Persistent structural priming from language comprehension to language production

Kathryn Bock a,*, Gary S. Dell a, Franklin Chang b, Kristine H. Onishi c

a University of Illinois at Urbana-Champaign, Beckman Institute, 405 North Mathews Avenue, Urbana, IL 61801, USA
b NTT Communication Science Laboratories, 2-4 Hikari-dai, Seika-cho, Kyoto 619-0237, Japan
c Department of Psychology, McGill University, Montreal, Quebec, Canada, H3A 1B1

Received 22 June 2005; revised 3 July 2006; accepted 4 July 2006

Abstract

To examine the relationship between syntactic processes in language comprehension and language production, we compared structural persistence from sentence primes that speakers heard to persistence from primes that speakers produced. [Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology: General, 129*, 177–192.] showed that the production of target priming structures increased the probability of spontaneously using the same structures to describe events in subsequent pictures that were semantically unrelated to the primes. These priming effects persisted across as many as ten intervening filler trials. The present studies replicated these results using auditorily presented primes to which participants only listened. The results indicated persistence of priming across all lags, with relative magnitudes of priming as large as
those observed by Bock and Griffin. The implication is that structural priming is persistent regardless of the modality in which language structures are experienced, underscoring the power of priming as an implicit learning mechanism. © 2006 Elsevier B.V. All rights reserved.

Keywords: Sentence production; Language production; Structural persistence; Implicit learning; Language comprehension; Syntax; Language learning

1. Introduction

The distinction between linguistic competence and linguistic performance is one of the most controversial in the study of human language. In parallel with related distinctions between knowledge and behavior in traditional learning research (Tolman, 1948), the competence–performance distinction rests in part on differences between what is known in principle and what is done in practice. Beyond this, however, there are diverging views about how competence shapes or participates in performance. In this work, we call on the relationship between language comprehension and language production to evaluate two views of the relationship between language knowledge and language use.

Linguistic competence can be construed as having everything to do with normal performance (Bresnan & Kaplan, 1984; Bybee, in press) or as having next to nothing to do with it (Newmeyer, 2003). The first position is called the strong competence hypothesis. It says that competence is a single system that constitutes the language user's internal description of linguistic knowledge. The system is not merely compatible with the demands of performance, but supportive of them: "The formal properties of...proposed linguistic representations [must be] related to the nature of the cognitive processes that derive and interpret them in actual language use" (Bresnan & Kaplan, 1984, p. 107). Because language users both understand and speak the same language, successful communication entails that at some level, the same knowledge participates in language comprehension and language production.

Strong competence contrasts with what we will call weak competence. The weak competence hypothesis doubts the utility of an account of language knowledge that serves the goals of linguistic theory as well as the goals of explaining the different facets of normal language performance, including language comprehension and production. Clark and Malt (1984) presented a persuasive version of a weak competence hypothesis, arguing that different kinds of knowledge underlie different kinds of language use. They observed that "A native Californian can understand...Australian, Indian, Scottish...yet not have the slightest [ability to produce] them. He can understand syntactic forms in these dialects, as well as in Shakespeare, Joyce, and even Bellow...He can understand a large number of words that he couldn’t use himself" (Clark & Malt, 1984, p. 200). In short, "comprehension and production [may] access distinct representations of linguistic knowledge, even though in normal people the two representations code much the same information and are closely coordinated" (Clark & Malt, 1984, pp. 200–201). Along similar
lines, neuropsychological evidence for the partitioning of linguistic information by modality in language disorders (Caramazza & Hillis, 1991) is consistent with a weak competence perspective.

From the standpoint of psycholinguistic theory, these stances imply different things for our understanding of the relationship between language comprehension and language production. Strong competence predicts that the same knowledge representations enter into performance regardless of the modality. Weak competence predicts that different kinds of knowledge support different kinds of linguistic performance. To evaluate competing predictions from these views about syntactic processes in particular, we used the phenomenon of structural persistence.

Structural persistence is a tendency to echo syntactic structures from recent experience, despite changes in the wording, in the wording, even in the language embodying the persistent structure (Bock, 1986; Bock & Loebell, 1990; Hartsuiker & Kolk, 1998; Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003; Pickering & Branigan, 1998; Potter & Lombardi, 1998). The use of “old structures in new clothes” can be observed naturalistically in speakers echoing structures from their own speech as well as structures from the speech of others (Gries, 2005; Szmarcinyi, 2004, 2005). This implies that production and comprehension can both affect the structural-syntactic underpinnings of language performance (Branigan, Pickering, & Cleland, 2000; Potter & Lombardi, 1998). An open question is whether the effects are similar in strength and duration. The answer to this question is important for our understanding of the links between comprehension and production and important to explaining how knowledge of language works with respect to the modalities it serves.

The importance of the issue derives in part from how it bears on the problem of language learning (Seidenberg & MacDonald, 2001). To the degree that linguistic knowledge and the cognitive skills of language comprehension and language production are separate achievements of language acquisition, they require separate explanations. What happens in language comprehension need not constrain what happens in language production, and vice versa, and the configuration of neither comprehension nor production need inform our understanding of language knowledge. However, if comprehension and production share cognitive resources that can be identified with aspects of language knowledge, each step in understanding how comprehension works (for example) offers a constraint on understanding how production works and a constraint on explaining the knowledge that supports normal language performance. Suppose that information accessible to language comprehension can become automatically accessible for language production, and vice-versa, without further learning. Suppose that the ability of a language user to understand a particular construction is a strong predictor of ability to produce that same construction. If so, the case for generalized, performance-relevant language knowledge becomes stronger.

Chang, Dell, and Bock (2006) proposed a model for a learning mechanism that is general enough and abstract enough to span language comprehension and language production. The focus of the Chang et al. model is the phenomenon of structural persistence. The model offers an account that calls on implicit, error-based learning as the foundation of structural persistence (Bock & Griffin, 2000). Implicit learning is a type of procedural memory that results from engaging a cognitive mechanism,
effectively tuning the operation of the mechanism and thereby changing it. Definitions of this kind of learning emphasize its tacit nature, its inaccessibility to consciousness, and its preservation in the face of loss of normal explicit memories of the performance that yields implicit learning. Regarding the last of these properties, the evidence for the persistent, implicit character of structural persistence is that it occurs in anterograde amnesics, at the same magnitude as in adults with unimpaired memory ability (Ferreira, Bock, Wilson, & Cohen, 2006).

There are alternatives to an implicit learning account that cast the workings of structural persistence in different terms. Some persistence phenomena are treated within approaches that emphasize social accommodation in general (Giles & Coupland, 1991) and communicative alignment in particular (Pickering & Garrod, 2004). There are well known short-term priming effects in cognition due to biases and expectations about upcoming information (Neely, 1991) and to the transitory nature of direct memory for language (Sachs, 1967). Credible links exist between these transient effects and structural priming, particularly in the strong relationship between lexical repetition and structural priming (Branigan et al., 2000; Levelt & Kelter, 1982; Smith & Weldon, 2001). The implicit learning account illuminates structural persistence as a basic and fairly primitive mechanism of language use that supports language learning and language change, as well as providing a scaffold upon which the short-lived dynamics of language performance can build. By itself, it cannot explain how structural repetition serves or facilitates communication and social interaction, or how speakers might make strategic use of parallel structure; for these things, other theories are essential. Instead, what the implicit learning view emphasizes is the automatic, nondeliberative character of structural repetition, its abstractness, and its duration. All of these properties contribute to generalization from previous language experiences to new experiences, which is fundamental to a learning process.

The present work extends existing evidence for the abstractness and duration of structural persistence by examining the duration of structural generalizations from language comprehension to language production. We tested the persistence of sentence structures across utterances when speakers heard but did not reproduce the priming structures. In the first experiment, we examined persistence from comprehension to production after immediate priming and after priming across interruptions by one to ten other utterances, and compared the strength and duration of persistence to existing results for within-modality, production-to-production priming. In the second experiment, we changed the priming procedure to reduce any effects of explicit or intentional memory encoding and self-monitoring. Both experiments used a picture-description priming paradigm (Bock, 1986) in which speakers experienced a long, mixed sequence of sentences and line-drawings of simple events. The primary task involved recognition memory for previously presented materials. The elements of the priming procedure were introduced as secondary components of the memory task: Speakers were asked to listen carefully to each auditorily presented sentence and to describe aloud, in one sentence, the event described in each picture. Fig. 1 illustrates a priming trial with a lag of two sentences between the priming sentence and target picture. Only a small subset of the sentences presented constituted priming sentences and a similarly small subset of the pictures constituted targets, thus embedding
the priming trials among fillers that elicited different sentence structures. In the first experiment, we compared the results of the priming manipulation to previous results from tasks in which speakers listened to and repeated auditorily presented sentences, creating a situation in which the impact of production could be compared to the impact of comprehension alone.

In earlier work using this paradigm, structural persistence has been found across as many as 10 intervening trials after speakers produced the priming structures (Bock & Griffin, 2000). Other work hints at even longer time courses from production priming (Bock & Kroch, 1989; Safran & Martin, 1997), although there is also evidence that its effects can be transient in some circumstances (Branigan, Pickering, & Cleland, 1999; but see also Branigan, Pickering, Stewart, & McLean, 2000). Here, we explore for how long and at what strength structures persist when they are understood but not produced, and compare the outcomes to a parallel set of data in which the primes were produced.

If the workings of syntactic structure differ in language comprehension and language production, differences in strength and duration would be predicted simply
because production priming entails three exposures to the prime: Speakers hear the prime, produce it, and hear themselves say it. Representations of the priming events in episodic memory could also magnify the impact of modality differences. Structural priming being in essence repetition priming of an unusually abstract kind (Bock & Griffin, 2000), the consequences of changing modalities may be similar to those found in work on other kinds of repetition priming. These consequences include sharply reduced magnitudes of priming and shortened durations (e.g., Roediger & Blaxton, 1987). Episodic accounts of the acquisition of new sentence structures likewise predict that a change from production to comprehension priming should reduce the magnitude of priming (Kaschak & Glenberg, 2004).

Alternatively, if the origins of structural persistence are in structural representations or mechanisms that subserve events in language comprehension as well as language production, the prediction of differences is less certain. Any episodic components to structural persistence would weigh in favor of modality differences regardless of the underlying nature of persistence, but there is little hard evidence for such components except in the immediate lexical enhancement of priming (Konopka & Bock, 2005; Levelt & Kelter, 1982). The most persuasive evidence for general structural mechanisms that are effective in both comprehension and production would be persistence that is similarly strong and similarly durable over time or intervening events.

2. Experiment 1

Experiment 1 used the materials, procedures, and designs from Bock and Griffin’s (2000, Experiments 1 and 2) shorter-lag and longer-lag experiments, making one important change. Participants heard sentences, including priming sentences, without repeating them aloud. Separate groups received the shorter- (Lags 1 and 2) and longer- (Lags 4 and 10) lag primes; both groups also received immediate (Lag 0) priming. In Experiment 1 we collected and combined the data from these lag conditions and then compared the results to those from Bock and Griffin’s production priming experiments, whose 240 participants were sampled from the same undergraduate population.

2.1. Method

2.1.1. Participants

There were 96 undergraduates from the University of Illinois at Urbana-Champaign in the short-lag conditions of the experiment and 144 in the longer-lag conditions. Remuneration for participation was a small payment or credit toward an introductory psychology course requirement. The 36 participants who failed to produce a minimum of 25% target or alternative responses in the picture-description task were replaced.

2.1.2. Materials

The materials and list arrangements were exactly as described in Bock and Griffin (2000). Examples of the two types of priming sentences and accompanying pictures are given in Table 1. The placeholders and neutral primes were intransitive sentences,
and the fillers represented a variety of sentence structures that were minimally related to the target structures.

The materials were a combination of spoken sentences and line drawings of simple events. There were 48 experimental pictures and 48 sets of auditorily presented priming sentences in the short-lag conditions and 36 experimental pictures and 36 sets of auditorily presented priming sentences in the long-lag conditions (the reduced number of experimental trials in the long-lag condition, as in Bock and Griffin’s work, accommodated the longer lists needed for the long-lag manipulation). The priming trials were separated by filler pictures, and intransitive placeholder sentences separated the prime and target trials to create the lag manipulations. Half of the experimental pictures were selected to elicit simple active and passive transitive sentences, and the other half were selected to elicit prepositional and double-object dative sentences. The normatively less frequent structures, the passives and the prepositional datives, were designated the target structures for purposes of scoring and analysis. (Relative frequencies were determined from unpublished description norms for the picture set.)

The priming sentences were transitive and dative sentences in the target (passive, prepositional dative) and nontarget (active, double-object dative) structures. Half of the primes were transitive and the other half were dative sentences. The transitive

<table>
<thead>
<tr>
<th>Type of event</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitive active prime</td>
<td><em>The government isn’t evacuating the embassy staff</em></td>
</tr>
<tr>
<td>Transitive passive prime</td>
<td><em>The embassy staff isn’t being evacuated by the government</em></td>
</tr>
<tr>
<td>Transitive target-eliciting picture</td>
<td>[Image]</td>
</tr>
<tr>
<td>Dative double-object prime</td>
<td><em>The team owner told the columnist an offensive joke</em></td>
</tr>
<tr>
<td>Dative prepositional prime</td>
<td><em>The team owner told an offensive joke to the columnist</em></td>
</tr>
<tr>
<td>Dative target-eliciting picture</td>
<td>[Image]</td>
</tr>
<tr>
<td>Neutral (intransitive) prime</td>
<td><em>The young man shaved too often</em></td>
</tr>
</tbody>
</table>
primes were actives or their full-passive counterparts, with the same content words in different syntactic structures. The dative primes were prepositional datives or their double-object counterparts, also with the same content words in different structures. The placeholder sentences consisted of intransitive (e.g., The real-estate agent blundered) and predicate adjective (e.g., The books were expensive) sentences. All of the materials were the same as used by Bock and Griffin (2000), including the same digital recordings and the same digitized images.

Each priming sentence from the transitive and dative sets was coupled with an experimental picture designed to elicit the same sentence type. Each transitive-eliciting picture was paired (in different lists) with both the active and passive versions of a prime, and each dative-eliciting picture was paired (in different lists) with both the prepositional and double-object versions of a prime. The constraints on these couplings were that (a) the primes and the expected picture descriptions did not share any content words; and (b) there were no obvious thematic or narrative relationships between the priming sentences and the anticipated descriptions. The materials were arranged and presented as mixed lists of pictures and sentences. Dative and transitive priming trials alternated at a distance throughout each list, so that participants never encountered two pictures of the same type on successive priming trials, and the two different forms of each prime type also alternated, yielding the rotation double-object dative, active, prepositional dative, passive.

Priming trials were separated from each other by one filler picture. For the purposes of the cover memory task, filler pictures and placeholder sentences were sometimes repeated in the course of the experiment. Including these repetitions, there were 288 events in the long-lag lists and 252 in the short-lag lists. Assignments of priming sentences to lists were counterbalanced so that every list contained only one sentence from each of the priming-sentence pairs, with equal numbers of sentences in each of the four priming forms (active, passive, prepositional, and double object) in each list.

2.1.3. Procedure

Fig. 1 illustrates the events on a lag-2 transitive priming trial, when two intransitive placeholders separated the auditory presentation of a passive prime (The car’s window was struck by a brick) from the presentation of the pictured event designed to elicit the passive target structure. Depending on lag condition, the number of placeholders separating the primes from the target events was 0, 1, or 2 in the short-lag conditions, and 0, 4, or 10 in the long-lag conditions. In the short-lag conditions, all priming sentences were either preceded by the two placeholders (lag 0), preceded by one and followed by another placeholder (lag 1), or followed by two placeholders (lag 2). In the long-lag conditions, all priming sentences were either preceded by ten placeholders (lag 0), preceded by six and followed by four (lag 4), or followed by all 10 (lag 10). A priming item in any of its alternative forms (e.g., active or passive) was always accompanied by the same unique set of placeholders, so that the list environment in which each item occurred was similar across conditions and lags. Different sets of placeholders accompanied different items.

On all sentence trials, the participants listened to the sentence and indicated whether or not the sentence had occurred previously in the course of the experiment.
On picture trials, the participants described what was happening in the depicted event, and then indicated whether the picture had occurred previously in the course of the experiment. All of the events during the experiment were controlled by a Macintosh Quadra 650 running PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants used the PsyScope button box to proceed through the experimental lists. Trials began when participants pressed a button on the button box. On sentence trials, this caused the message “Listen” to be displayed on the computer monitor; on picture trials the message “Describe” was displayed. Then a sentence was presented auditorily or a picture appeared, remaining until the pictured event was described. At the offset of the sentence trial the monitor displayed the question “Have you heard this sentence before?”, and at the offset of a description the question “Have you seen this picture before?” appeared. The questions remained until a “yes” or “no” response was made on the button box. Participants’ picture descriptions were recorded to digital audio tape.

Participants were run individually. They were told that their task was to listen to sentences and look at pictures, and to try to detect all of the repetitions of the sentences and pictures. They were also asked to describe what was happening in each picture. The instructions for the descriptions were to “use just one sentence and try not to use any pronouns,” in all other respects leaving the participants free to decide for themselves what to say and how to say it. Two examples (one picture and one sentence) were presented as part of the instructions, and two practice items (one picture and one sentence) followed the instructions. None of these pictures or sentences elicited the critical sentence constructions. After instructing the participant about the task, the experimenter monitored the session from an adjacent room.

The details of these procedures were the same as Bock and Griffin’s (2000) with one critical difference designed to eliminate structural persistence due to the speaker’s own production of the priming sentences. Participants simply listened to sentences on the sentence trials, rather than listening and then repeating.

2.1.4. Scoring and data analysis

Table 2 gives the raw numbers of target (passive or prepositional dative) and alternative (active or double-object dative) structures produced in each condition at each lag. These numbers represent the codings of the descriptions of the experimental pictures after transcription from the audio tapes of the experimental session.

Coding and data analyses were performed as in Bock and Griffin (2000), where the scoring system is described in detail. In broad outline, to be eligible for coding, picture descriptions had to include a clause containing the appropriate number of noun phrases mentioning key actors from the pictured event (two for transitive events and three for dative events) and a verb of the relevant type (transitive or ditransitive dative, respectively). For transitive pictures, utterances were scored as active, passive or other. To be coded as an active, the scored clause had to have an acceptable passive counterpart, and to be coded as a passive, the clause had to have an acceptable active counterpart. Passive clauses had to contain a passivized verb form (i.e., a main verb preceded by a form of be or get) followed by a by-phrase (i.e., a prepositional phrase headed with the preposition by). Datives were scored as
prepositional datives, double-object datives, or other. Prepositional datives had to have acceptable double-object counterparts, and double objects had to have acceptable prepositional counterparts. Structurally, prepositional datives contained a subject, a dative verb, a direct object noun-phrase, and after the direct object a prepositional phrase headed by to or for. Double-object datives contained a subject and two object noun-phrases. All utterances that fell outside of these four structural

Table 2
Numbers of target (passive or prepositional dative) and alternative (active or double-object dative) structures produced in each priming condition at each of four lags in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Lag group and prime type</th>
<th>Structure produced</th>
<th>Lag</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Short lags, Experiment 1</td>
<td>Target</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>171</td>
</tr>
<tr>
<td>Short lags, Experiment 2</td>
<td>Target</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>268</td>
</tr>
<tr>
<td>Long lags, Experiment 1</td>
<td>Target</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>66</td>
</tr>
<tr>
<td>Short lags, Experiment 1</td>
<td>Target</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>118</td>
</tr>
<tr>
<td>Short lags, Experiment 2</td>
<td>Target</td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>214</td>
</tr>
<tr>
<td>Long lags, Experiment 1</td>
<td>Target</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>37</td>
</tr>
<tr>
<td>Short lags, Experiment 1</td>
<td>Target</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>106</td>
</tr>
<tr>
<td>Short lags, Experiment 2</td>
<td>Target</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>205</td>
</tr>
<tr>
<td>Long lags, Experiment 1</td>
<td>Target</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>54</td>
</tr>
<tr>
<td>Short lags, Experiment 2</td>
<td>Target</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Alternative</td>
<td>43</td>
</tr>
</tbody>
</table>

Datives (target = prepositional dative; Alternative = double-object dative)

Transitives (target = passive; alternative = active)
categories were coded as other and omitted from analysis. In total, 78% of the trials were included in the analyses, in line with the 77% rate in Bock and Griffin’s Experiments 1 and 2 (2000). The rates reflect the loss of data due to participants’ use of picture descriptions falling outside the bounds of the responses of interest, which is normal when extemporaneous picture description is used (Bock, 1996).

Separate analyses were performed on the short- and long-lag data sets using participants as the random factor in the participant analysis and items as the random factor in the item analysis. For the long-lag data sets, as in Bock and Griffin (2000), participant pairs were used as the random factor to reduce the number of empty cells in the design (see Dell, Burger, & Svec, 1997; Ratcliff, Thapar, Gomez, & McKoon, 2004, for similar methods of grouping to address sparse data problems). The pairs were formed by randomly coupling data from individuals who received the same stimulus lists in the same order. Participant-pairs and items were both crossed with the factors of prime structure (same, different, or neutral with respect to the target structure) and lag between prime and picture presentation (0, 1, and 2 trials in the short-lag conditions and 0, 4, and 10 trials in the long-lag conditions). The results of the participant and item analyses were combined in minF' calculations. Statistically significant effects were associated with probabilities less than or equal to .05. Effects that were significant for participants and items but not by minF' are labelled as marginal-by-minF'. In all analyses, degrees of freedom were reduced when data from any cell were missing.

To evaluate effect sizes, 95% confidence intervals for planned pairwise comparisons were calculated using the mean squared error for the priming factor from the relevant analyses. The halfwidths of these intervals indicate whether and by how much a difference departs from zero.

2.2. Results

Fig. 2 summarizes the magnitudes of cross-modal persistence for the short- and long-lag conditions combined; Table 2 gives the results separately by lag condition. The corresponding data for intra-modality persistence are shown on the right-hand side of the graph (from Bock & Griffin, 2000).

Fig. 3 (upper panel) breaks out the cross-modal patterns at each lag, in comparison to the intramodal, production-to-production data on the bottom. The analysis of the crossmodal data yielded statistically significant effects of priming for the short-lag condition \( F_1(1,95) = 6.41, F_2(1,46) = 13.61, \text{ min } F'(1,141) = 4.36 \) and \( \text{min } F' \)-marginal effects for the long-lag conditions \( F_1(1,72) = 8.05, F_2(1,34) = 6.04, \text{ min } F'(1,82) = 3.45, p = .067 \). Priming did not interact with lag \( F_1(2,190) = 1.85, F_2(2,94) = 1.83 \) in the short-lag conditions and \( F_1(2,144) = 1.02, F_2(2,68) = 1.91 \) in the long-lag conditions. Overall, the target-primed condition differed from the alternative-primed baseline (with proportions of targets produced out of all targets and alternatives of .54 and .46, respectively), but only the target-primed condition differed from the neutral condition (.47 targets produced) when it was included in the analyses; that is, in the long-lag conditions. The halfwidth of the 95% confidence interval for pairwise planned comparisons of these means was .05, calculated from the participants analysis of the long-lag conditions.
To compare the effects of priming and lag in different modalities, we carried out an analysis of the combined cross- and intramodal data without the neutral condition. To do this, we pooled the lag-0 data from the short- and long-lag lists in both experiments and included just the 36 items that occurred on all lists in both experiments. The analysis was performed on items only, for several reasons. First, with 36 items, the item analyses are more conservative than the analyses by participants, which included 480 individuals (240 in each experiment), and are less distorted by missing data. Second, the items in the two modality conditions were identical, allowing them to serve as their own controls, whereas the participants in the two modality conditions differed. Third, the different numbers of participants and different designs of the short- and long-lag conditions of the experiments, including the use of participant pairs as the random factor in the long-lag conditions, made joint analysis at best unwieldy (for the same reasons, the conditions were analyzed separately above). Finally, the results from participant analyses in the component experiments uniformly revealed statistically significant effects of priming.

In the analysis of the combined data sets, there was a statistically significant effect of priming ($F_2(1,32) = 26.35$) but no effect of lag ($F_2(4,128) = .12$) and, more important, no interaction between priming and lag ($F_2(4,128) = .70$). The interactions of

![Proportion Targets Produced](chart)

**Fig. 2.** Overall proportions of targets produced after Target, Neutral, and Alternative primes when primes were comprehended or produced (production priming data from Bock and Griffin, 2000; error bars represent halfwidths of 95% confidence intervals for pairwise planned comparisons of differences between means).
modality and priming ($F_2(1,32) = .03$), of modality and lag ($F_2(4,128) = 1.20$), and modality, priming, and lag were also not significant ($F_2(4,128) = 1.47$). The only difference between comprehension and production priming that approached significance was a marginal main effect of modality ($F_2(1,32) = 3.88, p < .06$), due to the overall larger number of targets produced with production priming.

The target sentence types (transitives and datives) were not equivalent in their sensitivity to priming. The clearest evidence for this emerged in the combined analysis, which produced a significant interaction between priming and sentence type ($F_2(1,32) = 8.79$). The interaction reflects a difference between datives and transitives in the magnitude of priming: For the dative items, the priming effect was .10 (.46 vs. .36); for the transitives it was .02 (.61 vs. .59). The confidence interval for pairwise comparisons of means involved in the interaction was .087. However, there were no significant modulations of these differences in primability due to modality or lag (all $F_2 < 1.56$).

2.3. Discussion

Experiment 1 elicited structural persistence in descriptions of pictured events when primes were heard but not reproduced by speakers. This result replicates
previous demonstrations of persistence from language comprehension to language production (Branigan et al., 2000; Huttenlocher, Vasilyeva, & Shimi, 2004; Potter & Lombardi, 1998; Weiner & Labov, 1983) and adds to them two major findings. First, there was evidence of persistent priming from comprehension to production across multiple intervening utterances of other types. Second, taken together with Bock and Griffin (2000), the results of Experiment 1 imply that structural priming from comprehension to production can be roughly as strong and as durable as priming within production itself.

3. Experiment 2

Experiment 2 addressed an alternative to the implicit learning hypothesis about why crossmodal priming persists. A plausible objection to the methods employed in Experiment 1 is that the cover task, the running recognition memory test, encourages speakers to store priming sentences in long-term memory (as noted in Branigan et al., 1999). If explicit encoding efforts or the workings of a comprehension monitor are responsible for the persistence of priming, the duration of priming should decrease when encoding efforts are minimized, as they should be when priming structures do not occur among the to-be-remembered materials. This hypothesis was tested in Experiment 2. The experiment again examined priming from language comprehension to language production, but using a cover task that encouraged indirect encoding of the priming sentences.

The cover task in the experiment, the participants’ primary task, was recognition memory administered in a study-test arrangement. During study, only filler materials – pictures and auditory sentences – were presented as the to-be-remembered items (see also Bock, 1986, Experiment 1). During the test phase, which occurred a day after the study phase in order to accommodate the duration of the sessions, the memory test was presented. The foils in the test were the priming sentences and the target pictures, all of which occurred among the filler items that were repeated from the study phase. The priming sentences occurred at each of three lags relative to the target picture (0, 1, and 2). They occurred just once in the course of the test phase, and were simply judged as to whether they had occurred in the study set. If explicit, direct encoding is responsible for the persistence of crossmodal priming, it should be short-lived or absent under these conditions.

3.1. Method

3.1.1. Participants and materials

The participants were 144 undergraduates from the same source as Experiment 1, compensated in the same ways; 10 were replaced. The materials were those from the short-lag lists in Experiment 1.

3.1.2. Procedure

The experiment comprised two sessions separated by a day. In session 1, participants received the filler items (pictures and auditory sentences) from the short-lag
lists. The order was the same as in Experiment 1, but excluded all of the priming sentences, target pictures, and placeholders. Instructions were to describe what was happening in each picture and repeat each sentence, studying the materials well enough to be able to recognize them the next day. The duration of picture trials was set at 7 s. In session 2, the participants received the short-lag lists from Experiment 1, with modified instructions and without repetitions of fillers. After each auditorily presented sentence or picture description, participants indicated whether the sentence or picture had occurred during session 1, responding “yes” or “no”. Because none of the priming sentences or target-eliciting pictures had occurred in session 1, the correct response to them was always “no”. This made most of the overt experimental events during session 2 identical to those for the short-lag lists in Experiment 1.

3.1.3. Scoring and data analyses

Responses were scored and analyzed in the same way as Experiment 1. Table 2 gives the numbers of responses in each category.

3.2. Results

The priming results over lags are shown in Fig. 4. More targets were produced after priming, with an overall proportion of .41 in the primed condition compared to .35 in the alternative condition ($F_1(1,124) = 18.31$, $F_2(1,46) = 22.47$, $min F(1,146) = 10.09$). Lag was significant only in the items analysis ($F_1(2,248) = 2.55$, $F_2(2,92) = 3.27$, $min F(2,298) = 1.43$) and none of the interactions with lag approached significance (all $Fs < 1$).

The effect of sentence type was significant ($F_1(1,124) = 43.41$, $F_2(1,46) = 11.18$, $min F(1,71) = 8.89$, respectively), but the interaction between priming and sentence type was significant in the items analysis only ($F_1(1,124) = 1.87$, $F_2(1,46) = 8.11$, $min F(1,164) = 1.52$). The overall proportions of targets produced in the target-
alternative-primed conditions for datives were .34 and .23, respectively, compared to .49 and .46 for transitives. These effects did not interact with lag (all Fs < 1).

Accuracy on the memory test in session 2 was high, with 93.8% correct. Most of the errors occurred on filler items. For the priming sentences, all requiring a “no” response, the error rate was 1.6%.

3.3. Discussion

Using an indirect-memory-based priming procedure, Experiment 2 demonstrated that priming persists over intervening, unrelated utterances. The magnitude of priming with the incidental procedure was no smaller than when primes were experienced under memory instructions; in fact, the effects were numerically larger and tended to increase over lags. These findings argue against an interpretation of persistent priming as a byproduct of intentional memorization, study, or internal monitoring of the priming sentences.

Persistence from transitive priming was again marginally weaker than dative priming. The sentence-type difference seen here and elsewhere in the literature (Bock, 1986; Hartsuiker & Kolk, 1998) implies that sentence structures may not be equivalent in their susceptibility to priming.

4. General discussion

The experiments demonstrated structural persistence from comprehension to production that was comparable in magnitude to persistence from production to production. This adds to the evidence for crossmodality persistence. Most important, there were few discernible differences in the strength of persistence or its endurance over the long stretches of unrelated linguistic events that separated primes from target production.

Putting the results for comprehension priming together with those from production priming to give a broad picture of how structural persistence survives intervening events, Fig. 5 shows the combined pattern for sentence types and modalities from the present experiments and Experiments 1 and 2 in Bock and Griffin (2000). The 95% confidence intervals for the differences between the primed and unprimed conditions were calculated from the mean squared error for the persistence effect, with the error terms derived from analyses at each lag for the 36 items common to all lags in all experiments. At lag 0, the confidence interval reflects the analyses of the long- and short-lag item data in all experiments (encompassing five sets of data); at lags 1 and 2 they reflect the results from the short-lag lists (three datasets) and at lags 4 and 10 the results from the long-lag lists (two datasets). Only at lag 10 does the width of the confidence interval (.056) begin to exceed the magnitude of the condition differences (.055). The differences (and confidence intervals) at lags 1 through 4 were as follows: Lag 1, .079 (.035); lag 2, .076 (.043); and lag 4, .071 (.044).

Paradoxically, the magnitude of the persistence effect was smallest at lag 0, with a difference of .044 (confidence interval of .032). This weakness remains to be
explained. Other experiments have found similarly weakened immediate effects, notably for anterograde amnesics (Ferreira et al., 2006). One speculation about this immediate weakness attributes it to interference from efforts to remember the prime when the cover experimental task involves direct memory (cf. Poldrack et al., 2001). Consistent with this speculation, comparison of the immediate conditions on the short-lag lists in Experiments 1 (direct-memory priming) and 2 (indirect-memory priming) reveals effects at lag 0 of .006 and .047, respectively.

The major exception to the uniformity of the observed persistence was associated with differences between sentence types. As in previous work, persistence from transitive was weaker and more variable than persistence from datives (cf. Bock, 1986; Hartsuiker & Kolk, 1998). One of the factors responsible for this difference is fairly well understood. The two objects in dative sentences almost always contrast in animacy, with the direct object typically being inanimate and the indirect object typically being animate. Because animacy differentials can boost persistence above the structural effect (Bock, Loebell, & Morey, 1992), at least some of the disparities are likely to be attributable to the impact of conceptual support on the occurrence and duration of persistence (Griffin & Weinstein-Tull, 2003). This is not the whole story, however, because structures other than transitives and datives show immediate as well as longer-term persistence in the absence of conceptual or lexical enhancement (Ferreira, 2003; Konopka & Bock, 2005). It is as yet unknown how much or how long conceptual similarities affect persistence over interruptions by time and other factors.

A second component of the sentence-type difference may be linked to the distributional restrictions of the dative (Goldberg, 1995). Whereas almost all transitive verbs can be passivized, dative verbs are more limited in their structural possibilities (e.g., donate very rarely appears in the double-object dative). In addition, the alternative

---

**Fig. 5.** Overall proportions of targets produced across lags, pooled over experiments, sentence types, and priming modalities (error bars represent halfwidths of 95% confidence intervals for pairwise planned comparisons).
dative constructions have different semantic and statistical associations (Ferreira, 1996; Gropen, Pinker, Hollander, Goldberg, & Wilson, 1989). In theories where distributional learning supports the use and acquisition of syntactic structures (e.g., Chang et al., 2006), differential restrictions can serve to differentiate constructions in the syntactic system and contribute to variations in the strength of persistence.

The apparent indifference of structural persistence to the modality of experience (whether comprehended only or both comprehended and produced) stands in fairly striking contrast with the fragility of repetition priming in other domains. For example, repetition priming for single words is substantially reduced when the modality of initial exposure differs from the modality of test. Words presented visually show substantial priming when tested with a visual word-fragment completion task; words presented auditorily show a fraction of the priming when tested in the same way (Roediger & Blaxton, 1987). Obviously, the contrast between visual and auditory presentation differs from the contrast between heard and reproduced primes in several respects, among them the sheer number of tokens experienced during prime presentation and the act of generation. Because either of these things (additional tokens or generation) might be expected to enhance persistence (Brown, Jones, & Mitchell, 1996; Slamecka & Graf, 1978), especially if persistence has any basis in episodic memory, the absence of impact is all the more puzzling.

The leading explanation for the indifference of structural repetition to modality changes is the abstractness of the structural representations or procedures responsible for persistence. Language structure must be abstract and modality-neutral (or at least richly shared across modalities) in order to support normal communication: Speakers and listeners, writers and readers, understand one another when they know the same language and have command of its structural resources. Although structural persistence is a good example of how priming may affect the abstractions that underpin language use, there are indications that it is not the only one. Joordens and Becker (1997) found that selective priming of the semantic features of words yielded effects that survived over intervening unrelated words, unlike priming from the associative and categorical relationships between words.

Because of the durability and abstractness of structural persistence, the mechanism behind it is more readily understood as a form of learning than as a type of episodic activation or a short-term memory effect. The implicit nature of the learning is attested in evidence that individuals suffering from anterograde amnesia exhibit structural persistence at the same level as normal speakers (Ferreira et al., 2006). This fits well with persistence modeled in terms of procedural learning that arises during structural assembly in sentence formulation, as in Chang et al. (2006). Consistent with the present results, the Chang model instantiates the hypothesis that comprehension and production make use of a shared sequencing system, implemented as a simple recurrent network. Although the full architecture includes other systems that may not be shared, the common sequencing system can explain the similar magnitudes of persistence from comprehended and produced primes.

The results of the present experiments raise many questions, some of which stem from limitations of the work. In order to elicit natural speech, the paradigm does not
place prior constraints on what speakers say, and this extemporaneity adds noise and subtracts data. As a result, the priming effects are very coarse. Other methods for eliciting priming address these problems (Branigan et al., 2000; Chang, Bock, & Goldberg, 2003; Cleland & Pickering, 2003; Pickering & Branigan, 1998), and permit some questions about duration and modality effects to be addressed with more sensitivity. In one relevant study, Cleland and Pickering (2006) show that a different kind of cross-modal relation (whether priming was from written to spoken production, or vice-versa) yields no difference in the magnitude of immediate priming effects, in line with the present results.

Another limitation to the findings is the delay in the measurement of persistence, which has a number of consequences. An important one in the context of questions about duration is the inability to capture early changes in the strength of the effects, which may be substantial (Szmrecsanyi, 2004, 2005). Related to this is the nature of the measure, which taps only the asymptotic variations in which structure is produced, and not how it is produced. Other measures, including assessments of timing (Smith & Wheeldon, 2001) and disfluency, have the potential to capture more transient effects.

Even so, as they stand, our results are more compatible with a user-friendly version of the strong competence hypothesis. The claim is that the same structural knowledge or structural resources support language production and language comprehension. The findings stand as a challenge and an impetus for further work. They show once more that structural persistence is not transient, undergoing no sharp decline over brief stretches of time and other events. They replicate other demonstrations of persistence from structures experienced in comprehension to structures formulated in production, and add new evidence that this crossmodality persistence can be as strong and as durable as intramodal persistence. Experiment 2 substantiated the claim that the durability of persistence does not depend on special efforts to remember primes, instead hinting that extra mnemonic effort may be detrimental to persistence. This is potentially problematic for episodic or explicit-memory-based explanations of structural learning (Kaschak & Glenberg, 2004), and reduces the likelihood that persistence within modality-specific monitoring systems underlies the effects.

In good measure, the significance of the results is in their implications for those facets of human language that depend on flexibility in the mechanisms behind language performance. One of these is language comprehension. The theoretical perspective offered here predicts that structural persistence should occur during comprehension as well as in production, and more importantly that produced and comprehended primes should be equally effective in altering input processing. Such persistence would be predicted by a theory that posits fully shared structural resources across the modalities of language use (Pickering & Garrod, 2004). Although this remains to be fully evaluated, there is at least emerging evidence that structural persistence can be observed within comprehension (Branigan, Pickering, & McLean, 2005; van Gompel, Pickering, Pearson, & Jacob, in press).

Our results also have implications for language change and language acquisition. Because experience with unfamiliar structures can lead to their incorporation into a
speaker’s repertoire (Kaschak & Glenberg, 2004) and to increased use in normal speaking (Szmarcyni, 2004, 2005), even across languages (Hartsuiker et al., 2004; Loebell & Bock, 2003), structural persistence can be a vehicle of changes in language over long stretches of time. In acquisition, structural persistence supports the generalization of old structures to new utterances. Depending on how the rudiments of structure emerge and develop, structural persistence as a form of implicit learning offers a scaffold for children (Brooks & Tomasello, 1999; Huttenlocher et al., 2004; Savage, Lieven, Theakston, & Tomasello, 2003, 2006) and adults alike. If what persists is not only abstract but generalizable over the modalities of language use, it supports structural learning from hearing language as well as from speaking it.

References


