pword) - \sum_{nword \in Nwords} A(prime, nword)$$

The Pwords used, taken from Turney and Littman, were: *good, nice, excellent, positive, fortunate, correct,* and *superior*. The Nwords, also taken from Turney and Littman, were: *bad, nasty, poor, negative, unfortunate, wrong,* and *inferior*.

The strength of association *A* was calculated using Pointwise-Mutual Information (PMI, Church & Hanks, 1991). PMI can be interpreted as the ratio between the probability of seeing a prime with a positive/negative word in its context, and the probability of co-occurrence of the prime and a positive/negative word under independence (see Equation [b]):

[b] $$PMI(prime, pword) = \log_2 \frac{p(prime \,\&\, pword)}{p(prime)\,p(pword)}$$

Hence, the semantic valence tendency S*VT* of a prime calculated using PMI is as in Equation [c]:

[c] $$SVT\_PMI(prime) = \sum_{pword \in Pwords} PMI(prime, pword) - \sum_{nword \in Nwords} PMI(prime, nword)$$

Co-occurrence and single word probabilities were estimated calculating the number of hits on automated Google searches, thus using the World Wide Web as a large corpus to circumvent problems of data sparseness (Keller & Lapata, 2003; see Mittelberg et al. 2007, for a discussion). Word forms that could be ambiguously used in different word categories were eliminated. For example, for the verb *to harbor*, we retained the forms *harboring*, and *harbored*, and excluded the forms *harbor*, and *harbors*, which can also be used as nouns. This type of manual filtering was necessary because the noun *harbor* (i.e., port) does not necessarily prime negative words in its immediate context.

6.2 Results

Grouping word primes in two groups (positively and negatively oriented), a Mann-Whitney test indicated that the difference between the two groups was significant, $z(20)=3.73$, $p<.001$. This result suggests that the algorithm assigned words in the positive group consistently higher values of semantic orientation than words in the negative group. There was a perfect ranking, in that even the lesser positively oriented word was ranked above the lesser negatively oriented word (see Table 4). Overall, the results of Experiment 3 suggest three tentative but important considerations. First, the associative algorithm of Turney & Litman (2003) can be extended to infer the semantic valence tendency of words whose denotative meaning does not appear to signal a specific positive or negative orientation. For example, it is not a priori intuitive that the verb *to encounter* is associated with negative events. One reading of our results is thus that the connotative meaning of words arises from contextual use. In addition, the algorithm is sensitive to differential distributional uses of near-synonyms, such as *pure* versus *sheer*. The specific SVT_PMI value for *perfectly* (which was labeled positive, according to corpus studies) was –2.16, while *utterly* (which was labeled negative) had a value of -5.46. Likewise, in accord with preliminary 'eyeballing' concordance lines for the near-synonym adverbs *largely* and *markedly, largely* turned out to be more positively oriented (SVT_PMI= -1.85) than *markedly* (SVT_PMI= -2.46)[3].

A second consideration is that the algorithm used was successful at predicting semantic valence tendencies, despite its being a distributionally approximate method. The co-occurrence between a given prime and each of the Pwords and Nwords was calculated within a window of the whole text. Thus, given a very large corpus, and despite considerable noise in the sampling, the semantic valence tendency of a word can be extracted to a sufficient precision by a simple distributional analysis of the text environment.

A third consideration pertains to the psychological implications of our modeling efforts. From a psycholinguistics perspective, the algorithm suggests that native speakers would have enough evidence on a purely distributional basis to develop intuitions on the connotative dimensions of words without strong

[3] Note that here what counts as positive versus negative is not an absolute value above or below zero, but the relative value of two words compared to each other. In the the Mann-Whitney test, which uses a relative ranking procedure by ordering the words in descending order of value, there was a perfect ranking, in that all words labeled as positive appeared in the top rankings above all negatively labeled words.

denotational orientations (cf. *to encounter*, *to cause*, *largely*, *to consider*). Such intuitions can be developed on the basis of being exposed to distributional co-occurrences of the words in question with more clearly oriented positive and negative words (in our experiment exemplified by a few prototypical Pwords and Nwords).

7. General Discussion

The primary issue addressed in this study is the degree to which statistical structure of the mental lexicon can affect sentence processing. We have investigated the manner in which distributional patterns of co-occurring words may form units of meaning, on which native speakers capitalize in order to produce and understand language. We have focused on the tendency of words to be associated with other words connoted positively or negatively, as evidenced by corpus studies. In Experiment 1, native speakers of English provided sentence completions that were consistent with the semantic valence tendency of the last word of a given initial sentence fragment. This is evidence that speakers are sensitive to the general semantic orientation of a word, and thus naturally constrain their production to calibrate this knowledge, while concurrently they freely choose many different sentence continuations. We speculate below on the implications of this concurrent job of productivity and constraint.

In Experiment 2, we provided the first empirical results of lexical priming in sentence comprehension due to semantic valence tendencies. From the perspective of the receiver, knowledge of what lexico-semantic constraints are imposed on sentence continuations may help to facilitate fluent processing by creating an *implicit expectation* of possible word continuations. In Experiment 2, readers were significantly slower at processing words that violated the semantic valence tendency of a given word. These data support a view of sentence processing as a complex task involving an incrementally unfolding interpretation of words within their relevant context. At each point in time, expectations of likely upcoming material are computed based on partial information. Expectations can be seen as multiple probabilistic constraints internalized by the linguistic processor (MacDonald et al. 1994), and we have shown that semantic valence tendencies are one such constraint that can contribute to real-time fluent language processing. Finally, in Experiment 3 we have shown that it is possible to measure the semantic orientation of a word by a simple distributional analysis carried out over a large sample of language, thus providing an "existence proof" for the hypothesis that semantic valence tendencies can be induced from distributional patterns.

In the remaining portion of this paper, we consider some of the implications of our work, as well as limitations of the current studies. One contribution is that distributional information revealed by corpus studies was here shown to have a direct impact on mechanisms of sentence processing, and thus adds considerable psychological reality to these phenomena. Not only do semantic valence tendencies tell us a fact about the conventional usage of a language, they also tell us a fact about the human machinery that processes language, and thus have important implications for linguistics, computational linguistics, and psycholinguistics.

Another important aspect of the current work regards the preservation of generativity. Work on co-occurrence statistics (e.g. selectional restrictions in computational linguistics, Brent, 1991; collocations in corpus linguistics) is often perceived as involving mere lexical constraints. Psychologically, these phenomena are often regarded as peripheral in explaining language processing because they are assumed (simplistically, we would argue) to be dealt with by processes of rote memorization. On the contrary, we argue that the types of distributional patterns we have investigated afford the language system the necessary fluent generativity to understand and produce not only crystallized collocations (e.g. 'to cause damage' which has a high co-occurrence and is probably learned by rote), but also novel sentences and word combinations that conform to the general semantic valence tendency of a given word. This was shown to be true because the prime-target pairs in Experiment 2 had low probability of co-occurring in a very large corpus such as the Web. In both the linguistic and psycholinguistic traditions, generativity and constrained lexical selection have often been constructed as two opposing facets of language, one being the product of syntactic machinery, the other the product of associative memorization in the lexicon. We speculate here that in regard to semantic valence tendencies, we seem to be dealing with a sort of 'constrained semantic generativity' that emerges from the same statistical machinery that analyzes the linguistic environment. Although we have not yet provided a mechanistic account of how semantic valence tendencies could be learned, it is possible that the same statistical mechanism that is sensitive to individual collocation strengths (e.g. *cause problems*, *cause delays*, *cause troubles*, etc.) eventually accumulates enough evidence for a given word (e.g. *cause*) to compare the semantic distance between all the predicates that most frequently collocate with it (*problems*, *delays*, *troubles*, etc.), and to eventually find that the majority is close to the semantic dimension of negativity in hyperdimensional semantic space.

Our hypothesis of extended generalized units of meaning has further bearing on the nature of the bilingual brain. Many late second language (L2) learners attain high levels of language knowledge, and yet often produce sentences that sound 'non-native' (Pawley & Syder, 1983), such as 'Although tourism *causes economic improvement*, its operational costs must also be considered'. In this case, a Chinese L2 speaker appeared unaware of an extended unit of meaning 'cause + unpleasant word', whereas what he/she meant might have been rendered more naturally as 'Although tourism *leads to* economic improvement, …' arguably because *lead to* has a more neutral semantic valence tendency (this intuition can be checked against a corpus of English, see Sinclair, 1991). Even very proficient L2 speakers lag behind native speakers specifically in the degree of knowledge of language-specific selectional restrictions, and there is evidence that a correlation exists between language skill and fluency and knowledge of language-specific phraseology (Howart, 1998; Onnis, 2001). In work in progress, we are investigating whether late L2 learners may lack a great deal of language-specific knowledge about extended generalized units of meaning which impacts fluent on-line sentence processing. This should become particularly evident when the semantic valence tendency for cognate words with similar denotational meaning is different between languages, for instance the adjective *impressionante* in Italian is connoted negatively whereas *impressive* has a positive connotation in English. A few authors have highlighted how learning the different connotations of these pairs of cognate words in two languages may be hard for L2 learners (for English/Portuguese, see Sardinha, 2000; for English/Italian, see Partington, 1998). This fact has direct relevance on teaching practices of L2. Although most L2 teaching curricula now recognize the importance of what is not only grammatical, but also conventional, for speaking a foreign language, the focus is generally on frozen idiomatic expressions and collocations (Bahns & Eldaw, 1993; Lewis, 2000), and may overlook the existence of extended generalized and productive units of meaning. Even authoritative dictionaries and thesauri compiled by expert lexicographers often fail to recognize such semantic valence tendencies of words. Our statistical analyses of very large linguistic databases (Experiment 3) and our pilot psycholinguistic data (Experiments 1 and 2), however, suggest that several words may possess language-specific semantic valence tendencies that determine preferences for certain semantic sets of words.

From a methodological point of view, our study indicates that behavioral evidence and corpus-based computational analysis can be used as converging tools for the study of human cognition. It is particularly interesting that this also holds in a “connotational” domain such as semantic orientation, traditionally linked to human emotion more than to logical faculties. This suggests that distributional methods might have a wider relevance than what is sometimes claimed (e.g., French and Labiouse, 2002).

Before concluding, we would like to point out several limitations of the current work, which are currently being addressed in work in progress. One potential criticism of Experiment 2, in particular, concerns the relatively limited number of items administered to participants. This concern is indeed valid because it influences the generalizability of the effect to other items not used in this present study. That is, one might argue that the observed by-condition RT differences are specific to the very few prime-target tokens used here. Given the relatively specific nature of the items used in both study one and study two, and given the degree of linguistic control necessary in order to afford the ability to make valid inferences from the RT data, it is, of course, quite difficult to generate meaningful and usable sentence frames. A challenge for future research is to identify more words that have been hypothesized to contain some sort of semantic valence, and to systematically examine the effects of SVT violation on production and comprehension of downstream information.

More generally, our positively and negatively connotated forms have been selected based on the corpus linguistics literature and our own intuition. Future work should provide a more formal and controlled way to choose stimuli charged with semantic valence tendency.

Additionally, although the data here reveal a detrimental effect of inconsistency between the prime and the target, as evident in the increase in RTs from prime to target in the inconsistent word-pair condition, it is fair to consider why the opposite effect was not also observed for the consistent prime-target word-pairings. That is, if the SVTs of the prime words are facilitating the predictability of subsequently occurring word-forms, then an additional prediction might be that RTs should decrease in magnitude from the prime to the target in the consistent word-pairs, indicating that SVTs can actually facilitate on-line processing as well. As evident in Figure 1, however, such a trend was not observed. One potential cause for the lack of a facilitation effect in the RT data provided here might very well be that something of a "floor effect" occurred in the RTs associated with the sentence materials. Self-paced reading is a technique that affords the researcher one, maybe two, data points (button presses) per second. Therefore, when participants are reading simple sentences with no relevant (increase-evoking) anomaly, one might expect RTs to fall within the range observed here. That is, although some small beneficial facilitation effect might very well exist in the consistent prime-target pairings, the relatively coarse-grained temporal sensitivity of the self-paced reading technique might not allow for the observation of it. In future research, one might consider using techniques with better temporal sensitivity, such as the tracking of eye-movements while reading or the examination of the event-related potentials (ERPs) associated with the onset of "consistent" target words, in order to better understand the types of effects SVTs have in both the consistent and inconsistent prime-target word-pairs.

Finally, we decided to use Turney and Littman's algorithm because it is straightforward to implement, almost knowledge-free (only requiring a short list of good and bad "seed words") and effective. However, in future work we would like to explore other methods that would make SVT induction more cognitively plausible. In particular, we want to develop procedures that do not require hand-picked seeds, and that will be effective on input that is more similar to the one that children hear and read during language acquisition (e.g., corpora of child-directed speech and/or written materials used in primary education).

8. Acknowledgments

This work was supported by Grant # 5R03HD051671-02 from the National Institutes of Child Health and Human Development (NICHD) to L.O., M.J.S. and M.H.C., and by a Dolores Zohrab Liebmann Fellowship awarded to Thomas A. Farmer. Part of this work was carried out when L.O. and M.J.S. were at Cornell University.

References:

Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System,* 21, 1, 101-114.

Barker, C., & Dowty, D. (1993). Non-verbal thematic proto-roles. In A. Schafer (Ed.), *Proceedings of NELS 23*, vol. 1, (pp. 49-61). Graduate Student Linguistic Association, Amherst, MA.

Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum*, I, 1-14.

Brent, M. (1991). Automatic acquisition of subcategorization frames from untagged text. *Proceedings of the 29th annual meeting on Association for Computational Linguistics*, 209-214.

Church, K., & Hanks, P. (1991). Word Association Norms, Mutual Information and Lexicography. *Computational Linguistics*, 16, 1, 22-29.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1992). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.

de Groot, A. & Nas, G. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30, 90–123.

Elman, J.L. (1995) Language as a dynamical system. In R.F. Port and T. van Gelder (Eds), *Mind as motion: Explorations in the dynamics of cognition*, 195--223. Cambridge, MA: MIT Press.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*, 301-306.

French, R. M. and Labiouse, C. (2002). Four Problems with Extracting Human Semantics from Large Text Corpora. *Proceedings of the 24th Annual Conference of the Cognitive Science Society*.

Gaskell, M., & Marslen-Wilson, W. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, *45*, 220-266.

Harris, C. L. (1998). Psycholinguistic studies of entrenchment. In J. Koenig (Ed.), *Conceptual structure, discourse and language*. Stanford, CA: CSLI Publications.

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hopper, P. (1998). Emergent Grammar. In Tomasello, M. (ed.) *The new psychology of language*. Mahwah, New Jersey and London: Lawrence Erlbaum.

Howart, P. (1998). Phraseology and Second Language Proficiency. *Applied Linguistics 19*, 1, 24-44.

Jackendoff, R. (1997). Twistin' the night away. *Language,* 73, 3, 534-559.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. Journal of Experimental Psychology: General, 111, 228-238.

Keller, F., & Lapata, M. (2003). Using the Web to Obtain Frequencies for Unseen Bigrams. *Computational Linguistics*, 29, 3, 459-484.

Kemtes, K.A., & Kemper, S. (1997). Younger and older adults on-line processing of syntactic ambiguities. *Psychology and Aging*, 12, 362-371.

Kutas, M., & Hillyard, S. A. (1984). Brain potential during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.

Langacker, R. (1988). A usage-based model. In Rudzka-Ostyn, B. (Ed.) *Topics in cognitive linguistics*. Amsterdam: Benjamins.

Lewis, M. (Ed.) (2000). *Teaching Collocation: Further Developments in the Lexical Approach*. Hove, England: Language Teaching Publications.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Baker M., Francis G. and Tognini-Bonelli E. (Eds.) *Text and Technology: In Honour of John Sinclair*, 157-76. Amsterdam: John Benjamins.

MacDonald, M.C., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71-102.

McDonald, S.A., & Shillcock, R.C. (2003). Eye movements reveal the on-line computation of lexical probabilities. *Psychological Science*, 14, 648-652.

McRae, K., Ferretti, T.R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes: Special Issue on Lexical Representations in Sentence Processing*, 12, 137-176.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition, 33*, 1174-1184.

Mittelberg, I., Farmer, T. A., & Waugh, L. R. (2007). They actually said that? An introduction to working with usage data through discourse and corpus analysis. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. Spivey (Eds.), *Methods in cognitive linguistics: Ithaca* (pp. 19-52). Amsterdam/New York: John Benjamins.

Onnis, L. (2001). Fluency in native and non-native speakers. Published undergraduate dissertation. In A. Carli (Ed.) *Aspetti linguistici e interculturali del bilinguismo*. Milan: Franco Angeli, 20-139.

Partington, A. (1998). *Patterns and Meanings. Using Corpora for English Language Research and Teaching*. Amsterdam: Benjamins.

Pang, Bo and Lillian Lee. 2004. "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd ACL*, 271–278.

Pawley, A., & Syder, F.H. (1983). Two puzzles for linguistic theory: native-like selection and native-like fluency. In: Jack C. Richards and Richard W. Schmidt (eds.). *Language and Communication*, 191-226. London and New York: Longman.

Rogers, T. T. and McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA: MIT Press

Sardinha, T. (2000). Semantic prosodies in English and Portuguese: A contrastive study. *Cuadernos de Filología Inglesa,* 9,1, 93-110.

Sinclair, J.M. (1991). *Corpus, Concordance, Collocation*. Oxford: OUP.

Sinclair, J.M. (1996). The Search for Units of Meaning. *Textus,* 9, 75-106.

Spivey-Knowlton, M., Tanenhaus, M., Eberhard, K., Sedivy, J. (1998). Integration of visuospatial and linguistic information in real-time and real-space. In P. Olivier & K. Gapp (Eds.), *Representation and Processing of Spatial Expressions*. 201-214. Mahwah, NJ: Erlbaum.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language*, 2, 4–27.

Tabor, W., & Tanenhaus, M.K. (1999). Dynamical Models of Sentence Processing. *Cognitive Science*, 23, 4, 491-515.

Tanenhaus, M.K. Carlson, G., & J.C. Trueswell (1989) The role of thematic structures in interpretation and parsing, *Language and Cognitive Processes,* 4, 211-234.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.

Tanenhaus, M.K., & Trueswell, J.C. (1995). Sentence comprehension. In J.L. Miller & P.D. Eimas (Eds.) *Speech, language, and communication*. San Diego, CA: Academic Press.

Turney, P.D., and Littman, M.L. (2003). Measuring praise and criticism: Inference of semantic orientation from association, *ACM Transactions on Information Systems (TOIS)*, 21(4), 315-346.

Vigliocco, G., Vinson, D.P, Lewis, W. & Garrett, M.F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48, 422-488.

Zhu, X., & Rosenfeld, R. (2001). Improving trigram language modeling with the world wide web. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*. Salt Lake City, Utah.

Table 1

A random sample from the British National Corpus produced by searching for sentences containing the verb 'cause'. Brackets highlight the verb and its immediate noun to the right. Even a quick look reveals that the collocates of cause are negative.

…in the lung and the gut, <causing shortness> of breath and other problems …
…Every day the virus <causing AIDS> is infecting more young people …
…But some drugs <cause bad , disturbing flashbacks> . " I can't cope …
…Income Tax ? This can <cause problems> , since you agree under the terms…
…evidence that Iraqi forces had <caused the deaths> of babies by removing…
…varied , and some personal animosities <caused the alliance> to break up…
…an immoderate devotion to them <causes an infinite waste> of time , fatigues…
…to accept that he had <caused his brother> to suffer . In all this there…
…hey are marvellously done , and they have <caused a stir> of approval in this…
…canon of artistic detachment , but it can <cause controversy> . Heirs to the…
…clash between male and non-male that <causes all the trouble> . They are…
…it happens . He makes mistakes and <causes havoc> , in pursuit of the right…
…and such criticism can <cause considerable distress> to many people…
… to speak in public places even if it <causes an affray> , and opposing the…
…city' s Phoenix Park <caused particular concern> to Eire 's tourism industry…
…addressed in a professional manner can <cause catastrophic consequences>…

Table 2

Mean (SDs) human Semantic Valence Tendency Ratings over Fragments of sentences that followed the prime in Experiment 1.

Prime	Mean rating	SD	Valence tendency postulated (p=positive; n=negative)
PROVIDE	0.57	1.21	p
PURE	0.36	1.13	p
PERFECTLY	0.32	1.09	p
KNOWN FOR	0.18	1.44	p
PROFOUNDLY	0.13	1.39	p
SHEER	0.10	1.12	n
UTTERLY	-0.09	1.71	n
DEEPLY	-0.15	1.15	n
CLEARLY	-0.20	1.11	n
PEDDLE	-0.30	0.77	n
MARKEDLY	-0.35	1.19	n
REVEAL	-0.35	1.29	n
INVOLVED IN	-0.37	1.07	n
NOTORIOUS FOR	-0.39	1.22	n
ENCOUNTER	-0.41	1.06	n
BENT ON	-0.43	1.40	n
INCITE	-0.48	1.09	n
CONSIDERABLE	-0.55	1.15	p
HARBOR	-0.65	1.23	n
EXPRESS	-0.65	0.97	n
PATENTLY	-0.67	1.34	n

CAUSE	-0.73	1.13	n
COMMIT	-0.97	1.28	n

Table 3

Means (SDs) associated with the control t-tests in study 2.

Prime-Target Pairing	Plausibility (Scale of 1-7)	Length of Target Word	Frequency of Target Word	Log-Frequency of Target Word	Log-Frequency Of Bigram
Consistent	4.85 (.49)	7.17 (1.17)	1283 (1023)	6.76 (1.11)	1.92 (1.5)
Inconsistent	4.60 (.85)	7.5 (1.05)	1502 (1097)	7.05 (.86)	1.00 (0.94)

Table 4

Semantic Valence Tendency Ratings generated by the algorithm in Experiment 3.

Prime	Valence Tendency generated by algorithm	Valence tendency postulated (p=positive; n=negative)
PROVIDE	2.66	p
IMPRESSIVE	-0.26	p
CONSIDER	-1.39	p
LARGELY	-1.85	p
BROADLY	-1.90	p
CONSIDERABLE	-2.01	p
PURE	-2.03	p
PERFECTLY	-2.16	p
EXPRESS	-2.32	n
DEEPLY	-2.38	n
MARKEDLY	-2.46	n
ENCOUNTER	-2.55	n
COMMIT	-2.74	n
CAUSE	-2.84	n
VOICE	-3.78	n
HARBOR	-3.98	n
FICKLE	-4.85	n
PEDDLE	-4.99	n
INCITE	-5.10	n
UTTERLY	-5.46	n
PATENTLY	-6.70	n

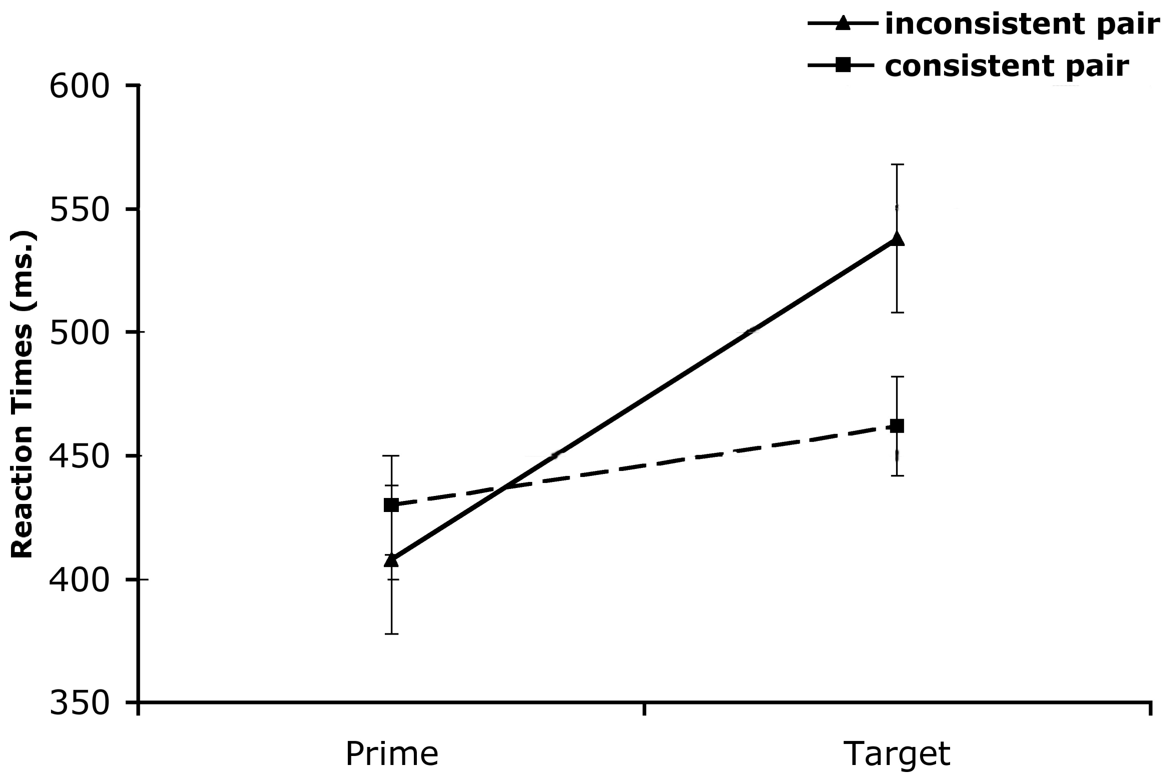

Figure 1. Mean Reading Times associated with reading prime and target words in the self-paced reading task (Experiment 2).