Causal priming: How a language production mechanism guides representation

Caitlin M. Fausey (cmfausey@psych.stanford.edu)
Psychology Department, Stanford University
Stanford, CA 94305 USA

Neal Snider (snider@stanford.edu)
Linguistics Department, Stanford University
Stanford, CA 94305 USA

Lera Boroditsky (lera@psych.stanford.edu)
Psychology Department, Stanford University
Stanford, CA 94305 USA

Abstract

How do people talk about causal events? One constraint on causal event descriptions may be the local linguistic environment in which the description occurs. In this paper, we report results from one corpus study and one experimental study that examined the role of priming in the production of agentive language such as "*She broke the vase*" and non-agentive language such as "*The vase broke*". In both natural language use as well as in a more constrained experiment, English speakers were more likely to describe a causal event using the same kind of expression as had been used in a previous causal event description than to change the form of their description. Given evidence that people remember and reason about causal agents differently depending on whether they comprehend agentive or non-agentive language (Fausey & Boroditsky, 2007; submitted), understanding the language production mechanisms that create agentive and non-agentive linguistic environments is an important piece of a broader puzzle about how people integrate linguistic and non-linguistic sources of information in the representation and processing of causal events.

Keywords: Priming; causal language; causal explanation

Introduction

Imagine a woman tidying up rooms in her house, listening on the phone to a friend who is recounting the morning events with her toddler: "*First the cheerios stuck to his face, then the bib slipped off, and then the spoon dropped on the floor*". Upon entering her own daughter's bedroom, the woman sees a broken vase on the floor. Does she exclaim "*The vase broke!*" or "*She broke the vase!*"? Would her reaction be different if her friend had just told her, "*First, he stuck the cheerios to his face, then he slipped off the bib, and then he dropped the spoon on the floor*"?

Given the ubiquity of causal events in people's lives, and variation in their possible description, examining causal language use is likely to provide interesting insights about language production and comprehension, and also the relations between event construal and language. In this paper, we test whether local linguistic context influences how English speakers describe causal events.

Why examine causal language production? This question about language production is motivated by findings that the form of a causal event description impacts other inferences about the event. For example, event descriptions influence whether people more readily imagine causal antecedents or consequents (e.g., Majid, Sanford & Pickering, 2007), infer one or two events in a causal chain (e.g., Wolff, 2003), or attribute outcomes to the described agent or patient (e.g., Semin & Marsman, 1994). Recent findings also suggest that agentive event descriptions like "*She broke the vase*" lead people to punish agents more harshly, and to better remember agents, than do non-agentive descriptions like "*The vase broke*" (Fausey & Boroditsky, in preparation; submitted). Why do different event descriptions lead to different reasoning and memory for causal events?

The effects of causal language on causal event construal are presumably based in people's experience learning and using language across various non-linguistic contexts. Beyond this broad statement, however, we understand very little about the mechanisms that lead to the interesting language comprehension effects. For a more detailed account, we may need to better understand how causal language is produced in a range of non-linguistic contexts. Examining production may help us to understand the contexts in which agentive and non-agentive expressions are likely to be used and when they are especially likely to influence causal event construal. One relevant production mechanism may be linguistic priming.

Priming. Previous research has suggested that the form of an utterance is influenced by the form a previous utterance. For example, people are more likely to describe an event using the passive voice if they have recently heard a passive description than if they have recently heard an active description (e.g., Bock, 1986; Sankoff and Lanberge 1978). We purposely use the phrase "linguistic" priming, and not "structural" or "syntactic" priming in this paper because we remain agnostic about the nature of the priming effects that we report. Previous research has suggested that the production of active versus passive transitives (e.g., Bock, 1986; Sankoff and Lanberge, 1978), prepositional versus object datives (e.g., Bock, 1986) and relative clause attachment (e.g., Scheepers, 2003) may be primed. Could the

agentive and non-agentive language use that impacts attribution and memory for causal agents also be primed?

Priming causal event descriptions in natural language and experiments. One may examine linguistic priming from at least two perspectives. First, it is interesting to consider a particular kind of language use – such as causal verbs that may appear in both agentive and non-agentive expressions – and analyze patterns in language use *per se*. Second, it is interesting to consider the perspective that people might take on causal events – such as attending to causal agents or not – and use language production as one behavioral indicator of a particular perspective. In this paper, we examine priming in both ways, using a natural language corpus to address the first question and a more constrained experiment to address the second question.

Both corpus analyses and experiments have advantages and disadvantages as paradigms for examining language production. Corpus analyses are more ecologically valid and also may include more tokens per speaker than constrained experiments. With respect to priming, it may be easier to find evidence of priming in natural language use than in experimental language use because of many levels of contexts that could mutually reinforce each other to lead to priming (e.g., Pickering & Garrod, 2004). Though experimental contexts are often limited to few items and utterances per participant and are rarely fully conversational, it is possible to control the non-linguistic (and linguistic) context of experiments and therefore they are amenable to testing precise mechanistic hypotheses. Converging evidence from both natural and experimental language use, of course, would most strongly provide evidence for the behavior of interest.

Therefore, we did both a corpus study and an experiment. Both are analyses of fairly unconstrained language use. The corpus study used natural dialogue that was constrained only by the general topic of conversation. The experiment merely required participants to describe a picture however they saw fit. Many experimental priming paradigms aim to study language production, but tightly constrain possible utterances (e.g., sentence completion in Pickering & Branigan, 1998). The data presented here are as naturalistic as possible.

By examining whether agentive and non-agentive expressions may be primed, we broaden our understanding of what kinds of expressions are susceptible to local linguistic context effects and also reveal a potential mechanism for generating descriptions that may influence one's own and others' subsequent reasoning. Language production is both a reflection of the current event representation as well as influential in continuing to construct that representation and influence further processing. Can linguistic context influence the production of agentive and non-agentive language?

Study 1: Causal language in the wild

This study investigates whether causal language is primed in the course of natural language use. The data come from a corpus of conversations. If priming is evident in natural language data, it will be strong evidence that representations involved in processing causal events persist over time.

Data

The corpus used in this study was the manually parsed Penn Treebank (release 3, approximately, 800,000 words, see Marcus et al., 1999) portion of the Switchboard corpus (Godfrey et al., 1992), a syntactically-annotated corpus of conversational speech. All intransitive and monotransitive forms of 24 alternating agentive/non-agentive verbs were automatically extracted using *tgrep2*[1]. These 24 verbs were chosen because they were found to alternate in a cross-linguistic study of agentive/non-agentive verbs by Haspelmath (1993). Haspelmath's study actually examined 31 verbs, but only 24 of theses are found in the Treebank Switchboard. The 24 verbs in the study are: *begin, boil, break, burn, change, close, connect, develop, dry, fill, finish, freeze, gather, improve, melt, open, rock, roll, sink, split, spread, stop, turn, wake*, with *change* occurring most often in the dataset (266 times) and *rock* occurring the least often (2 times), with a mean of 37.2 occurrences per verb. There are a total of 889 tokens in the dataset.

Each token was automatically coded as agentive if it occurred monotransitively in active voice (where the agent of the action was explicitly mentioned), and automatically coded as non-agentive if it occurred intransitively (with no agent explicitly mentioned). Next, the form of the prime construction (agentive or non-agentive) was added to the database. The prime is defined as the nearest previous usage of one of the 24 alternating verbs. Primes that occurred in the same utterance as the target were excluded, as these are more likely to be speech errors or stylistic repetitions. Prime-target distances ranged from 1 turn to 332 turns, with a mean of 60.78.

Method

To test the effect of priming in choice of agentive versus non-agentive descriptions, the data set was analyzed with mixed logit models (Bates & Sarkar, 2006; Breslow & Clayton, 1993). Mixed logit models can be thought of as an extension of logistic regression that include modeling of random effects. Inclusion of random effects, such as speaker or participant, is necessary to generalize beyond the participants in the current data set (Clark, 1973).

The dependent variable was structure choice, agentive or non-agentive, with agentive coded as the positive choice. The primary independent variable of interest was the prime construction, which had three levels: agentive, non-agentive, and no-prime. The no-prime cases occur when an agentive or non-agentive verb occurs at the beginning of the conversation, so the prime is unknown. The no-prime tokens act as a baseline, allowing one to determine whether there are independent effects of both agentive and non-agentive primes relative to a baseline level of agentive versus non-agentive use. In addition to the fixed effect of the prime

[1] tedlab.mit.edu/~dr/Tgrep2/

construction, random effects of speaker and target verb lemma were added. Controlling for the random effect of speaker enables one to generalize beyond the speakers in this sample, allowing each speaker a different base rate of agentive language use. There were 264 different speakers in this sample, with a mean of 9.3 tokens per speaker. The random effect of target verb lemma was added because a corpus study of structural priming (Gries, 2005) has argued that priming only occurs on target verbs that are not strongly biased. Some verbs in the database were quite strongly biased towards agentive use (e.g., *fill* occurred agentively in in 88% of 34 tokens) or non-agentive use (e.g., *dry* occurred non-agentively in 80% of 15 tokens). Including a random effect of verb lemma controls for the fact that target verbs have different rates of usage in agentive and non-agentive forms.

Results

We report the coefficient for the independent variable and its levels of significance. Coefficients in logistic regression models are given in log-odds (the space in which logistic models are fitted to the data). For categorical factors, significant positive coefficients mean that the positive level (agentive in this case) is more likely in the tested level of the variable than in the other level. For example, if the coefficient of prime=agentive is positive, then having a prime that is agentive makes an agentive form more likely in the target description. Negative coefficients mean the opposite. Below, we also report the difference in odds between conditions (as the name suggests, odds are simply $e^{log-odds}$). Odds range from 0 (for proportions of 0) to positive infinity (for proportions of 1), with proportions of 0.5 corresponding to odds of 1. Odds are a multiplicative scale, so we talk about an x-fold increase or decrease in odds between conditions.

The coefficients in log-odds and standard errors associated with the remaining factors are given in the second and third column of Table 1. The corresponding odds coefficients are given the fourth columns. The fifth and sixth columns summarize the Wald's Z statistic, which tests whether the coefficients are significantly different from zero. Finally, the last two columns give the χ^2 over the change in data likelihood (Δ_x (Λ)) associated with the removal of the predictor (x) from the final model. The latter test is more robust against collinearity in the model (Agresti, 2002). The χ^2 value, which literally corresponds to the difference in the model's data likelihood without the predictor, can be seen as a measure of the predictor's importance in the model. The Wald-test is included because it implicitly tests the directionality of the effect (unlike the χ^2 over the change in data likelihood).

As Table 1 illustrates, people are 1.9 times as likely to describe an event using agentive language when they have recently heard a similarly agentive event description, compared to when there was no prime. Further, they are 29% less likely to describe an event agentively if they have just heard a non-agentive description. The priming factor

Table 1. Mixed logit model results of corpus analysis.

Predictor	***Parameter estimates***			***Wald's test***		$\Delta_x(\Lambda)$***-test***	
(independent variable)	Log-odds	S.E.	Odds	*Z*	*p*	χ^2	*p*
Agentive prime	0.66	0.17	1.9	3.8	<.001	12.8	<.001
Nonagentive prime	-0.34	0.17	0.71	-2.0	<.05		

significantly improves the model even while including random control factors like speaker and target verb. This effect is represented graphically in Figure 1. Speakers were more likely to describe events using agentive language after agentive primes (n = 174) than after non-agentive primes (n = 99) and more likely to describe events using non-agentive language after non-agentive primes (n = 117) than after agentive primes (n = 81).

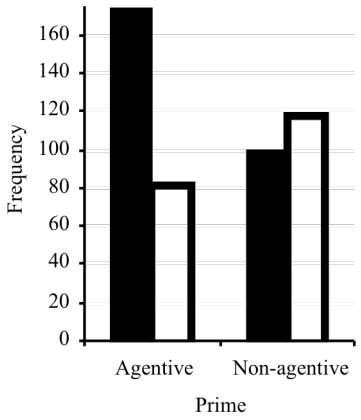

Figure 1: Causal event descriptions in natural language.

Discussion

This study provides the first psycholinguistic evidence that people are more likely to use agentive language in natural conversation if they have previously heard an agentive description, even when controlling for effects of individual speakers and verbs. The same is true for priming non-agentive descriptions. These results lends support to the notion that priming is a psychologically real mechanism in actual language use and could contribute to reinforcing particular patterns of description both locally as well as over a lifetime of language use.

Although same-utterance repetitions were excluded in this analysis in order to minimize the presence of non-priming related repetitions, structures are repeated in natural language use for a variety of stylistic reasons that may not be related to priming (e.g. speech errors, stylistic repetition, etc., see Szmrecsanyi, 2005). In order to provide more evidence that the use of agentive descriptions can indeed be primed, we also conducted a more constrained experiment.

Study 2: Causal language in the lab

The way that people describe causal events during natural language use appears to depend on the local linguistic environment: With respect to agentive and non-agentive forms, a particular causal event description is more likely to match the form of the previous causal event description than to change forms. The results of our corpus analysis suggest that agentive and non-agentive descriptions of causal events can be linguistically primed. In natural language corpora, however, each target description can only be primed by either an agentive prime description or a non-agentive description. That is, it is impossible to manipulate the prime status for any given target description. Because different causal events were described following agentive and non-agentive language, patterns revealed by our corpus analysis provide only weak evidence that linguistic primes may alter event construal. Could descriptions of the same causal event be changed by linguistic primes? An experimental paradigm constrains the environment of the language user so that answering this question is possible: participants in our second study described the exact same event following minimally-paired linguistic primes.

In this study, participants viewed a pair of pictures depicting the beginning and end states of a causal event – either paint-splattering or vase-breaking – and then they described the event. We hypothesized that people's descriptions of the causal event would match the form of an unrelated prime sentence that they had read prior to viewing the event depiction. Assuming that people's language production is related to how they construe the beginning and end state pictures of the event, this paradigm may reveal one way in which local linguistic context shapes causal event construal.

Our corpus analyses suggest that an implicit priming mechanism may operate during natural language use of agentive and non-agentive expressions. Can the production of these expressions be linguistically primed even when participants describe the very same causal event?

Participants

233 students at the University of California, Merced participated for course credit.

Materials

Linguistic primes. Participants read one prime sentence, drawn from a set of eight sentences. The full set consisted of an agentive and non-agentive description of four events (see Table 2).

Visually-depicted events. Participants described a pair of pictures depicting either a paint-splattering event or a vase-breaking event. Each visual depiction consisted of a beginning-state frame and an end-state frame (see Figure 2).

Design

After reading either an agentive or a non-agentive prime sentence, participants viewed a pair of pictures and described the visually-depicted event. Primes and pictured events were fully crossed.

Table 2. Linguistic prime stimuli.

Agentive primes	**Nonagentive primes**
He popped the balloon.	The balloon popped.
He opened the umbrella.	The umbrella opened.
He unfastened the necklace.	The necklace unfastened.
He blew out the match.	The match blew out.

The dependent measure was whether people described the pictured event using agentive or non-agentive language. In addition to the influence of the agentivity status of the prime, item effects of the four prime sentences and the two pictured events were also analyzed.

Procedure

Participants completed a two-sided survey that was presented among several other unrelated surveys. On the front side of the page, participants read one sentence and were asked to "please continue the story for another sentence or two" on blank lines that appeared below the prime sentence. This encouraged them to actually process the prime sentence. On the back side of the page, participants saw the beginning and end of a causal event and were asked to "please describe this event".

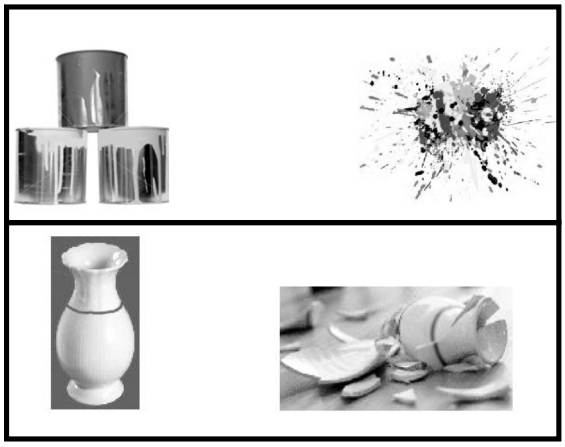

Figure 2: Pictured events.

Results

Coding. Each event description was coded as agentive or non-agentive. Transitive sentences, both active (71%) and passive (29%), were coded as agentive. Intransitive sentences were coded as non-agentive. All sentences were coded by the first author and by an independent coder, with high reliability (κ = .87). See Table 3 for example responses.

Analyses. Data were submitted to a chi-square analysis. As predicted, people were more likely to describe an event using agentive language following an agentive prime (n = 73) than following a non-agentive prime (n = 53), and to describe an event using non-agentive language following a non-agentive prime (n = 67) than following an agentive prime (n = 40), $\chi^2(1) = 9.79$, p = .002 (see Figure 3a).

Table 3. Example causal event descriptions.

Agentive responses
Somebody broke the vase.
The vase was on a table until someone knocked it down.
Someone took the paint and splattered it on the wall.
Somebody knocked the paint over and made a mess.
Nonagentive responses
The vase broke.
The pretty antique vase broke and shattered into pieces.
Paint cans exploded.
The paint was in the buckets then it spilled onto the wall.

The effect of agentive and non-agentive primes held for both the vase-breaking and the paint-spilling events, though it is clear that English speakers had different overall biases with respect to agentivity in describing these events. As shown in panels (b) and (c) of Figure 3, English speakers preferred to describe the paint-splattering event using agentive language and the vase-breaking event using non-agentive language. Note, however, that the effect of an agentive or a non-agentive prime was in the same direction for both events, marginally significant for paint-splattering ($\chi^2(1) = 2.90$, $p = .089$) and signficant for vase-breaking ($\chi^2(1) = 6.71$, $p = .01$). Further, all four prime sentences influenced causal event descriptions in the same way.[2]

Discussion

In a controlled experimental setting, people were more likely to talk about the agent of a causal event in the presence of an unrelated event description that was agentive rather than non-agentive. Descriptions of the very same visual stimulus depended on the local linguistic context.

General Discussion

Our studies provide evidence that agentive and non-agentive causal event descriptions can be primed. In both experimental settings and in natural dialogue, people are more likely to mention an agent when describing a causal event when they have recently encountered an agentive event description than when they have recently encountered a non-agentive event description. This result is especially compelling because it comes from converging evidence in both conversational and controlled experimental settings.

One implication of these findings is the need to consider language production in theories about the relationship between language use and causal reasoning. Priming clearly influences linguistic choices in everyday conversation using agentive and nonagentive event descriptions and other research suggests that these expressions impact further reasoning about causal events (e.g., Fausey & Boroditsky, 2007; submitted; in prep). Thus, priming may be a mechanism by which linguistic experience influences reasoning more generally. . Future experiments will

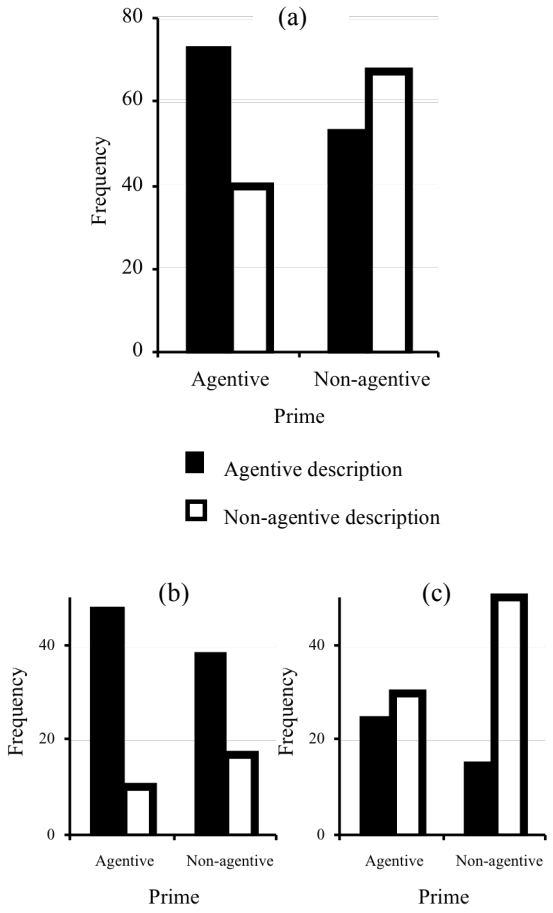

Figure 3. Primed descriptions of causal events.
(a) overall, (b) paint-splattering,(c) vase-breaking.

combine the to-date separate approaches of production and comprehension in order to more directly understand the relationship between linguistic priming and causal reasoning.

This work has further implications for priming research in general, as we have shown (particularly in the corpus study) that intransitive expressions do indeed prime. Intransitives have often been used as fillers and baselines in syntactic priming experiments (e.g., Bock & Griffin, 2000; Griffin & Weinstein-Tull, 2003), but our results show that in some contexts intransitives do prime. The nature of the non-linguistic context may partially explain the potency of intransitive primes in our experiment: Language users saw an event depiction without an agent and thefore could felicitously be described by an intransitive expression. Interestingly, participants could infer a causal agent of the event, and were more likely to do so in the presence of an unrelated agentive expression. In future experiments, we will examine a wider range of prime and target stimuli and in future corpus studies, we will explore how intransitives of non-causal verbs may affect causal language use.

What influences the description of causal events? Our results suggest that whether someone describes a causal

[2] Only one prime-event stimulus combination deviated from the reported effects: Following the umbrella prime sentence, only three of 28 people described the paint-splattering event using non-agentive language.

event using agentive language like "*She broke the vase"* or non-agentive language like "*The vase broke"* depends in part on the local linguistic context.

Acknowledgments

We thank the citizens of Cognation and SPLAT! attendees for helpful discussions of this work. We thank Teenie Matlock, Justin Matthews and Vijay Vanchinathan for help with data collection and coding.

References

Agresti, A. (2002). *Categorical Data Analysis*. New Jersey: Wiley & Sons.

Bates, D., & Sarkar, D. (2006). lme4: Linear mixed-effects models using S4 classes. R package version 0.995-2.

Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355-387.

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.

Breslow, N.E., & Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association 88 (421)*, 9–25.

Clark, H.H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12(4)*, 335–359.

Fausey, C.M., & Boroditsky, L. (in preparation). Speaking of accidents: "He did it" invites more punishment than "It happened".

Fausey, C.M., & Boroditsky, L. (submitted). English and Spanish speakers remember causal agents differently. Paper submitted to the *30th annual meeting of the Cognitive Science Society*, Washington, D.C.

Fausey, C.M., & Boroditsky,L. (2007). Language changes causal attributions about agents and objects. Poster presented at the *29th annual meeting of the Cognitive Science Society*, Nashville, TN.

Godfrey, J., Holliman, E., & McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. In *IEEE ICASSP*. 517-520.

Gries, S.T. (2005). Syntactic priming: A Corpus-based Approach. *Journal of Psycholinguistic Research, 34(4)*, 365–399.

Griffin, Z.M., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language*, 56, 537-555.

Haspelmath, M. (1993). More on the typology of inchoative/causative verb alternations. In *Causatives and transitivity* (Ed. B. Comrie & M. Polinsky), John Benjamins: Amsterdam.

Majid, A., Sanford, A. J. & Pickering, M. J. (2007). The linguistic description of minimal social scenarios affects the extent of causal inference-making. *Journal of Experimental Social Psychology, 43*, 918-932.

Marcus, M.P., Santorini, B., & Marcinkiewicz, M.A. (1994). Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics, 19*, 313–330.

Pickering, M.J., & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39(4)*, 633-651.

Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169-226.

Sankoff, D., & Laberge, S. (1978). Statistical dependence among successive occurrences of a variable in discourse.In D. Sankoff and S. Laberge (Eds.): *Linguistic Variation: Methods and Models*. Academic Press, New York.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, 89, 179—205.

Semin, G.R., & Marsman, G. (1994). "Multiple inference-inviting properties" of interpersonal verbs: Event instigation, dispositional inference, and implicit causality. *Journal of Personality and Social Psychology, 67(5),* 836-849.

Szmrecsanyi, B. (2005). Language users as creatures of habit: a corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory, 1*, 113–149.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition, 88*, 1-48.