Words with Friends:
Effects of Associative and Semantic Relationships on Subject-Verb Agreement Errors During Sentence Production

by Darrell J. Penta

B.A. in English and Psychology–Sociology, University of Massachusetts Boston
M.A. in Applied Linguistics, University of Massachusetts Boston
M.A. in Psychology, Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 21, 2017

Dissertation directed by

Dr. Neal J. Pearlmutter
Associate Professor of Psychology
Dedication

For Jina,
who shows me every day
the real meaning of dedication
Acknowledgements

Think where man’s glory most begins and ends,
and say my glory was I had such friends.
— Y.B. Yeats

If the past five years have taught me nothing, they’ve taught me that I am loved, supported, and inspired by the greatest group of people in the world. I’d like to thank each of those people now, but then I’d need to publish this dissertation in several volumes, and trust me, nobody wants that. There are far too many deserving people that I’m not able to thank by name here. I thank them now in spirit.

To my advisor, Dr. Neal J. Pearlmutter: Thank you now and in perpetuity for being my advisor. I am humbled to have had the chance to work with you. I simply can’t convey how much I admire you as a person and as a scientist, nor how much I’ve learned from you. Thank you for imparting me with a small fraction of the staggering amount of insight that you have about psycholinguistics, language, research, and statistics. And late 80s/early 90s Cambridge trivia. You faced the awesome challenge of turning a language arts guy into a language science guy, and you did so without losing your sense of humor. It was always clear to me that I could come to you at any time and for any reason, and that you’d be there to help me no matter how long it took. I sincerely hope I can return the favor to you some day. Thank you again, Neal. For her vicarious support, thank you also Jodie Silverman. To Ruby Pearlmutter, whose pictures made me smile whenever I visited her dad’s office: Thanks!

To my dissertation committee members, Dr. Joanne Miller and Dr. Judy Hall: A heartfelt thank you to you both. You are the “dream team” of dissertation committees, which is perfect because there’ve been a few nightmares along the way. You probably don’t hear this enough (at least not from me): Our department is a special place because you are both great at what you do. Joanne, I doubt very much if I’ll ever not be in awe of you. You told me when I first started the program that your door was always open, and I found that to be true on every occasion. Judy, you really deserve your own paragraph, and probably even your own chapter. But to squeeze a few thoughts into this tight space: The support that you’ve given me since I started the program has meant the world to me. And if you ever feel like you just want to have a ten hour conversation about anything, look no further than me.
To Yvonne Malcolm, Maribel Pereira, and Gina Gonsalves: Thanks, times a million. To my Cognition Area people, past and present, each of whom holds a special place in my heart: My deepest appreciation to you all. You all make *geek* seem *chic*. Please think of me fondly every time you use the Keurig machine.

To the SPLers: A most humble thank you. Amy DiBattista, how lucky I was to have landed in your office five years ago! You are a marvelous person, and I’m proud to consider you as my friend. Yingzhao Zhou, I’ve truly enjoyed getting to know you, and I look forward to hearing of the many successes you’re sure to have in the future.

I am eternally indebted to the phenomenal RAs who worked so tirelessly on my research, including: Ammielia Cash, Shreya Divati, Catherine Balboni, Camila Demaestri, Rylie Ellam, Leila Habib, Monika Izdebski, Mikhail Kuznetsov, Josie Lewis, Erin Lopez, Alex More, Geetha Nichanametla, Kathy Niu, Ian Redpath, Talis Reks, and Maria Van Thienen. To Liz Dovenberg, Gizem Kullukçu, Zoë Winkworth, and Claudia Ortiz: I hope you learned as much from me as I learned from you all.

To the many teachers who have transformed me: Thank you. Thank you especially to my previous master’s thesis advisors, Dr. Charles Meyer and Dr. Corinne Etienne, for investing so much time and energy in me, and to my English professor, Dr. Neal Bruss, for, among many other things, telling me to learn an “employable skill, like computer programming.” Professors Lloyd Schwartz, Askold Melnychuk, Thomas O’Grady, and the late Pepi Leistyna: Thank you.

To my loyal friends Mike DiGregorio, Joe Cirino, John Pollock, Chris and Elizabeth Sweeney, Brian MacNeil, and Hyungmin Jun: Thanks. I wouldn’t trade you guys for the world.

To my amazing parents, Carol and John: Thank you, and I love you. You are both the most compassionate, unselfish, and generous people I’ve ever met. I have no idea how you do it. I am honored to call you my heroes, and blessed to call you Mom and Dad. By the way. Mom, I know this dissertation is no substitute for a grandchild, but you’re welcome to take it home and dissertation-sit it anytime. To my twin brother, Mike: I can’t imagine how uninteresting and lonely my life would be if I hadn’t been yoked to you from birth. To the world’s best sister, Dr. Lisa D’Amico, and the world’s best brother-in-law, John D’Amico: Yes, I’ll come have a drink with you now. To my non-twin brothers, Jay, Jeff, and Greg. Sorry you didn’t each get your own sentence here, but I’m pretty sure you’re not planning
To my loving aunt, Barb Austin: You bring a special kind of joy to my life. Thank you for everything.

And to my dear uncle and friend, Dr. James Austin: Well, old chum, if you’re reading this, then I’d say we’ve both managed to achieve another little victory in spite of ourselves! For me, being able to tell you—I mean, you and especially you—that I’ve made it this far means that at least two of my life’s wishes have been granted. I want you to know that the times I have spent in your company are among the happiest moments of my life. You have no idea how much I adore you. Thank you for being my teacher, my friend, and on so many occasions, my inspiration. Gracias, amigo. I don’t know what happens in the next chapter of our little story, but I have a good feeling that it’s going to involve a tranquil sea, a steady breeze, and a conversation that goes on until the bosses tell us to give it rest. If by some stroke of good luck you get to all the good parts before me, just grab my dad and have ‘em get some beverages ready.

Finally, to the lovliest person I have ever known, my very beautiful, very loving, and very patient wife, Jina Park Penta: 사랑하는, 제 아내 지은에게 깊은 감사의 말을 전하고자 합니다. I owe you the biggest thank you of all. 지난 5년 동안 당신의 현신적인 희생이 없다면, 제가 추구하고자 하는 꿈을 이루지 못했을 것입니다. You are, and will always be, the brightest part of my everyday. 힘들고 지쳐서 포기하고 싶을 때마다, 당신의 존재는 저에게 큰 힘이 되었으며 학위과정을 지속할 수 있게 하는 원동력되었습니다. 당신이 없이는 오늘 이 자리가 없었던 것입니다. 때문에 저는 본 학위과정을 당신과 함께 끝마친 것이라고 생각합니다. I’m proud of what I accomplished, but I’m only here because of your many sacrifices. You are my anchor. I may not say it or show it every day, but I believe I’m the luckiest person in the world to have you next me. 가장 신뢰할 수 있는 친구로써 그리고 많은 영감을 주는 동료가 되어 주셔서 감사합니다. 저에게 준 아낌없는 사랑에 대해 다시 한번 감사의 마음을 전합니다. 그 동안 당신에게는 쉽지 않았다는 것을 잘 알고 있습니다. 하지만 이전에 약속 했듯이, 저는 당신을 사모님왕비 만들거야.

And now, for the next two lifetimes at least, it’s all about you honey. 세상 그 누구보다 널 사랑해—Your 노비/오빠.
Abstract of Dissertation

The sentence production system transforms preverbal messages in the mind of a speaker into coherent grammatical utterances. During this process, which unfolds rapidly, the system has to link meaning information from the speaker’s message to appropriate lexical and grammatical information from the speaker’s memory. It usually does so with fluency and accuracy, but there are some predictable circumstances under which the system is prone to processing failures. Researchers have examined prototypical failures, including subject-verb agreement errors, to learn more about how the system coordinates the flow of information from the mind to the mouth of the speaker under routine conditions.

Subject-verb agreement in English refers to the conventional requirement that sentence subjects and verbs agree in grammatical number (e.g., The book is; The books are). Bock and Miller (1991) found that participants tend to make agreement errors when completing subject noun phrase sentence preambles (e.g., The key to the cabinets) where a singular head noun (HN; key) is followed by a plural local noun (LN; cabinets). This is referred to as the mismatch effect.

The three experiments in this dissertation used the preamble completion task to investigate influences on mismatch effects from two types of word relationships: association, which refers to the probability that two words will co-occur irrespective of their meaning (e.g., apple–computer; Penta & Pearlmutter, 2015); and semantic relatedness, which refers to similarity between two words on one or more dimensions of meaning (e.g., apple–orange; Barker, Nicol, & Garrett, 2004; Penta & Pearlmutter, 2015).

Penta and Pearlmutter (2015) found that preambles with associated HN–LN pairs produced larger mismatch effects than preambles with unassociated nouns. Experiments 1–2 examined
whether mismatch effects would differ when either the HN, or the LN, or neither noun of a
given preamble (e.g., *The highway by the new school(s)*) was associated with a pair of external
“prime” nouns (HN primes: *road, interstate*; LN primes: *campus, university*) presented prior
to the preamble in a sentence (Exp. 1) or at the end of a word list (Exp. 2). In Experiment 1,
mismatch effects were smallest when the prime nouns were associates of the HN and largest
when they were associates of the LN. This suggests that mismatch effects were modulated
by the associative relationship type. Experiment 2 replicated the mismatch effect, but there
were no significant differences in mismatch effects as a function of the associative relationship
type.

Experiment 3 asked whether mismatch effects would be larger for preambles with animacy-
matched versus -mismatched HN–LN pairs (e.g., *doctor–villager(s); hospital–village(s)*
vs. *doctor–village(s); hospital–villager(s)*) when the nouns were minimally related on other
semantic dimensions. Ratings data from a large online survey of synonymy were used to
minimize the amount of similarity between nouns that was not otherwise attributable to
differences in animacy relatedness. Agreement errors were produced more often following
animacy-matched nouns, but mismatch effects did not differ significantly between conditions.
This is consistent with past results (Penta & Pearlmutter, 2015) showing that increased
semantic relatedness for animacy-matched nouns does not predict larger mismatch effects (cf.
Barker, Nicol, & Garrett, 2001). The results of Experiment 3 suggest that animacy-based
relatedness does not have a strong influence on mismatch effects.
# Table of Contents

Dedication i

Acknowledgements ii

Abstract vi

List of Figures ix

List of Tables x

Introduction 1
  1.1 Grammatical and notional number in agreement 6
  1.2 Non-number related factors in agreement 10
  1.3 Semantic relatedness and association 12

Chapter 2: Effects of Association 20
  2.1 Experiment 1 26
  2.2 Experiment 2 37

Chapter 3: Effects of Semantic Relatedness 44
  3.1 Experiment 3 49

Chapter 4: General Discussion 59

References 68

Appendix A: Experiment 3 Stimuli 82

Appendix B: Synonymy Rating Instructions 84
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bock and Levelt’s (1994) model of the language production process.</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Barker’s (2001) schematic of activation patterns between the semantically related nouns <em>canoe</em> and <em>sailboat</em> and the unrelated nouns <em>canoe</em> and <em>cabin</em>.</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Adaptation of Bock and Levelt’s (1994) sample lexical-semantic network for the prime word <em>spider</em> and its associated target word <em>web</em>.</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>Experiment 1 mismatch effects by prime condition.</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>Experiment 2 mismatch effects by prime condition.</td>
<td>41</td>
</tr>
<tr>
<td>6</td>
<td>Experiment 3 mismatch effects by HN–LN animacy.</td>
<td>55</td>
</tr>
</tbody>
</table>
List of Tables

1. Experiment 1 error rates and response counts by condition. 32
2. Experiment 2 error rates and response counts by condition. 41
3. Experiment 3 stimuli and synonymy ratings by condition. 50
4. Experiment 3 error rates and response counts by condition. 55
Sentence production refers to the process through which a speaker’s internal thoughts are translated into external speech, and the cognitive faculties that underlie this process are referred to as the sentence production system. From a subjective standpoint, producing a sentence may seem unremarkable. Speakers occasionally have difficulty thinking of something to say, but once they have a message in mind, they tend to have very little difficulty generating a fluent utterance (e.g., Bock & Ferreira, 2014; Bock & Garnsey, 1999; Bock & Huitema, 1999). What makes this remarkable from a researcher’s standpoint, however, is that the sentence production system is able to rapidly and accurately translate messages into speech despite the fact that most utterances have novel configurations of semantic, grammatical, and phonological information (Bock, 1991).

The capacity of the system to support fluency without sacrificing flexibility during production rests largely on the fact that sentences are formulated incrementally and through a division of labor (cf. Bock, 1987b; Dell, 1985; V. Ferreira, 1996; Garrett, 1976, 1980; Kempen & Hoenkamp, 1987; Levelt, 1989). A general framework of the production system, from Bock and Levelt (1994), is presented in Figure 1. An important assumption reflected in Bock and Levelt’s model is that production processes are grouped into three different processing domains: Operations at the message level prepare the semantic content of a to-be-uttered sentence according to the meaning information that a speaker intends to convey; at the grammatical encoding level, words suitable for expressing the message are selected and ordered in phrases following the grammar of the speaker’s language; finally, at the phonological encoding level, the sound (or gesture) patterns of the sentence are specified as a motor plan for articulation (Bock, 1987b, 1995b, 1999; Bock & Griffin, 2000; Levelt, 1989, 1992).
Figure 1. Adapted from Bock and Levelt’s (1994) model of the language production process. The three main processing levels are the message level, grammatical encoding, which is further separated into functional and positional processes, and phonological encoding. Details about separate subprocesses within both the message level and phonological-encoding level are omitted here (and in the original depiction of the model). The downward pointing arrows represent the top-down, incremental flow of information through the system.
Separating out production responsibilities through an architecture such as the one shown in Figure 1 is advantageous in many respects. For example, it allows overt production of an utterance to be initiated before a fully-fledged production plan has been worked out (e.g., Konopka & Brown-Schmidt, 2014), and it enables the system to make on-the-fly adjustments to a plan in progress without bringing the entire process to a halt (e.g., V. Ferreira, 1996; Konopka, 2012). Because formulation is a dynamic process, however, these advantages do not come without significant processing costs. One issue, for example, is that the system must continuously solve the problem of how to form connections between elements of a sentence within and across processing levels (Bock, 1995b). This challenge is compounded by the fact that processing decisions within the system have to be made rapidly and in accordance with internal and external constraints that affect how, when, and what information can be processed at any point during formulation (Bock, 1982, 1991; Butterworth, 1980). One can appreciate the magnitude of this problem by considering that the conceptual components of a speaker’s message, which may be unordered and contemporaneous, must ultimately be linked to a temporally unfolding, linearly ordered sequence of speech sounds (Bock, 1987a; Lashley, 1951).

Sentence production research is concerned with understanding how the system manages to orchestrate the flow of information from the mind to the mouth of the speaker during normal utterance formulation. Because psycholinguistic processes are not directly observable, researchers have pursued questions about the production system by analyzing patterns in its output. A particularly valuable source of information to this enterprise is data from speech errors (e.g., Baars, 1992; Fromkin, 1971; Garrett, 1975). Errors are relatively rare occurrences in natural speech (e.g., Bock, 1991), and it has become increasingly common for researchers to elicit speech errors in the laboratory (Baars, 1992; Baars, Motley, & MacKay, 1975; Lapointe & Dell, 1989; A. S. Meyer, 1992).

The three experiments in this dissertation follow in the tradition of using speech error data in order to make inferences about how the system maintains and coordinates information
during normal utterance production. More specifically, they explore questions about how different types of relationships between words affect the production of subject-verb agreement errors. The experiments use a well-established sentence-completion task that is designed to elicit errors of subject-verb agreement in speakers’ sentences. The task is the same as in most studies (Barker et al., 2001; Bock & Cutting, 1992; Bock, Eberhard, Cutting, Meyer, & Schriefers, 2001; Bock & Miller, 1991; Eberhard, 1999; Gillespie & Pearlmutter, 2011, 2013; K. R. Humphreys & Bock, 2005; Solomon & Pearlmutter, 2004; Penta & Pearlmutter, 2015; Thornton & MacDonald, 2003; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco, Butterworth, & Semenza, 1995): Participants are presented with test preambles, which they repeat or read aloud and for which they provide an ending of their own creation; participants’ utterances are transcribed, coded, and analyzed; and error rates, calculated as the number of errors out of the total of number-inflected verbs produced in participants’ responses, are computed and compared to quantify differences in effects across conditions.

Speech errors in production research

The value of agreement errors and other types of speech errors to production researchers stems from the fact that they display a surprising degree of regularity, despite being typologically diverse (e.g., Fromkin, 1971; Garrett, 1975). This is especially true of a class of errors in which phrases, words, or sounds are unintentionally added, deleted, exchanged, combined, wrongly ordered, or incorrectly selected for inclusion in an utterance. Seminal works involving analyses of errors from corpora of spontaneous speech (e.g., Fay & Cutler, 1977; Fromkin, 1971; Fry, 1969; Garrett, 1975; Nooteboom, 1967; Stemberger, 1985) showed that speakers’ errors tend to deviate from well-formedness in predictable ways. For example, in the case of word substitution errors, in which an unintended substitute word (e.g., restaurant) is produced in place of an intended target word (e.g., rhapsody) in a speaker’s utterance (e.g., Lizst’s second Hungarian (*restaurant vs. rhapsody); Dell, 1995, p. 193), the substitute and the target are usually from the same grammatical class (e.g., nouns substitute other nouns, but
not adjectives; e.g., Fay and Cutler, 1977; Garrett, 1975; Nooteboom, 1967). Substitution 
errors may also be influenced by factors such as similarity in the sound or number of syllables 
between the substitute and target word (e.g., restaurant and rhapsody are both three-syllable 
words with the same initial consonant). Semantic or associative connections between the 
words involved in the error and other words in the utterance (e.g., restaurant and rhapsody 
are both associated with Hungarian; Dell, 1995, p. 194) also affect substitution errors (cf. 
Bock, 1995b; Dell, 1995; Hotoph, 1980).

Many researchers surmise that similar such constraints shape not only the form of some 
speech errors, but also the form of normal sentences (e.g., Bock, 1990; Butterworth, 1980, 
1981; Dell, 1995; Dell, Chang, & Griffin, 1999; Dell & Reich, 1981; Garrett, 1980, 2015; Harley, 
1984; Kempen & Hoenkamp, 1987; Levelt, Roelofs, & Meyer, 1999). Because subject-verb 
agreement errors tend to occur under predictable circumstances, they are well-suited for 
investigations of sentence production, especially the processes that link conceptual information 
to grammatical information during utterance formulation (Bock, 1995a; Bock & Middleton, 
2011; Bock & Miller, 1991; Eberhard, Cutting, & Bock, 2005).

Subject–verb agreement errors

In English, subject-verb agreement traditionally refers to a match between the subject 
(e.g., The book) and the main verb of a sentence (e.g., is) on the binary feature of grammatical 
number. The standard distinction is between singular subjects, which refer to singletons 
and take singular verbs, and plural subjects, which refer to multiples and take plural verbs. 
Agreement is only overtly indicated on a subset of verbs that exhibit inflectional variations 
(cf. The book is interesting vs. The books are interesting), and subject-verb agreement errors 
are evidenced by utterances in which a subject and a verb of this type do not match in 
number (e.g., The book *are interesting). As a linguistic device, agreement allows speakers to 
signal relationships between sentence elements that are connected in thought but that may 
be separated by long distances in speech. As a component of sentence production, agreement
processes are a conduit through which some of the relational features of a speaker’s message are formally encoded.

Several studies of agreement errors have been concerned with distinguishing influences on agreement from factors involving grammatical information versus factors involving meaning information (for an overview, see Bock & Middleton, 2011). Among the factors that have been investigated, grammatical number and notional number have repeatedly been shown to predict differences in the size of mismatch effects.

### 1.1 Grammatical and notional number in agreement

**Grammatical number**

One of the most reliable means of eliciting agreement errors is through the use of noun phrase–prepositional phrase (NP+PP) preambles containing nouns with conflicting grammatical number information. The first empirical demonstration of this was reported in Bock and Miller (1991), who observed a pronounced asymmetry in the distribution of error rates in participants’ preamble completions: In three experiments, participants produced substantially more agreement errors when completing preambles featuring a singular head noun (HN; e.g., *key*) followed by a plural local noun (LN; e.g., *cabinets*), as in (1a), than when completing preambles featuring a plural HN followed by a singular LN, as in (1d).

\[
\begin{align*}
(1) & \text{ a. The key to the cabinets (SP)} \\
& \text{ b. The key to the cabinet (SS)} \\
& \text{ c. The keys to the cabinets (PP)} \\
& \text{ d. The keys to the cabinet (PS)}
\end{align*}
\]

Since Bock and Miller’s (1991) study, the effects of independent variables on agreement are usually quantified in terms of differences in the size of the mismatch effect between conditions, where mismatch effects are calculated as the difference in the error rates for each preamble mismatch condition and its number-matched control (i.e., 1a–1b, and 1d–1c). Because error
rates in the plural HN preamble versions of test stimuli are generally at floor, the experiments in this dissertation follow the custom of using only singular HN preamble versions with a manipulation of grammatical number on the LN (i.e., *singular vs. plural*).

The robustness of the mismatch effect asymmetry indicates that the agreement mechanism is differentially sensitive to conflicts of grammatical number information from the nouns in preambles such as those in (1). One hypothesis attributes this to differences in how number information is marked on nouns or NP constituents at the grammatical-encoding level (e.g., Bock, 2004; Eberhard, 1997; Eberhard et al., 2005). The main idea is that most common nouns are unspecified for number, or else specified as grammatically singular by default. Plurality is therefore the non-default, or *marked*, case. The contrast is important for agreement to the extent that the agreement mechanism uses evidence of plurality in the process of computing the number of the subject NP (e.g., Eberhard et al., 2005). Ordinarily, it derives a plural-agreeing verb when the HN is plural (e.g., in plural–singular preambles like 1d), or a singular-agreeing verb when the HN is unmarked. Occasionally, however, agreement is erroneous computed using the plural information from the LN (e.g., in singular–plural preambles, like 1a) as the number for the subject NP, resulting in the retrieval of a non-agreeing plural verb.

Several studies provide evidence supporting this hypothesis. For example, Eberhard (1997) reported attenuated mismatch effects for preambles containing a singular head NP with a number-marked determiner versus an unmarked determiner (e.g., *ONE key to the cabinets* vs. *THE key to the cabinets*), suggesting that the explicit indicators of singularity on the HN make plural agreement less likely. Bock et al. (2001) found that invariant-plural LNs, which do not have a standard singularized version (e.g., *bubble* vs. *bubbles*; cf. *sud* vs. *suds*), were more likely than regular singular LNs to elicit agreement errors, but considerably less likely to do so than normal plural LNs, which suggests that markedness contrasts are a reliable cue to the agreement mechanism for computing the subject’s number.

The effects of grammatical number variations on agreement appear also to be largely
independent of variations in the morphophonological correlates of number (but see Hartsuiker, Schriefers, Bock, & Kikstra, 2003, for evidence from German and Dutch; see Vigliocco et al., 1995, for evidence in Italian): Bock and Eberhard (1993), for example, reported comparably-sized mismatch effects for preambles containing regular versus irregular grammatically plural LNs (e.g., kids vs. children). Plural-sounding but grammatically singular LNs, on the other hand, were no more likely to induce agreement errors than normal singular nouns (e.g., cruise vs. crew; cf. crews).

**Notional number**

Notional number, which refers to a speaker’s conceptual number valuation for the referent or referents of a to-be-uttered message, also has an influence on agreement production (Bock, Nicol, & Cutting, 1999; Eberhard, 1999; Gillespie, 2011; K. R. Humphreys & Bock, 2005; Vigliocco et al., 1996; Vigliocco et al., 1995). Several studies have examined mismatch effects for a variety of nouns in English whose notional and grammatical number information are incongruent or inconsistent. For example, a (bipartite) noun such as suspender is notionally singular in that speakers conceive of suspenders as a single object, albeit one with two symmetrical parts. For the purposes of agreement, however, suspender is grammatically plural (cf. My suspenders are are too tight vs. *My suspenders is too tight), and bipartite LNs such as this tend to elicit plural agreement errors despite their notional singularity (Bock, Eberhard, & Cutting, 2004). By comparison, collective nouns, such as committee(s), have both a singular sense, which delineates the referent as a unitary entity, and a plural sense, which delineates the individual constituents comprising the referent. As local nouns, most grammatically singular collectives are not more likely than regular singular LNs to increase the size of the mismatch effect (cf. Bock et al., 2006; Bock & Eberhard, 1993; Bock et al., 2001; Bock et al., 1999); however, Bock et al. (2006) found that some plural collective LNs produce unusually large mismatch effects compared to regular plural LNs.

Conflicts between notional and grammatical number can also arise at the level of the
subject NP. For example, some subject NPs encourage a notionally plural interpretation
and reliably elicit more plural than singular verbs in both natural and experimental settings
(Bock et al., 2006; K. R. Humphreys & Bock, 2005). Eberhard (1999) demonstrated this
using distributed-referent preambles, observing larger mismatch effects for preambles in
which a singular HN was construed as being instantiated across multiple tokens of a plural
LN referent (e.g., *The label on the bottles*) compared to single-referent preambles (e.g., *The
key to the cabinets*). Other researchers have replicated this finding, with effect sizes being
modulated by factors such as the conceptual concreteness or imageability of the preamble
(cf. Gillespie, 2011; Vigliocco et al., 1996; Vigliocco et al., 1995; Vigliocco & Hartsuiker,
2002). Although researchers disagree about the exact nature of the contribution that it makes
during agreement computation, there is consensus that notional number information from
a message interacts with other meaning- and grammar-based properties of an utterance to
affect agreement (cf. Berg, 1998; Bock et al., 2001; Bock & Middleton, 2011; Eberhard et al.,

In summary: For number-mismatched preambles in English, grammatically plural LNs,
including those that do not bear the regular formal indicators of plurality, are more likely to
promote agreement errors than singular LNs, including those whose phonological properties
might otherwise ordinarily indicate plurality. And agreement is also influenced by the notional
number information in a speaker’s message. These findings are critically important in that
they establish some of the most common sources of agreement errors (Bock & Eberhard, 1993;
Bock et al., 2001; Bock et al., 1999; Eberhard et al., 2005). They have also formed the basis
for several different theories about the underlying properties of the agreement mechanism
(for an overview, see Bock & Middleton, 2011).

The studies in this dissertation are not directly concerned with providing evidence in favor
of a particular theory of agreement. The experiments join a small but expanding group of
studies that have investigated effects on agreement from factors that do not directly involve
grammatical number or notional number information. Evidence from two separate lines of
research showing effects on agreement from plausibility (Thornton & MacDonald, 2003) and semantic integration (e.g., Gillespie, Pearlmutter, & Shattuck-Hufnagel, 2013; Solomon & Pearlmutter, 2004) have been influential in constraining theories of agreement.

1.2 Non-number related factors in agreement

**Plausibility**

In a set of agreement studies conducted by Thornton and MacDonald (2003), participants made passive sentence constructions from NP+PP preambles that were presented with an accompanying verb. The researchers manipulated the plausibility of different subject–predicate patterns, which affected the rate of agreement errors in participants’ completions: Mismatch effects were larger when both the HN and the LN of the preamble were plausibly predicated by a given verb, as in (2a), where *PRAISED* is a plausible predication for both *album* and *composer*, versus when only the HN of the same preamble was plausibly predicated by a different verb, as in (2b), where *PLAYED* is a plausible predication of *album* only. On the researchers’ interpretation of these findings, speakers’ plausibility valuations constitute one of many possible non-grammatical, non-number-related constraints on agreement production (see also Haskell & MacDonald, 2003; Pollard & Sag, 1988).

(2) a. The album by the classical composer(s) were PRAISED

   b. The album by the classical composer(s) were PLAYED

**Semantic integration**

In a different set of studies, Solomon and Pearlmutter (2004) found consistently larger mismatch effects for preambles that were more versus less semantically integrated (henceforth, *integrated/integration*), which was defined as the degree to which the elements of a to-be-uttered sentence or phrase are linked at the message level. Integration varies as a function of how the HN and LN relate to each other in the specific context denoted by a
preamble, with stronger relationships corresponding to tighter conceptual connections.

For example, in (3a), the preposition of signifies that the chauffeur performs a specific function with respect to the actors, but in (3b), the preposition with most likely signifies an accompaniment relationship. On norming surveys, participants rated preambles like (3a) as more integrated than preambles like (3b), consistent with Solomon and Pearlmutter’s (2004) assumption that the links between referents in a speaker’s to-be-communicated message should be stronger for functional relationships than for accompaniment relationships. Similarly, ratings of integration were higher for (4a), where with designates topping(s) as an attribute of pizza, than for (4b), where the same preposition signals an accompaniment relationship between beverage(s) and pizza.

(3) a. The chauffeur of the actor(s) (Integrated)
   b. The chauffeur with the actor(s) (Unintegrated)

(4) a. The pizza with the yummy topping(s) (Integrated)
   b. The pizza with the tasty beverage(s) (Unintegrated)

Solomon and Pearlmutter (2004) remained neutral as to how agreement might be implemented, but they hypothesized that the effects of integration on agreement are due to changes in the timing at which the elements of the message are encoded. Essentially, the HN and the LN in a more integrated preamble will be planned closer together in time than in a less integrated preamble such that information about the nouns, including information about their grammatical numbers, is more likely to be concurrently active and, therefore, overlappingly accessible to the agreement mechanism (i.e., on the assumptions of incremental-planning accounts of production; e.g., V. Ferreira, 1996). In such a scenario, integration increases the likelihood of interference from the LN during agreement computation, which increases the probability that a speaker will produce an agreement error.

Subsequent experiments supported a timing-based mechanism for the effects of integration on agreement. For example, DiBattista and Pearlmutter (2011) observed an increase in
phrase and word-ordering errors for more versus less integrated visual stimuli in a picture description task, consistent with Solomon and Pearlmutter’s (2007) hypothesis that elements planned closer together in time would be more likely to participate in such errors. There is also evidence for downstream effects of integration on production processes that follow from the predictions of a timing-based mechanism. Gillespie et al. (2013) analyzed word durations and prosody of the recorded utterances of participants from two prior sentence-completion experiments, and they found shorter spoken word durations for the words intervening the preamble nouns as a function of increased integration. This finding suggests that concepts that are planned closer together in time at the message level are uttered with less temporal separation.

The evidence from research on semantic integration underscores three important points: First, sentence formulation is bound on a temporal dimension. Second, the successful translation of a message into an utterance involves a set of specially-timed operations that are vulnerable to disruptions or alterations (see also Dell, 1986; Dell, Burger, & Svec, 1997; Gillespie, 2011; Gillespie & Pearlmutter, 2011; Nicol, 1995). Third, the timing at which information is received, transformed, and output at one level of the system has direct consequences on processing at subsequent levels (e.g., Bock & Warren, 1985; Levelt, 1989). One implication relevant to models of agreement is that properties of a speaker’s message can alter the time-course of production processes in a way that directly affects agreement errors.

### 1.3 Semantic relatedness and association

**Semantic relatedness**

Barker et al. (2001) reported larger mismatch effects for preambles featuring semantically related nouns (henceforth, related/relatedness) versus unrelated nouns (e.g., canoe–sailboats vs. canoe–cabins). The researchers proposed that relatedness increases the size of the mismatch effect by altering the activation profiles of preamble elements within a spreading activation
network. The changes are not related to timing per se, but rather, to the magnitude or strength of activation (cf. Solomon & Pearlmutter, 2004, 2007, for consideration of a similar mechanism for semantic integration; and see Solomon & Pearlmutter, 2007, for a rejection of such a mechanism). Specifically, the relative levels of activation for more versus less related HN-LN pairs are increased because the spread of activation between the nouns is mutually reinforcing. As the activation level of the LN increases, so does the level of activation on its plural feature. This increases the potential for the plural LN to interfere with agreement computation and thus raises the probability of agreement errors.

A hypothetical pattern-activation network is illustrated in Figure 2 for the related nouns canoe and sailboat (A), and separately, for the unrelated nouns canoe and cabin (B). The individual nodes at the message level represent atomic semantic features in accordance with the view that, at some level of abstraction, concepts decompose into such features (cf. Bierwisch & Schreuder, 1992; Dell, 1985; Dell et al., 1999; Levelt et al., 1999). The nodes at the lemma level correspond to the nouns’ semantic-syntactic representations in the functional-processing component of the grammatical-encoding level in Bock and Levelt’s (1994) model (see Figure 1). The network depicted in Figure 2 departs from the model in Figure 1 in allowing for feedback between the message level (= conceptual level, in Figure 2) and the grammatical-encoding level (cf. Dell, 1985, 1986; Dell, Schwartz, Martin, Safran, & Gagnon, 1997; Levelt et al., 1999). Thus, on this model, related nouns are mutually reinforced as activation spreads from the HN to the LN by way of their common semantic feature nodes, and from the LN back to the HN over the same channel(s).

1The semantic-syntactic representations of words at this stage are traditionally referred to as lemmas, hence the label in Figure 2; lemmas are distinguished from lexemes, or the representations of a word’s morphophonological properties (Kempen & Hoenkamp, 1987).
Figure 2. Adaptation of Barker’s (2001) schematic of activation patterns between the semantically related nouns canoe and sailboat (A) and the unrelated nouns canoe and cabin (B), which are represented as nodes in the lemma level. Nodes in the message level (in Barker’s original figure, the conceptual level) represent semantic features. Activation spreads bidirectionally between nodes across the levels. Heavier outlines correspond to greater activation. The plural feature linked to sailboat is more highly activated than the plural feature on cabin.

Relatedness vs. integration

When the conceptual elements of a to-be-uttered message are more integrated or more related versus less integrated or less related, speakers are more likely to produce agreement
errors. Both integration and relatedness involve properties of message-level representations, neither involves a direct manipulation of number meaning, and both conceivably affect agreement computation by changing the nature of information flow through the production system. Although the factors are orthogonal in principle—with integration varying as a function of how two nouns are construed within a specific preamble, and relatedness varying as a function of the meanings of the nouns independent of their context—they are likely to be confounded in practice in that related nouns should be more tightly linked at the message level than unrelated nouns.

The design of three of Solomon and Pearlmutter’s (2004) semantic integration experiments controlled for relatedness by matching the local noun across integrated and unintegrated versions of a preamble; additionally, Penta and Pearlmutter (2015) separately collected relatedness ratings of Solomon and Pearlmutter’s Experiment 4 stimuli, where integration was manipulated by varying the content of the LNs. Analyses of those data indicated that integrated preambles were slightly more related than unintegrated preambles, but the difference was not significant. Thus, the effects of integration on agreement were clearly not dependent on variations in relatedness. Less clear was whether semantic integration contributed to Barker et al.’s (2001) observed effect of relatedness, and more generally, how these factors might interact to affect agreement.

To examine the relationship between integration and relatedness, Penta and Pearlmutter (2015, Experiment 2) fully crossed these factors along with a manipulation of local noun number in 24 test preambles. Norming data confirmed the suspected confound between the variables, and the consequence of trying to control for the confound was a weaker manipulation of integration in the experiment than had been achieved in previous experiments. Standard by-subjects ($F_1$) and by-items ($F_2$; Clark, 1973) ANOVAs on error rates indicated a significant difference in mismatch effects for related preambles compared to unrelated preambles. On subsequent analyses using empirical logit-weighted linear regressions (Barr, 2008), increased relatedness and increased integration both reliably predicted more plural agreement errors.
when the effect of the other factor was held constant.

**Relatedness vs. association**

Penta and Pearlmutter (2015) confirmed an effect of relatedness on agreement errors in their Experiment 1, replicating the results reported by Barker et al. (2001). However, they also noted that the factor was operationalized differently between the respective experiments. Namely, Barker and colleagues described the related HN–LN pairings in their preambles as having “substantial overlap of semantic features” (p. 101), represented by synonyms and taxonomic category coordinates (e.g., *canoe–sailboat*). In contrast, Penta and Pearlmutter used norming surveys to measure relatedness for noun pair candidates, and the set of preambles that were included in the experiment represented many types of relationships (e.g., category coordinates; part-whole relations: *camera–lens*; functional relations: *hammer–nail*). This raised the question of whether different types of semantic relationships would have different potentials to affect agreement, which Penta and Pearlmutter’s Experiment 2 sought to answer.

Although this question had not previously been addressed in the context of agreement research specifically, similar questions have been investigated in production studies involving word-priming effects. One of the enduring issues in that line of research is how to reliably distinguish effects due to relatedness, which in its broadest sense entails a connection between words based on their meanings (e.g., *motor–engine*), from those due to association, which is a measure of the probability of two words co-occurring independent of what the words mean (e.g., *mechanic–engine*).

Although relatedness and association are continuous rather than dichotomous factors, and despite being regularly confounded in practice (Balota, 1994; Ferrand & New, 2003; Hutchison, 2003; Lucas, 2001; Moss, Ostrin, Tyler, & Marslen-Wilson, 1995), priming experiments often maintain the distinction because competing theories of lexical-semantic memory make different predictions about their effects on lexical access (cf. Collins & Loftus, 1975; Fodor, 1983; Forster, 1979; Masson, 1995). Models of agreement do not hinge on such theories, Penta and Pearlmutter (2015), but research into whether either factor is capable of producing priming.
effects in the absence of the other has resulted in a substantial amount of conflicting evidence. Penta and Pearlmutter preserved the distinction between relatedness and association in the design of their second experiment on the assumption that it had the potential to yield the most informative and interesting results.

In Penta and Pearlmutter’s (2015) Experiment 2, singular and plural LN versions of 24 preambles such as those in (5) were used to test for an effect of relatedness on agreement in the absence of association, and a separate group of 24 preambles such as those in (6), also with singular and plural LN versions, were used to test for effects of association on agreement in both the presence and absence of relatedness.

(5) a. The saw by the oily wrench(es) (Category coordinates)  
b. The saw by the oily rag(s) (Control)

For preambles such as those in (5a), the HN–LN pairs were semantically related category coordinates (e.g., saw–wrenches) that were, critically, not associates of each other. For preambles in the unrelated control condition, as in (5b), the same HN was paired with a LN to which it was neither related nor associated (e.g., saw–rags). The category coordinate status of noun pairs in (5) was confirmed through published norms (Van Overschelde, Rawson, & Dunlosky, 2004), and association norms (Nelson, McEvoy, & Schreiber, 2004) were used to ensure the noun pairs in both (5a) and (5b) were completely unassociated. Additionally, a large relatedness difference between conditions was verified through separate norming rating surveys.

(6) a. The jar by the sticky lid(s) (Property relations)  
b. The jar by the sticky cookie(s) (Associates)  
c. The jar by the stale bagel(s) (Control)

For preambles like the ones in (6), the HN was paired with three different LNs, representing two different association conditions and a control condition. In the first association condition, preambles such as (6a) contained a pair of associated nouns that were also related to each
other, with the LN being a property of the HN (e.g., *jar-lids*; henceforth, *property relations*). In the other association condition, preambles such as (6b) were associated but unrelated (e.g., *jar-cookies*; henceforth, *associates*). Such “pure” associates normatively co-occur but otherwise have few semantic features in common. Finally, for control preambles such as (6c), the LN was neither associated nor related to the HN (e.g., *jar-bagels*). The presence of association for nouns in the first two conditions, and the absence of association for nouns in the third condition, was confirmed through published norms (Nelson et al., 2004). Separate rating surveys were used to confirm differences in relatedness between conditions. Property relations were rated as significantly more related than associates, and both were rated as more related than the controls.

The results of Penta and Pearlmutter’s (2015) Experiment 2 showed clear differences in the size of the mismatch effect as a function of relationship type. Both types of associated preambles (e.g., 6a and 6b) produced significantly larger mismatch effects compared to the control preambles (e.g., 6c). In contrast, there was no difference in the size of the mismatch effect between category coordinates (e.g., 5a), which were related but unassociated, and their unrelated controls (e.g., 5b). Thus relatedness had no independent effect on agreement in the absence of association. Also, the presence of relatedness had no additive effect on agreement for associated preambles. Numerically, the largest mismatch effect was produced by the pure associates (e.g., 6b), which was marginally greater than the size of the effect for property relations (e.g., 6a).

The lack of an effect on agreement for the unassociated category coordinates in Penta and Pearlmutter’s (2015) Experiment 2 contrasted with the results reported by Barker et al. (2001). One possible explanation for this is that Barker et al.’s manipulation of relatedness was confounded with association, considering that nouns with a high degree of semantic overlap are often also strongly associated (Balota, Yap, & Cortese, 2006; Hutchison, 2003), or that their manipulation was confounded with integration. Barker et al. (see also Barker, 2001) claimed that the nouns used in their preambles were not associated, but they did not
describe how their stimuli were selected and normed, and Penta and Pearlmutter were not able to access the full set of materials to evaluate the possibility of a confound. Penta and Pearlmutter interpreted their results as evidence that association between preamble nouns is sufficient to increase the size of the mismatch effect.

Experiment 1 follows up on Penta and Pearlmutter’s (2015) work by continuing the exploration of association. It asks whether association can influence mismatch effects when the associative relationship is spread between nouns in different sentences constructions. The research question is informed by predictions involving a priming-based mechanism through which association might influence mismatch effects. A secondary goal of the Experiment is to evaluate the merits of the proposed mechanism. The experiment incorporates elements of classic priming studies in a novel version of the sentence completion task. Singular or plural LN preambles are presented after sentences containing a pair of nouns that are associates of the preamble’s HN, or a different pair of nouns that are associates of the same preamble’s LN. In the control condition, the external nouns and the preambles are unassociated.
2. Effects of Association

Penta and Pearlmutter (2015, Experiment 2) found a reliable increase in the size of the mismatch effect for preambles with associated versus unassociated HN–LN pairs. Outside of these findings, however, there is no direct evidence for effects of association on agreement. By comparison, there is an extensive priming literature documenting manipulations of association to investigate semantic memory and lexical access (cf. Abdel Rahman & Melinger, 2009; Ferrand & New, 2003; Hutchison, 2003; Lucas, 2001). Numerous experiments have demonstrated that on lexical decision tasks (Lupker, 1984; D. E. Meyer, Schvaneveldt, & Ruddy, 1974; Moss et al., 1995; Perea & Gotor, 1997; Seidenberg, Waters, Sanders, & Langer, 1984; Shelton & Martin, 1992) and on word- and picture-naming tasks (Alario, Segui, & Ferrand, 2000; Irwin & Lupker, 1983; Lupker, 1988; Sperber, McCauley, Ragain, & Weil, 1979), participants respond faster to a target word when it is presented after an associated prime stimulus versus an unassociated prime or no prime.

The basic spreading-activation mechanism through which associative priming is thought to affect lexical access (cf. Abdel Rahman & Melinger, 2007, 2009; Collins & Loftus, 1975; Levelt et al., 1999)—namely, by increasing the rate at which information becomes available for processing—is readily incorporated in popular models of sentence production (Bock & Levelt, 1994; Dell, 1986; Dell et al., 1999). An example of such a mechanism is illustrated in Figure 3: (A) depicts part of a lexical-semantic network for the associated words spider and web. Lines between nodes represent links over which activation may spread, with dashed lines designating associative links. The box in (B) symbolizes the event of reading the prime word spider. The dotted arrow connecting (B) to (A) means that reading the prime word causes the relative activation levels of the prime’s lexeme-level representations to increase, which initiates the spread of activation between the linked nodes in the network.
Figure 3. Adaptation of Bock and Levelt’s (1994) sample lexical-semantic network for the prime word *spider* and its associated target word *web*, in (A). Lines between nodes represent links over which activation may spread. Associative links, which are not in Bock and Levelt’s original figure, are shown with dashed lines. The lines imply the possibility of associative links at any given level (Bock & Levelt, 1994; Cutting & Ferreira, 1999; Dell, 1986; Griffin, 2002), but they do not imply that links are necessarily specified at every level. (B) represents the event of reading the prime word *spider*. The dotted arrow connecting (B) to (A) means that reading the prime word causes the relative activation levels of the prime’s lexeme-level representations to increase, which initiates the spread of activation between the linked nodes in the network.

The lexical-semantic network in Figure 3 shows associative relationships between nodes at each of the three levels of the system, which is intended to indicate the possibility of links at any given level. Spreading-activation based accounts of production (e.g., Bock & Levelt,
1994; Cutting & Ferreira, 1999; Dell, 1986; Griffin, 2002) have been fairly noncommittal with respect to where associative connections occur in the system (though this is a highly contentious issue in the literature on lexical access and semantic memory; e.g., Abdel Rahman & Melinger, 2007; Fodor, 1983; Masson, 1995). The predictions for Experiment 1 do not depend critically on any configuration for association, but the priming scenario below makes reference to links at every level for the purposes of illustration.

Thus, a hypothetical priming scenario within a network such as the one shown in Figure 3 starts when the prime word *spider* is read. Reading the prime word activates information about its morphophonological properties at the lexeme level (i.e., its *lexeme*). From here, activation spreads to the prime’s semantic-syntactic representation at the lemma level (its *lemma*), and through association, to the target word’s lexeme. The target’s lemma receives activation from its lexeme, and through association, from the prime’s lemma. Activation from the prime’s lemma spreads to its conceptual-level correlates (its *concept*), while the target’s concept receives activation from its own lemma and from the prime’s concept. On this scenario, information about the target is increased relative to its baseline level. Thus, on tasks in which the primed target should be produced, participants’ responses are speeded because information about the word has been pre-activated, which facilitates encoding processes.

Priming effects are regularly observed on naming tasks when association is manipulated, including when the prime–target pairs are pure associates (cf. Abdel Rahman & Melinger, 2009; La Heij, Heikoop, Akerboom, & Bloem, 2003; Sailor, Brooks, Bruening, Seiger-Gardner, & Guterman, 2009). This raises a question about whether the effects of association on agreement are the downstream consequences of associative priming. Experiment 1 explores this possibility using a sentence completion task designed to prime the HN (e.g., *cathedral*) or, separately, the LN (e.g., *street(s)*) of the same preamble (e.g., *The church near the winding street(s)*) with different nouns to which they are respectively associated (e.g., *cathedral, chapel* for the HN; *avenue, boulevard* for the LN). If the effects of association on agreement are brought about through priming, it should be possible to obtain differences in the pattern of
mismatch effects as a function of the prime condition.

There is evidence of priming effects on the production of full sentences (i.e., as opposed to single words) when associated prime-target stimuli are used. In Bock (1986a), participants described pictured events in which, on test trials, an agent (i.e., the source of an action) was shown performing a transitive action on a patient (i.e., the recipient of the action). Before each trial, the participant heard and then repeated a prime word whose target was either the agent or the patient of the picture. For example, given an image of lightning striking a church, the agent *lightning* would be primed by *thunder* in the agent-prime condition, and the patient *church* would be primed by *worship* in the patient-prime condition. Bock consulted published association norms when creating the prime-target pairings, but as association data were unavailable for many of the test pairs, semantically related, but potentially unassociated, pairs were also used.

The main finding from two experiments was that primed targets were more likely to appear earlier than unprimed targets in participants’ descriptions. Specifically, primed agents were produced more often as subjects of active sentences than as objects of the preposition in passive sentences (e.g., *Lightning struck the church* vs. *The church was struck by lightning*, when *lightning* was primed by *thunder*), and primed patients were more likely to appear as subjects of passive sentences than as objects of active sentences (e.g., *The church was struck by lightning* vs. *Lightning struck the church*, when *church* was primed by *worship*). The results were essentially replicated by Konopka and Meyer (2014) in a separate set of picture-description experiments. In those experiments, prime-target stimuli pairings were all normatively associated.

In addition to establishing that different target nouns from a single to-be-uttered sentence can be primed by an associated noun that is not part of the sentence, these results provide evidence that changes in the time-course of production processes induced through priming can alter the form of an utterance. The mechanism described by Bock (1986a) for the effects of priming on word-order differences in participants’ utterances elaborates on the
mechanism in Figure 3 by characterizing the effects of priming on grammatical-encoding level operations: Priming a target noun makes information about the noun available to the system for processing sooner than would be the case if it had not been primed. Information that is available for processing earlier during sentence formulation is more likely to be produced earlier in an utterance, which makes it more likely to be selected as the sentence subject (e.g., Bock, 1982, 1986a; Bock & Warren, 1985; Branigan, Pickering, & Tanaka, 2008; Levelt & Maassen, 1981). The results of Experiment 1 have the potential to clarify whether, through a comparable mechanism, associative priming can yield differences in the size of the mismatch effect when the HN or the LN of a preamble is primed relative to when neither is primed.

Evidence compatible with this possibility was reported in Gillespie and Pearlmutter (2011, Experiment 1). Participants in that experiment were presented with preambles such as those in (7), which contained two local NPs. Across versions, either the first LN (e.g., buckles) was plural, as in (7b); the second LN (e.g., straps) was plural, as in (7c); or neither was plural, as in (7a). The researchers found an attenuation in the size of the mismatch effect for preambles such as those in (7c) compared to those in (7b), suggesting that the potential for the LN to interfere with agreement processing decreases as linear distance—and temporal separation—between the HN and LN increases. These data support the prediction of a smaller mismatch effect for primed-HN preambles relative to preambles in which the HN is not primed.

(7)  a. The backpack with the plastic buckle on the leather strap
    b. The backpack with the plastic buckles on the leather strap
    c. The backpack with the plastic buckle on the leather straps

Penta and Pearlmutter’s (2015) finding of larger mismatch effects for preambles with associated versus unassociated nouns suggests that it should be possible to obtain a similar result in Experiment 1. The supposition is that a primed- versus an unprimed-LN should be available relatively sooner for planning, which should increase the probability of the HN and the LN being planned closer together in time. Given evidence for increased mismatch effects
when preamble nouns are planned with more temporal overlap (e.g., Gillespie, 2011; Nicol, 1995; Solomon & Pearlmutter, 2004), the prediction on this scenario is that mismatch effects will be largest for primed-LN preambles.

The predicted mismatch effect differences described above reflect the assumption that associative priming directly affects the processing of the targeted preamble noun without necessarily affecting the processing of the non-targeted noun. One alternative to this direct priming account is a scenario in which priming the HN through association changes the rate at which both of the preamble nouns are planned. For example, if encoding a primed-HN is easier than encoding an unprimed-HN, processing of the LN may begin earlier as a result. This would be consistent with evidence suggesting that two nouns from the same NP are more likely to be planned in parallel when the first noun is relatively easy to encode (e.g., when it is a high frequency word; cf. Konopka, 2012; Smith & Wheeldon, 2004; Veenstra, Meyer, & Acheson, 2015). Thus, association may directly influence the planning of the HN by priming it for production, but this may indirectly affect the timing of planning for the LN, increasing the likelihood that both preamble nouns will be planned closer together in time.

On this indirect priming scenario, mismatch effects should still be greater for primed-LN preambles than for primed-HN preambles, assuming that the potential for the LN to interfere with agreement computation will be more likely when the LN is primed than when the HN is primed. In contrast to the direct priming account, however, primed-HN preambles should generate larger mismatch effects than unprimed preambles because interference from the LN is more likely in cases when the HN is primed versus when neither noun is primed. In this case, the size of the mismatch effect for the primed HNs should fall between that of the Neutral preambles and primed-LN preambles.

Experiment 1 uses singular HN–singular LN and singular HN–plural LN preambles, which appear as the start of the third sentence in a three-sentence vignette. The opening sentence of the vignette is the same across versions. Association is manipulated by varying the content of a pair of nouns in the second sentence such that they are either both associates of the
preamble HN, or of the LN, or of neither the HN nor the LN.

Predictions

Following numerous other agreement experiments (e.g., Bock & Miller, 1991; Penta & Pearlmutter, 2015), the manipulation of local noun number is expected to generate more agreement errors after plural-LN than after singular-LN preambles, which would be evidenced by a main effect of this factor. A significant interaction between local noun number and prime type would indicate an effect of the associative priming manipulation on mismatch effects. The predictions on the direct priming account are decreased mismatch effects when the HN is primed, and increased mismatch effects when the LN is primed. The indirect priming account predicts that mismatch effects will be largest when the LN is primed, smaller when the HN is primed, and smallest when neither of the nouns is primed. Reliable differences in mismatch effects would be determined on the basis of an interaction between local noun number and prime type in the separate analyses comparing the primed-HN preambles to the Neutral preambles, the primed-HN preambles to the primed-LN preambles, and finally, the primed-LN preambles to the Neutral preambles.

2.1 Experiment 1

Method

Participants

Three hundred eleven Northeastern University undergraduates participated in the experiment for course credit. Excluded from all analyses were data from 15 participants who were non-native English speakers, from six participants who contributed too few usable trials, and from one participant who was unable to complete the experiment. Data from six participants were lost due to equipment failure, six participants’ data were not properly recorded, and one participant’s data were lost due to experimenter error. Analyses were conducted on data
from the remaining 276 participants.

**Materials and design**

Twenty-four 3-sentence vignettes were constructed as test stimuli. The preamble was always the start of the third sentence (e.g., as in 10) and its head and local noun were not associates of each other (based on the Nelson et al., 2004, association norms). Preambles like that in (10), which featured a standard local noun-number manipulation, were presented after an opening sentence (e.g., as in 8) and a prime sentence (e.g., as in 9a-c). The three different prime sentences were created by varying whether two grammatically singular nouns therein were normatively associated to the head noun of the preamble (e.g., HN; as in 9a), the local noun (e.g., LN; as in 9b), or neither noun (e.g., Neutral; as in 9c).

The prime sentences differed minimally across the three versions of each vignette, and most differed only in the noun primes. ANOVAs on mean (arcsine-transformed) proportions of association (Nelson et al., 2004) confirmed no difference in association as a function of the prime condition (e.g., HN vs. LN), the local noun number condition (e.g., singular vs. plural), the prime’s serial sentence position (e.g., first vs. second-occurring), nor as a function of their various interactions. Between conditions, the prime sentences were closely matched for length in characters, syllables, and phonemes. The preambles within a vignette differed only in the local noun’s number.

(8) The old town is a perfect model for postcards. (Opening)

(9) a. The white cathedral and the grand chapel are always decorated. (HN)
   b. The wide avenue and the grand boulevard are always decorated. (LN)
   c. The white fortress and the grand castle are always decorated. (Neutral)

(10) The church near the winding street(s) (Preamble)

To safeguard against the possibility of priming subject-verb agreement patterns (e.g., Haskell, Thornton, & MacDonald, 2010), the tense and number of prime sentence verbs were balanced across vignettes: Present tense verbs (e.g., *are*, as in 9) were used in the prime
sentences in 12 of the vignettes. Of these, six featured plural-agreeing verbs (e.g., *are*), and six featured singular-agreeing verbs (e.g., *is*). Past tense verbs (e.g., *were*) were used in the remaining 12 vignettes. Of these, four featured plural-agreeing verbs, four featured singular-agreeing verbs, and four were uninflected for number (e.g., *looked*). Opening sentences were identical across versions while varying across vignettes.

The critical stimuli were combined with 48 three-sentence and 8 four-sentence filler vignettes, the latter of which were included to discourage participants from attending more carefully to the third sentence than to the preceding sentences on each trial. A preamble was the last sentence in every filler vignette. The opening and middle sentence(s) of all filler vignettes varied considerably in content and structure. The preamble fragments in 16 of the filler vignettes were plural head NP+PP preambles that were similar in structure to the preambles in the critical stimuli. Eight of these preambles appeared with plural local NPs, and eight appeared with singular local NPs. Two preambles from each version were used in the four-sentence vignettes, and the remaining were used in the three-sentence vignettes.

Preambles in the remaining filler vignettes contained a variety of structures, none of which were NP+PP structures, and approximately one-quarter of which did not consist of a subject NP (e.g., *Now that he’s in high school*). The critical item vignettes and the filler vignettes were combined to create six counterbalanced lists. Each list contained all of the filler vignettes and exactly one version of each critical item vignette. Each list was seen by 45–47 participants.

**Apparatus and procedure**

All participants were run individually. Participants were instructed to read each sentence aloud and, for preambles, to create a sentence completion as quickly and as naturally as possible. Participants received no additional instructions and were provided feedback only if the rate of their speech slowed considerably at any time during the experiment, in which case the experimenter reminded the participant to increase the speed of his or her response. The experiment consisted of a total of 80 trials preceded by 5 practice trials.
On individual trials with critical item vignettes, a fixation cross appeared on the left edge of the computer screen for 1000 ms. This was followed by the opening sentence, presented in white text, which remained until the participant pressed the spacebar to advance to the prime sentence (or to the continuation sentence in non-critical item trials, as described below). The prime sentence (or continuation sentence) also appeared in white text and was displayed until the participant pressed the spacebar to advance to the vignette-terminal preamble. The preamble was presented in yellow text approximately one inch below where the opening and prime sentence appeared, and it remained on screen for the larger of 1000 ms or 40 ms per character. After each preamble was displayed, a blank screen appeared for 3000 ms, followed by a message that instructed participants to press the space bar to begin the next trial.

Individual trials for the 48 three-sentence filler vignettes were identical in structure to the critical item trials. The opening sentence appeared first, followed by the continuation sentence, which was followed finally by the preamble. Trials for the 8 four-sentence filler vignettes differed from the others only in that they included a second continuation sentence, which replaced the first continuation sentence when participants pressed the space bar. The second continuation sentences were presented in white text in the same location as the first continuations, and they also remained on the screen until the participant pressed the spacebar to advance to the preamble.

The apparatus was identical for Experiments 1–3. The stimuli were presented on IBM-compatible computers running the MicroExperimental Laboratory (MEL) software package (Schneider, 1988). Participants’ responses were recorded onto CD-R for analysis, using a professional-quality microphone connected to a Mackie 1202-VLZ Pro mixer/preamp and an Alesis Masterlink ML9600 CD Recorder.

**Scoring**

Scoring was identical for Experiments 1–3. All participants’ preamble completions for critical item trials were transcribed and assigned to one of four scoring categories, except in cases where a participant made no response. Correct responses were those in which a
participant properly repeated the entire preamble, said the preamble only once, and provided an inflected verb that was correctly marked for number as the first word of the completion. Responses were scored as errors if all of the above conditions were met except that the verb was incorrectly marked for number. Responses were scored as uninflected if all of the criteria for correct responses were met, but the verb was not overtly marked for number. Finally, responses were scored as miscellaneous for trials in which a participant failed to correctly repeat the preamble, a verb did not follow immediately after the preamble, or the criteria for any other category were not met. The no-response cases were noted and are excluded from all analyses.

Trials on which a participant’s response included a dysfluency during or immediately after the preamble were recorded as such; in cases where a participant produced a dysfluency but proceeded to complete a sentence in accordance with one of the first three scoring categories, both the dysfluency and the scoring category were noted. Dysfluencies in miscellaneous responses were not counted separately.

**Results**

Across 6,624 trials, there were 4,216 correctly inflected responses, 313 agreement errors, 1,036 uninflected responses, 1,044 miscellaneous cases, and 15 trials with no response. Table 1 shows agreement error rates and counts of each response type by prime condition and local noun number for the critical item preambles.

All of the analyses described for Experiments 1–2 were performed both on raw (untransformed) and on arcsine-transformed error rates (the proportion of error responses out of error plus correct responses), calculated both with and without the dysfluent cases. Except where noted, the statistical patterns were identical for the analyses on both raw and transformed error proportions, and when dysfluency cases were excluded. Therefore, the statistics reported for Experiments 1–2 are for analyses on untransformed error proportions with dysfluency cases included.

Experiment 1 analyses were initially conducted after data from 185 participants had
been collected. Subsequently, however, it was discovered that, for both singular and plural versions of three of the critical items, the HN prime sentences had been misclassified as LN prime sentences, and vice-versa. As a result, the experimental lists were not properly counterbalanced, and data from the remaining subjects were collected.

To evaluate the extent to which counterbalancing may have influenced the results of Experiment 1, the data were submitted to ANOVAs with local noun number (singular vs. plural) and prime type (HN vs. LN vs. Neutral) included as within-subjects factors, counterbalancing included as a two-level between subjects factor, and participants ($F_1$), or separately, items ($F_2$; Clark, 1973), as the random factor.

The ANOVAs indicated that there was no main effect of counterbalancing ($p > .10$); no interaction between counterbalancing and local noun number ($p > .10$); no interaction between counterbalancing and prime type ($F_s < 1$); and no three-way interaction ($F_s < 1$). Consequently, it is assumed that the counterbalancing issue is not likely to have directly influenced the outcome of Experiment 1. The results reported below are therefore from analyses conducted on the full data set with prime type coded correctly based on the stimuli actually presented to the subjects.

The agreement error data were analyzed with a series of 2 (LN number: singular vs. plural) × 3 (Prime: HN vs. LN vs. Neutral) ANOVAs with participants ($F_1$) and with items ($F_2$; Clark, 1973) as the random factor. Figure 4 shows the pattern of mismatch effects.

Participants produced more agreement errors following plural LN preambles compared to singular LN preambles ($F_1(1,275) = 90.40, MS_e = 474.92, p < .001; F_2(1,23) = 33.80, MS_e = 106.59, p < .001$), confirming the expected main effect of local noun number. There was a significant main effect of prime type ($F_1(2,550) = 7.92, MS_e = 203.35, p < .001; F_2(2,46) = 6.58, MS_e = 25.93, p < .01$). Participants produced fewer errors overall when the HN was primed compared to when the LN was primed ($F_1(1,275) = 13.01, MS_e = 229.76, p < .001; F_2(1,23) = 8.63, MS_e = 32.35, p < .01$), and fewer errors overall when the HN was
Table 1

Experiment 1 Error Rates and Response Counts by Condition

<table>
<thead>
<tr>
<th>Prime</th>
<th>Noun Number</th>
<th>Error Rate</th>
<th>Error</th>
<th>Correct</th>
<th>Uninflected</th>
<th>Misc</th>
<th>No Resp</th>
</tr>
</thead>
<tbody>
<tr>
<td>HN SS</td>
<td>0.92 (0.35, 0.45)</td>
<td>7 (1)</td>
<td>757 (83)</td>
<td>172 (21)</td>
<td>162</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>8.89 (1.21, 1.42)</td>
<td>64 (11)</td>
<td>656 (74)</td>
<td>181 (27)</td>
<td>204</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LN SS</td>
<td>1.97 (0.60, 0.90)</td>
<td>15 (1)</td>
<td>746 (80)</td>
<td>183 (27)</td>
<td>158</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>14.48 (1.55, 1.97)</td>
<td>109 (8)</td>
<td>644 (61)</td>
<td>155 (17)</td>
<td>192</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Neutral SS</td>
<td>2.04 (0.48, 1.16)</td>
<td>16 (2)</td>
<td>770 (75)</td>
<td>187 (24)</td>
<td>130</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>13.69 (1.46, 1.74)</td>
<td>102 (8)</td>
<td>643 (55)</td>
<td>158 (15)</td>
<td>198</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Total | 313 (31) | 4216 (428) | 1036 (131) | 1044 | 15 |

*Note.* Error rates are reported as percentage errors out of errors plus correct responses. Standard errors of the mean computed from the analyses by participants and by items are in parentheses. Response counts include dysfluency counts, in parentheses. HN = Head noun; LN = Local noun; SS = singular head, singular local noun; SP = singular head, plural local noun; Misc = Miscellaneous; No Resp = No response.

Figure 4. Mismatch effects by prime condition. Error bars indicate ± 1 SEM, computed from the analyses by participants.
primed compared when neither noun was primed ($F_1(1,275) = 10.07, MS_e = 163.18, p < .01$; $F_2(1,23) = 10.00, MS_e = 23.06, p < .01$).

The critical interaction between prime and local noun approached significance in the main analyses (all $ps < .07$), and it was significant in the analyses with dysfluencies excluded ($F_1(2,550) = 3.75, MS_e = 257.32, p < .05$; $F_2(2,46) = 5.07, MS_e = 22.37, p < .05$). When dysfluencies were included, the interaction was significant on the by-items analysis of untransformed error rates (all $ps < .05$), and otherwise.

Differences in mismatch effects between the prime conditions were tested using separate 2 $\times$ 2 ANOVAs, all of which included LN number (singular vs. plural) as a fixed effect, and all with participants ($F_1$) and with items ($F_2$; Clark, 1973) as the random factor. Thus, the other fixed effect term was either specified for the comparison of primed-HN preambles to the Neutral preambles, for the comparison of primed-LN preambles to Neutral preambles, or for the comparison of primed-HN preambles to Neutral preambles.

On these analyses, mismatch effects were numerically smaller for primed-HN preambles ($F_1(1,275) = 40.69, MS_e = 211.26, p < .001$; $F_2(1,23) = 15.56, MS_e = 40.97, p < .001$) compared to the Neutral controls ($F_1(1,275) = 46.70, MS_e = 331.90, p < .001$; $F_2(1,23) = 23.42, MS_e = 57.28, p < .001$). The difference was reliable only on the by-items analyses when dysfluency cases were included ($F_1(1,275) = 2.33, MS_e = 217.11, p = .128$; $F_2(1,23) = 4.84, MS_e = 13.36, p < .05$), and significant in both the by-subject and by-items analyses when dysfluencies were excluded (all $ps < .03$).

There was a significant difference in the size of the mismatch effect in the analyses comparing the HN and LN prime conditions ($F_1(1,275) = 5.57, MS_e = 215.06, p < .05$; $F_2(1,23) = 7.71, MS_e = 18.40, p < .05$). Participants produced larger mismatch effects in their completions to primed-LN preambles ($F_1(1,275) = 53.90, MS_e = 372.32, p < .001$; $F_2(1,23) = 36.99, MS_e = 47.89, p < .001$) than in their completions to primed-HN preambles.
Mismatch effects did not differ reliably between the LN prime preambles and the Neutral controls (all $ps > .36$).

Uninflected responses (with and without dysfluencies) were analyzed with a set of 2 (LN number: singular vs. plural) $\times$ 3 (Prime: HN vs. LN vs. Neutral) ANOVAs with participants ($F_1$) and with items ($F_2$; Clark, 1973) as the random factor. These analyses indicated no interaction between local noun number and prime type (all $ps > .19$). An analogous set of ANOVAs were used to analyze counts of miscellaneous responses, which also indicated no interaction between local noun number and prime type (all $ps > .31$).

Discussion

Two findings emerged from Experiment 1. The first was the replication of the singular–plural mismatch effect asymmetry reported in Bock and Miller (1991) and in numerous other agreement studies (e.g., Eberhard et al., 2005; Solomon & Pearlmutter, 2004). The second was that association between the prime nouns and their preamble target nouns affected the rate of agreement errors in participants’ sentence completions. Numerically, participants produced the fewest agreement errors when the primes were associates of the HN, and the most agreement errors when the primes were associates of the LN.

More critically, mismatch effects were reliably smaller for primed-HN versus primed-LN preambles, and they tended to be smaller for primed-HN preambles in comparison to the Neutral condition. This suggests that the associative relationship to the HN decreased the probability of agreement errors. These results extend Penta and Pearlmutter’s (2015) finding of increased mismatch effects for associated versus unassociated preamble nouns by showing that association can influence mismatch effects even when the relationship is spread across two sentences. They also go beyond Penta and Pearlmutter by showing that association affects agreement by reducing the mismatch effect when the HN, rather than the LN, is part of an associative relationship.
On the direct priming scenario described in the Introduction, association between the prime nouns and HN was predicted to decrease the size of the mismatch effect relative to cases in which the HN was not primed. The significant difference in mismatch effects between the primed-HN and primed-LN preambles, and the trend towards a significant difference between the primed-HN and Neutral preambles are consistent with this prediction. The results are inconsistent with the prediction on the indirect priming scenario that primed-HN preambles would increase the size of the mismatch effect relative to the unprimed control preambles. The assumptions underlying this prediction followed from evidence indicating that two nouns in a conjoined-noun phrase are more likely to be planned in parallel when the first noun is easy to process (Konopka, 2012; Smith & Wheeldon, 2004; Veenstra et al., 2015), and also that mismatch effects tend to be larger when information about both preamble nouns is overlappingly active during agreement computation (e.g., Gillespie, 2011; Nicol, 1995; Solomon & Pearlmutter, 2004). In Experiment 1, mismatch effects were smallest for primed-HN preambles, suggesting that association can prime the production of the HN without necessarily affecting the timing of planning of the LN.

One interpretation of these results is that reading the prime sentence activates information about the prime nouns (e.g., cathedral; chapel) within the lexical-conceptual network. Some information about the HN (e.g., church), which is associatively linked to the primes, will receive activation as well. Because a production plan for the preamble can be initiated once a sufficient amount of information about the HN is available—irrespective of whether any information about the LN is available (see Allum & Wheeldon, 2007, 2009; Konopka & Meyer, 2014; Wheeldon, Ohlson, Ashby, & Gator, 2013, for similar proposals concerning planning complex noun phrases)—a primed versus an unprimed HN is more likely be processed with greater temporal separation from the LN. One consequence is that information about the LN, including number information, is less likely to be active at the time that agreement is being
computed. As a result, there is a smaller chance that the LN will interfere with agreement when the HN is primed versus when it is not primed, which predicts smaller mismatch effects for these preambles.

The proposed priming mechanism for association’s influence on mismatch effects also predicted that association between the prime nouns and the LN would result in larger mismatch effects in comparison to the other conditions. Primed-LN preambles did yield the largest mismatch effects, but these were not reliably different from the unprimed control preambles. One possible reason for the generally weak effects of associative priming on mismatch effects observed in Experiment 1 is that the time interval between the presentation of the primes and their targets (i.e., the stimulus onset asynchrony; SOA) was too large. On many spreading activation models of lexical access (cf. Abdel Rahman & Melinger, 2009; Janssen, Schirm, Mahon, & Caramazza, 2008), the activation level of a primed element is assumed to decay rapidly, returning to baseline after two to three seconds (Chang, Dell, & Bock, 2006; Levelt et al., 1999). This is accounted for in the design of many traditional priming studies, where SOA is manipulated as an independent variable (e.g., Abdel Rahman & Melinger, 2009; La Heij, Dirkx, & Kramer, 1990; Lupker, 1988; Xavier-Alario, Segui, & Ferrand, 2000).

In Experiment 1, the amount of material separating the primes and their respective targets was not controlled. HNs were always separated from the second of the two primes by a minimum of one word (i.e., The), while LNs were always separated by a minimum of five words: the HN’s determiner, the HN, a preposition, the LN’s determiner, and the LN’s adjective. The prime sentences terminated with a prime noun in nine of the vignettes, but in the 15 remaining vignettes, the second prime noun was followed most frequently by 2 or 3 words (\( M = 2.7; \) min. = 1; max. = 5). The additional discourse material that followed most primes in Experiment 1 could therefore have resulted in lengthy SOAs, which may have
reduced the effectiveness of associative priming to influence agreement.

Experiment 2 explores this possibility by modifying the materials and procedure of Experiment 1. Only the prime nouns and preambles from each vignette are presented on critical trials, and the number of intervening words between the primes and preamble nouns is minimized. Also, whereas in Experiment 1, participants controlled the timing of the progression from the prime sentences to the target preambles, the rate of presentation in Experiment 2 is controlled automatically by the computer. If long SOAs resulted in preamble nouns being only weakly primed in Experiment 1, presenting the preambles with minimal separation from the primes in Experiment 2 should increase the relative strength of the priming effect, and potentially clarify the pattern of mismatch effects.

2.2 Experiment 2

Method

Participants

One hundred four Northeastern University undergraduates participated in the experiment for course credit. Excluded from all analyses were data from 8 participants who were non-native English speakers, and from 3 participants who contributed too few usable trials. Data from 3 participants were lost due to recording-equipment failure. Data from the remaining 90 were included in all analyses.

Materials

The preambles and prime sentences from the 24 vignettes from Experiment 1 were the source for the materials in Experiment 2. The six versions of each Experiment 2 stimulus item were created by pairing the singular and plural LN versions of each preamble (e.g., as in 12) with only the two (bare) nouns from each prime condition sentence (e.g., as in 11a for
the HN prime condition; 11b for the LN prime condition; and 11c for the Neutral condition).

(11) a. CATHEDRAL–CHAPEL (HN)
b. AVENUE–BOULEVARD (LN)
c. FORTRESS–CASTLE (Neutral)

(12) The church near the winding street(s) (Preamble)

Preambles from the 16 plural head filler items from Experiment 1 were used as plural head filler preambles in Experiment 2. The eight plural-head fillers with singular LNs and the eight plural-head fillers with plural LNs were each randomly assigned to one of two groups. The fillers in one group were paired with two “prime” nouns that were unassociated to either of the nouns in the preamble (Nelson et al., 2004). In the other group, the items were paired with two nouns that were both associates of the HN and/or the LN.

The preambles from the 40 miscellaneous filler items from Experiment 1 were used as miscellaneous filler preambles in Experiment 2. Of these, 12 items were selected at random and paired with two nouns that were both associates of each other and also of one of the content words in the preamble (Nelson et al., 2004). The experimenter selected each content word arbitrarily; if two normatively associated nouns could not be identified for the selected word, another word was chosen. For each of these preambles, a second version was created in which only one noun of the prime pair was pluralized. Another 12 items were selected at random and paired with two nouns that were associates of each other but that were not associated with any of the content words in the preamble. Plural versions were created for all but one of these items (due to experimenter error, one singular noun was used in both versions) by pluralizing one of the nouns. The remaining 16 items were paired with two nouns that were neither associated with each other nor associated with any of the content words in the preamble. A second plural version of these items was also created by pluralizing one of the nouns.
An additional 96 pairs of nouns were created as fillers and were not paired with preambles. In half of these no-preamble fillers, the nouns in the pair were associates of each other, and in the other half, they were not associated (Nelson et al., 2004). A second version of each pair was made by pluralizing only one noun in the pair.

The critical items and all filler items were combined to create 24 counterbalanced experimental lists. Each list contained 176 items: one of the six versions of each of the 24 critical items, all of the 16 plural-head fillers, one version (singular or plural) of each of the 40 miscellaneous fillers, and one version (singular or plural) of each of the 96 no-preamble fillers.

Procedure

All participants were run individually. Participants received instructions to read each word aloud as soon as it appeared on the screen and, for preambles, to create a completion as quickly and as naturally as possible. From the participants’ perspective, the stimuli were a list of words with sentence fragments randomly interspersed. Participants received no additional instructions, and were provided feedback only if the rate of their speech slowed considerably at any time during the experiment, in which case the experimenter reminded the participant to increase the speed of his or her response. The experiment consisted of 80 trials preceded by 6 practice trials.

An individual trial consisted of a space-bar press, followed by a fixation cross, followed by any number of prime pairs (mode = 3 pairs, or 6 words) presented one word at a time for 500 ms, followed by either a critical item preamble or a filler preamble, each of which was presented immediately after its own prime pair, for the larger of 1000 ms or 40 ms per character. After each preamble was displayed, a blank screen appeared for 3000 ms, followed by a message that instructed participants to press the space bar to begin the next trial.

Results
Across 2,152 trials, there were 1,383 correctly inflected responses, 22 agreement errors, 395 uninflected responses, 317 miscellaneous cases, and 35 trials with no response. Table 2 shows agreement error rates and counts of each response type by prime condition and local noun number for critical item trials. The data were submitted to a series of 2 (LN number: singular vs. plural) × 3 (Prime: HN vs. LN vs. Neutral) ANOVAs on (arcsine- and untransformed) agreement error rates, with participants ($F_1$) and with items ($F_2$; Clark, 1973) as the random factor. Figure 5 shows the pattern of mismatch effects.

Mismatch effects were virtually identical across the prime conditions. There was the expected main effect of local noun number ($F_1(1,89) = 9.47, MS_e = 112.49, p < .01; F_2(1,23) = 17.54, MS_e = 20.86, p < .001$). Participants produced more agreement errors following plural LN preambles compared to singular LN preambles. There was no effect of prime type ($F_1(2,178) = 0.28; F_2(2,46) = 0.32$), and no interaction between prime type and local noun number. There were no main effects of any factor, nor any interactions between the factors, in the analyses of uninflected responses (all $p$s > .21) and miscellaneous responses ($p$s > .12).

**Discussion**

The results of Experiment 2 were quite clear. The significant head–local mismatch effect reflected that participants only made agreement errors when completing plural LN preambles. Mismatch effects did not differ reliably in any of the analyses and were numerically very close in size. This suggests that association did not influence mismatch effects when the prime nouns were presented (serially) in a list immediately before the preambles.

Experiment 2 was motivated in part to investigate the possibility that decreasing the prime–target SOAs would strengthen the associative priming effects on mismatch effects observed in Experiment 1. To the extent that the relatively short SOAs used in Experiment 2 did not correspond to differences in mismatch effects as a function of the association manipulation, this possibility is not supported by the results of Experiment 2.
Table 2

Experiment 2 Error Rates and Response Counts by Condition

<table>
<thead>
<tr>
<th>Prime</th>
<th>Noun</th>
<th>Error Rate</th>
<th>Error</th>
<th>Correct</th>
<th>Uninflected</th>
<th>Misc</th>
<th>No Resp</th>
</tr>
</thead>
<tbody>
<tr>
<td>HN</td>
<td>SS</td>
<td>0.00 (0.00, 0.00)</td>
<td>0 (0)</td>
<td>215 (24)</td>
<td>74 (6)</td>
<td>53</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>2.94 (1.31, 1.58)</td>
<td>7 (0)</td>
<td>231 (22)</td>
<td>66 (6)</td>
<td>48</td>
<td>9</td>
</tr>
<tr>
<td>LN</td>
<td>SS</td>
<td>0.00 (0.00, 0.00)</td>
<td>0 (0)</td>
<td>235 (21)</td>
<td>70 (7)</td>
<td>42</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>3.21 (1.40, 1.08)</td>
<td>8 (0)</td>
<td>241 (22)</td>
<td>67 (6)</td>
<td>39</td>
<td>3</td>
</tr>
<tr>
<td>Neutral</td>
<td>SS</td>
<td>0.00 (0.00, 0.00)</td>
<td>0 (0)</td>
<td>235 (20)</td>
<td>64 (12)</td>
<td>43</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>3.00 (0.97, 0.88)</td>
<td>7 (1)</td>
<td>226 (24)</td>
<td>54 (8)</td>
<td>58</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>22 (1)</td>
<td>1383 (133)</td>
<td>395 (45)</td>
<td>283</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

Note. Error rates are reported as percentage errors out of errors plus correct responses for critical items. Standard errors of the mean computed from the analyses by participants and by items are in parentheses. Response counts include dysfluency counts, in parentheses. HN = Head noun; LN = Local noun; SS = singular head, singular local noun; SP = singular head, plural local noun; Misc = Miscellaneous; No Resp = No response.

Figure 5. Experiment 2 mismatch effects by prime condition. Error bars indicate ± 1 SEM, computed from the analyses by participants.
One possible explanation for why associatively primed-LN preambles showed little evidence of influencing mismatch effects in Experiments 1–2 is that, even with the reduced SOAs in Experiment 2, SOAs were generally too long. This follows from the fact that the (second) prime nouns and LNs were always separated from each other by five words (i.e., the length of the preamble up to the point of the LN). One interpretation on this explanation is that priming effects on agreement for primed-LN preambles were unlikely to have been obtained in either experiment because the activation levels of the LN decay while the preamble initial material is being planned and produced.

An alternative possibility follows from the fact that HNs are always produced in advance of LNs (i.e., when they are produced correctly) for NP+PP preambles. In order for the production system to produce the HN, a highly active (primed) LN might need to be suppressed or inhibited (e.g., Dell & O’Seaghdha, 1992; MacKay, 1987; Shattuck-Hufnagel, 1979, but cf. Caramazza, 1997; Dell, 1985; Harley, 1999; Levelt et al., 1999; Stemberger, 1985; Tree, Hirsh, & Monsell, 2005). Similar claims have been made in the priming literature, for example, to account for findings (Joordens & Besner, 1992) showing that producing one or two nouns between the presentation of a prime noun and target nouns has an additive effect on naming latencies beyond what is predicted by the increased SOAs from the “filler” nouns (Damian & Als, 2005; Joordens & Besner, 1992, cf.). This is argued to be because having to produce the non-prime nouns first forces the system to suppress the primed-targets’ activation levels.

Primed-HN preambles generated the smallest mismatch effects in Experiment 1, but in Experiment 2, they generated mismatch effects of about the same size as those of the primed-LN and the Neutral conditions. This result was surprising, and it suggests the need for a reevaluation of the SOA question. La Heij et al. (1990) observed that associative priming effects (on picture naming tasks) are most reliably obtained at negative SOAs between 2000–450 ms (and possibly smaller negative SOAs for word-naming tasks; see Perea & Gotor,
1997), but that associative priming becomes less effective more rapidly as negative SOAs approach 200 ms. Systematically varying SOAs as a factor in a design similar to the one used in Experiment 2 could help to clarify whether the priming-based mechanism for association’s effects on agreement is suitable.

In summary, the main goal of Experiments 1 and 2 was to explore association’s potential to influence mismatch effects when the associates are not part of the same preamble. The second goal was to evaluate the merits of a priming-based mechanism through which association might produce its effects on agreement. The results of Experiment 1 suggest that it is possible to modulate the rate of agreement errors with association, even when the associative relationship spans two sentences. Supporting the predictions of the direct priming mechanism, mismatch effects were significantly smaller for primed-HN preambles versus primed-LN preambles. However, the results of both experiments suggest that associative priming does not predict differences in mismatch effects at shorter SOAs.

While there are many interesting open question about the influence of association on mismatch effects, Experiment 3 turns its attention to another factor: semantic relatedness.
3. Effects of Semantic Relatedness

Animacy is argued to be one of the primary semantic dimensions along which humans’ conceptual knowledge is structured (cf. Caramazza & Shelton, 1998; Carey, 2009; Keil, 1981; Opfer & Gelman, 2010). Animacy-based distinctions are also ubiquitous in human languages (e.g., Comrie, 1989; Zaenen et al., 2004), serving as one of the “organizing principles for grammar and discourse” (Dahl, 2008, p.105). Several production studies have observed animacy-related biases in speakers’ utterances, including a preference for making animate entities subjects of active sentences (e.g., F. Ferreira, 1994; Gennari, Mirković, & MacDonald, 2012; McDonald, Bock, & Kelly, 1993), and a tendency to produce animate nouns earlier in sentences than inanimate nouns (e.g., Bock, 1986b).

There is also evidence that animacy relatedness between preamble nouns increases the probability of agreement errors in participants’ sentence completions: Barker et al. (2001) crossed LN number (singular vs. plural), HN animacy (animate (A) vs. inanimate (I)), and LN animacy (A vs. I) in standard NP+PP preambles such as those in (13). Their main finding was that mismatch effects were larger for preambles with animacy-matched nouns (AA, II) versus animacy mismatched nouns (AI, IA). The inanimate-matched (II) preambles generated larger mismatch effects than the animate-matched preambles (AA), and mismatch effects were smallest for animate-HN–inanimate-LN (AI) preambles. Barker et al. (2001) proposed that animacy makes (animate) HNs less prone to interference from LNs (cf. Bock & Miller, 1991).
Barker et al.’s (2001) interpretation of these results involved a mechanism analogous to the one illustrated in Figure 2 for the effects of semantic relatedness on agreement: In a spreading-activation network, lexical–conceptual representations of animacy-matched nouns are connected through their common links to semantic animacy features. When information about the HN is activated, the activation spreads over these links to the LN, which spreads activation back to the HN in a mutually reinforcing manner. One consequence of this is that the activation level of a (plural) LN’s number feature is also reinforced. In some network models of agreement (e.g., Eberhard, 1997), active plural information from a source other than the HN can disrupt normal agreement and lead to plural agreement errors. Thus, the mechanism predicts larger mismatch effects for animacy-matched versus -mismatched nouns, and the mismatch effects in Barker et al. (2001) were consistent with that prediction.

As the network model depicted in Figure 2 implies, conceptually related nouns tend to have more features in common than nouns that are less related (cf. Abdel Rahman & Melinger, 2009; Bock & Levelt, 1994; Collins & Loftus, 1975; Levelt et al., 1999). Barker et al. (2001) predicted that, on average, animacy-matched nouns would be more related than animacy-mismatched nouns. Consistent with their predictions, Barker et al. found that animacy-match reliably correlated with higher rated HN–LN synonymy on their analyses of data from a post hoc survey of their test preambles (see also Gennari et al., 2012; G. W. Humphreys, Mirković, & Gennari, 2016, for similar evidence of a confound). This confound suggests that animacy relatedness is a good predictor of the overall relatedness of nouns. The results of their Experiment 1 suggest that when these factors are varied together across

(13) a. The girl behind the teacher(s)  (AA)
b. The girl behind the desk(s)       (AI)
c. The blackboard behind the desk(s) (II)
d. The blackboard behind the teachers(s) (IA)
preamble versions, they affect the probability of agreement errors in participants’ sentence completions.

In a separate study, Barker et al. (2001) reported larger mismatch effects for preambles with HN–LN pairs that were matched in animacy but that differed in their amount of relatedness (i.e., based on rated synonymy). On Barker and colleagues’ interpretation of these data, the relevant factor for influencing mismatch effects is relatedness, of which animacy relatedness is a special case. However, Penta and Pearlmutter (2015) manipulated HN–LN relatedness between nouns that were matched in animacy and normatively unassociated, and they found no difference in mismatch effects as a function of relatedness. They argued that some additional factor(s), such as association, may have been inadvertently confounded with relatedness in Barker et al.’s study. Thus, the evidence concerning the influence of relatedness on agreement errors is contradictory.

Experiment 3 represents an effort to address several issues whose resolution could clarify questions about the role of relatedness in increasing mismatch effects. The specific research question is whether preambles with animacy-matched versus -mismatched nouns increase mismatch effects when additional relatedness between the nouns is minimized. If animacy-match, on its own, is sufficient to increase mismatch effects in the current experiment, there would be several implications.

First, Penta and Pearlmutter (2015) interpreted their data as showing that relatedness in general does not increase mismatch effects, but their results, and the results of Barker et al.’s (2001) second experiment, leave open the possibility that differences in animacy relatedness might. This is because their preamble nouns were matched on animacy in the related and unrelated preamble versions. Thus, this implies the possibility that Penta and Pearlmutter’s design was “hiding” an effect that might have been detected with a manipulation of animacy match.
Second, the results of Barker et al.’s (2001) first experiment are consistent with the claim that animacy-match, on its own, can account for the mismatch effect patterns obtained in that study. Although their manipulation of animacy relatedness was confounded with synonymy, it may be the case that differences in synonymy are not very predictive of mismatch effects beyond what is predicted by animacy-match. Thus, a replication of Barker et al.’s findings of increased mismatch effects for animacy-matched versus -mismatched preambles would lend support to their hypothesis.

The mechanism proposed by Barker et al. (2001) for the effects on agreement from relatedness makes straightforward predictions: Mismatch effects will be larger for animacy matched versus mismatched nouns. The assumption is that activation spreads reciprocally between animacy-matched nouns in the lexical-conceptual network because the nouns are linked by their common animacy information. A consequence of this is that information about the LN, including its plural number information, is more highly activated, which may interfere with agreement computation. Barker et al. (2001) also observed that agreement errors tended to be less likely following preambles with animate versus inanimate HNs, which supports the prediction for smaller mismatch effects overall for preambles with animate HNs in Experiment 3.

On the other hand, if animacy match does not increase mismatch effects in Experiment 3 when the preamble nouns are minimally similar on other dimensions, which Penta and Pearlmutter’s (2015) results predict, this could indicate either that the manipulation(s) of relatedness in Barker et al.’s (2001) experiments involved unintentional manipulations of one or more additional factors, or that Penta and Pearlmutter’s manipulation of relatedness differed from Barker et al.’s manipulation in critically important ways.

An additional goal of Experiment 3 is to evaluate the possibility that differences in how relatedness was measured in Barker et al. (2001) and Penta and Pearlmutter (2015),
respectively, contributed to the observed differences in the influence of relatedness on mismatch effects. Specifically, in the survey Barker et al. used both to confirm the confound between animacy and relatedness in their Experiment 1, and separately, to confirm the relatedness manipulation in their Experiment 2, they asked participants to rate noun pairs based on semantic overlap, where the maximum rating was anchored to “synonymy” (Barker, 2001). In contrast, on two separate surveys, Penta and Pearlmutter did not instruct participants to specifically use synonymy as a criterion for judging the stimuli pairs, and the maximum value of their rating scale was anchored to “very related.”

Analyses of separately obtained synonymy ratings of Penta and Pearlmutter’s (2015) (Experiments 1–2) stimuli showed a positive correlation between relatedness and synonymy \( r = .64, p < .001 \). The correlation notwithstanding, relatedness and synonymy are not patently interchangeable, assuming that synonymy is a more constrained measure of relatedness, and that it implies similarity in semantic features. Because the design of Experiment 3 depends critically on minimizing similarity between the preamble nouns, synonymy was chosen as the measure of relatedness in Experiment 3.

A large online survey (described below) of synonymy was conducted as part of the stimuli norming process. Animacy match (i.e., \( \text{match} \) vs. \( \text{mismatch} \)) was regressed against mean synonymy ratings. The residual synonymy measure was taken to indicate the amount of HN–LN relatedness that was not accounted for by animacy match. The stimuli in Experiment 3 were normed using this measure in order to vary animacy match while controlling for differences in relatedness vis-à-vis rated synonymy. To assess the independent contribution of animacy match on agreement with this control in place, local noun number, HN animacy, and animacy match are manipulated within NP+PP preamble sets similar to those used in Barker et al. (2001).
3.1 Experiment 3

Predictions

Barker et al.’s (2001) findings predict larger mismatch effects for preambles with animacy-matched nouns (e.g., AA, II) in comparison to preambles with the animacy-mismatched nouns (e.g., AI, IA). Evidence of larger mismatch effects for animacy-matched versus -mismatched nouns in the form of a significant interaction between LN number (singular vs. plural) and Animacy match (match vs. mismatch) would be taken as support for the hypothesis. Barker et al.’s findings also support the prediction that mismatch effects will be smaller for animate-HN preambles (e.g., AA, AI) relative to inanimate-HN preambles (e.g., II, IA). Support for the prediction would be evidenced by an interaction between HN-animacy (animate vs. inanimate) and local noun number. Finally, the LN number manipulation is predicted to generate more agreement errors following preambles with plural versus singular local nouns (Bock & Miller, 1991), which would be confirmed by a main effect of the factor.

Method

Participants

One hundred eighteen Northeastern University undergraduates participated in the experiment for course credit. Excluded from all analyses were data from four participants who were non-native English speakers, and from five participants who contributed too few usable trials. Data from three participants were excluded due to audio recording-equipment failure. Data from eight participants were lost due to experimenter error. Data from the remaining 98 were included in all analyses.

Materials and design
### Table 3

Experiment 3 Stimuli and Synonymy Ratings by Condition

<table>
<thead>
<tr>
<th>HN–LN Animacy</th>
<th>Noun Number</th>
<th>Preamble</th>
<th>Synonymy Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Original</td>
</tr>
<tr>
<td>AA SS</td>
<td>The doctor near the poor villager</td>
<td>1.92 (0.54)</td>
<td>0.03 (0.43)</td>
</tr>
<tr>
<td>SP</td>
<td>The doctor near the poor villagers</td>
<td>1.83 (0.52)</td>
<td>−0.01 (0.43)</td>
</tr>
<tr>
<td>AI SS</td>
<td>The doctor near the poor village</td>
<td>1.31 (0.19)</td>
<td>−0.04 (0.23)</td>
</tr>
<tr>
<td>SP</td>
<td>The doctor near the poor villages</td>
<td>1.27 (0.25)</td>
<td>−0.04 (0.27)</td>
</tr>
<tr>
<td>IA SS</td>
<td>The hospital near the poor villager</td>
<td>1.37 (0.34)</td>
<td>0.02 (0.35)</td>
</tr>
<tr>
<td>SP</td>
<td>The hospital near the poor villagers</td>
<td>1.32 (0.33)</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>II SS</td>
<td>The hospital near the poor village</td>
<td>1.75 (0.75)</td>
<td>−0.14 (0.59)</td>
</tr>
<tr>
<td>SP</td>
<td>The hospital near the poor villages</td>
<td>1.75 (0.70)</td>
<td>−0.09 (0.56)</td>
</tr>
</tbody>
</table>

*Note.* The synonymy rating scale was 1 (not at all synonymous) to 7 (highly synonymous); standard deviations are in parentheses. A = animate; I = Inanimate; SS = singular head, singular local noun; SP = singular head, plural local noun.

Forty critical items like that shown in Table 3 were selected from 54 candidate items on the basis of norming data. Each began with a head NP (e.g., *The doctor, The hospital*) followed by the same locative preposition (e.g., *near*) within items, followed by a local noun phrase containing an adjective (e.g., *poor*), the determiner *the*, and the local noun (e.g., *villager(s), village(s)*). The head noun was always singular, and the eight different versions were created by varying head noun animacy (animate; A vs. inanimate; I), (head–local noun) animacy match (match vs. mismatch), and local noun number (singular vs. plural). Animate nouns denoted humans in 38 of the stimulus sets, and they denoted animals in the remaining 2. Inanimate nouns always denoted non-human artifacts or natural kinds. A list of the test stimuli is given in Appendix A.

All nouns and adjectives from the critical item preambles were closely matched across conditions on length in characters, phonemes, and syllables, and on word frequency (based on the 51-million word SUBTLEX-US corpus; Brysbaert, 1996). Preambles were designed to
be plausible and not to encourage a distributive reading (e.g., Eberhard, 1999). Published association norms (Nelson et al., 2004) were consulted to ensure that there was no association between any of the nouns within a given critical item.

In addition to the critical items, 144 filler items were included. Thirty-two fillers had the same structure as the critical items but with plural heads. Sixteen of these had animate HNs, and sixteen had inanimate HNs. Four of the animate HN fillers had animate singular LNs, four had animate plural LNs, four had inanimate singular LNs, and four had inanimate plural LNs. Four of the inanimate HN fillers had animate singular LNs, four had animate plural LNs, four had inanimate singular LNs, and four had inanimate plural LNs. The other 112 fillers had a variety of structures, but none had structures like the critical items. The critical items and filler items were combined in eight counterbalanced lists, each containing all of the fillers and exactly one version of each critical item.

**Synonymy rating survey**

As the goal of the study was to evaluate whether animacy match between preamble nouns can increase the size of the mismatch effect when the nouns are not otherwise semantically related, controlling for the known confound between animacy match and semantic relatedness (e.g., Barker et al., 2001) was critically important. Following Barker et al. (2001), relatedness was defined in terms of synonymy in Experiment 3. Ratings of synonymy were obtained in an online survey conducted through the Amazon Mechanical Turk service (Buhrmester, Kwang, & Gosling, 2011). Two hundred seven participants provided data for the survey. A total of 1,877 noun pairs, including all the HN–LN pairs from Experiment 1, Experiment 3, and Penta and Pearlmutter (2015, Experiments 1 and 2); all the prime–preamble-noun pairs from Experiment 1; and 633 filler noun pairs were rated using the instructions presented in Appendix B. Filler pairs were included to maximize the range of relatedness across the set.

The nouns in a pair were presented side-by-side in capital letters (e.g., a. KITTEN b. CAT).
Participants were instructed to rate how synonymous the words were on a scale from 1 (not at all synonymous) to 7 (highly synonymous), and they indicated their choice by clicking a radio button on the screen. Each pair was presented only once to a given participant and in randomized order by Mechanical Turk (Buhrmester et al., 2011). Participants were free to discontinue their participation at any time (the mean number of ratings per participant was 164). They were paid $0.01 per rating. Seventeen or 18 ratings were obtained for each noun pair.

The HN–LN pairs for the 40 critical items in Experiment 3 were normed for synonymy by 167 of the 204 participants who contributed to the survey following the procedure described above. Table 3 shows the mean synonymy ratings and standard deviations by condition for the stimuli. The data were submitted to a 2 (LN number: singular vs. plural) × 2 (HN animacy: A vs. I) × 2 (Animacy match: match vs. mismatch) ANOVA on synonymy means, with items as the random factor. The results were consistent with Barker et al. (2001; see also Barker, 2001): Animacy-matched items were more related than animacy-mismatched items ($F(1,39) = 48.97, MS_e = 0.40, p < .001$). Singular-singular versions of noun pairings were more related than singular-plural versions ($F(1,39) = 4.41, MS_e = 0.03, p < .05$). There were no other main effects and no interactions between any of the factors (all $ps > .21$).

To remove from the synonymy measure the proportion of its variance that could be accounted for by animacy match, the HN-LN pairs from the 54 candidate preambles were combined with singular and plural versions of 50 animacy-matched and -mismatched filler pairs (described above). Mean synonymy for all items was then regressed against animacy-match as a two-level, effects-coded factor (1 = match, 0 = mismatch). Table 3 shows the residualized synonymy means and standard deviations by condition for the stimuli.

Reanalyses of the data using residualized synonymy means again confirmed the difference in relatedness for singular–singular items versus singular–plural, with the former being more
related than the latter ($F(1,39) = 32.43$, $MS_e = 0.002$, $p < .001$). Critically, there was no
difference in relatedness between animacy-matched and animacy-mismatched items ($F(1,39) < 1$, $MS_e = 0.39$, $p > 0.54$). This was interpreted as a manipulation of animacy relatedness
with a control for additional relatedness (i.e., rated synonymy) between the preamble nouns.

Procedure

Each participant was run individually. Participants read each preamble aloud as soon
as it appeared on the computer screen and provided their own ending to form a complete
sentence. They were instructed to create completions as quickly and as naturally as possible,
but received no additional instructions otherwise. There were 184 trials preceded by five
practice trials.

Each individual trial began with a space-bar press, followed by a fixation cross, which
appeared on the left edge of the computer screen. After 1000 ms, a preamble appeared on
the screen positioned such that its first character was in the same place as the fixation cross.
Preambles appeared for 40 ms/character or 1000 ms, whichever was greater. If a participant’s
rate of speech slowed considerably at any point during the experiment, the experimenter
encouraged the participant to respond more quickly.

Results

Across 3,920 trials, there were 1,822 correctly inflected responses, 88 agreement errors,
1,149 uninflected responses, 807 miscellaneous cases, and 54 trials with no response. Agreement
error rates and counts of each response type by HN–LN animacy and local noun number are
shown in Table 4. Analyses were performed on untransformed and on arcsine-transformed
mean error proportions, calculated both with and without the dysfluent cases, aggregated
separately over participants and items. Except where noted, the statistical patterns were
identical for the analyses on both raw and transformed error proportions, and when dysfluency
cases were excluded. Therefore, the statistics reported for Experiment 3 are for analyses on untransformed error proportions with dysfluency cases included.

The separate omnibus ANOVAs for the error data, uninflected responses, and miscellaneous responses were a series of 2 (LN number: singular vs. plural) × 2 (HN animacy: A vs. I) × 2 (Animacy match: match=AA,II vs. mismatch=AI,IA) ANOVAs, with participants ($F_1$) and with items ($F_2$; Clark, 1973) as the random factor. Figure 6 shows the pattern of mismatch effects.

The plural-singular mismatch effect was confirmed by a main effect of local noun number ($F_1(1,97) = 24.73, MS_e = 542.53, p < .001; F_2(1,39) = 40.93, MS_e = 97.07, p < .001$). There was no main effect of HN animacy ($p$s $> .17$), except on the by-items analysis on arcsine-transformed error proportions with dysfluencies excluded ($F_1(1,97) = 0.93, MS_e = 0.30, p = .336; F_2(1,39) = 4.75, MS_e = 0.13, p < .05$). The main effect of Animacy match was significant in the by-items analyses ($F_2(1,39) = 7.49, MS_e = 61.65, p < .01$), but there was no sign of an effect in the by-participants analyses ($F_1(1,97) = 1.61, MS_e = 230.61, p = .207$). Errors were more common for animacy-matched (AA, II) versus animacy mismatched (AI, IA) preambles.

The interaction of HN animacy and LN number was significant also only in the analyses by-items ($F_1(1,97) = 2.46, MS_e = 281.24, p = .12; F_2(1,39) = 5.22, MS_e = 99.64, p < .05$). Mismatch effects were smaller for animate HN preambles ($F_1(1,97) = 9.65, MS_e = 415.05, p < .01; F_2(1,39) = 8.41, MS_e = 96.24, p < .01$) compared to inanimate HN preambles ($F_1(1,97) = 24.72, MS_e = 408.72, p < .001; F_2(1,39) = 36.67, MS_e = 100.47, p < .001$).

There was no interaction between Animacy match and HN animacy (all $p$s $> .21$), nor between Animacy match and LN number (all $p$s $> .19$). Finally, there was no interaction between LN number, HN animacy, and Animacy match (all $p$s $> .33$).

Analyses of uninflected responses indicated no significant interactions between LN number,
### Table 4

Experiment 3 Error Rates and Response Counts by Condition

<table>
<thead>
<tr>
<th>HN–LN Animacy</th>
<th>Noun</th>
<th>Number</th>
<th>Error Rate</th>
<th>Error</th>
<th>Correct</th>
<th>Uninflected</th>
<th>Misc</th>
<th>No Resp</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>SS</td>
<td>1.67</td>
<td>(0.69, 1.53)</td>
<td>3 (0)</td>
<td>177 (18)</td>
<td>214 (32)</td>
<td>89</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>8.29</td>
<td>(2.31, 1.99)</td>
<td>15 (3)</td>
<td>166 (15)</td>
<td>168 (15)</td>
<td>135</td>
<td>6</td>
</tr>
<tr>
<td>AI</td>
<td>SS</td>
<td>1.06</td>
<td>(1.14, 0.71)</td>
<td>2 (0)</td>
<td>186 (15)</td>
<td>206 (19)</td>
<td>85</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>5.70</td>
<td>(2.50, 1.68)</td>
<td>11 (2)</td>
<td>182 (14)</td>
<td>190 (17)</td>
<td>103</td>
<td>5</td>
</tr>
<tr>
<td>IA</td>
<td>SS</td>
<td>0.34</td>
<td>(0.20, 0.25)</td>
<td>1 (0)</td>
<td>294 (25)</td>
<td>91 (8)</td>
<td>91</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>7.53</td>
<td>(2.12, 1.80)</td>
<td>22 (1)</td>
<td>270 (24)</td>
<td>88 (8)</td>
<td>108</td>
<td>2</td>
</tr>
<tr>
<td>II</td>
<td>SS</td>
<td>0.68</td>
<td>(0.42, 0.68)</td>
<td>2 (0)</td>
<td>293 (20)</td>
<td>103 (6)</td>
<td>86</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>11.19</td>
<td>(2.75, 2.14)</td>
<td>32 (4)</td>
<td>254 (26)</td>
<td>89 (11)</td>
<td>110</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>88 (10)</td>
<td>1822 (157)</td>
<td>1149 (116)</td>
<td>807</td>
<td>54</td>
</tr>
</tbody>
</table>

**Note.** Error rates are reported as percentage errors out of errors plus correct responses for critical items. Standard errors of the mean computed from the analyses by participants and by items are in parentheses. Response counts include dysfluency counts, in parentheses. HN = Head noun; LN = Local noun; SS = singular head, singular local noun; SP = singular head, plural local noun; Misc = Miscellaneous; No Resp = No response.

![Mismatch effects by HN–LN animacy](image)

**Figure 6.** Experiment 3 mismatch effects by HN–LN animacy. Error bars indicate ± 1 SEM, computed from the analyses by participants.
Discussion

Experiment 3 tested the hypothesis that preambles with animacy-matched nouns (AA, II) would create larger mismatch effects than preambles with animacy-mismatched nouns (AI, IA) when additional relatedness between the nouns was controlled. Analyses of the error data indicated that preambles with animacy-matched nouns did not reliably increase mismatch effects compared to preambles with animacy-mismatched nouns. This is consistent with Penta and Pearlmutter (2015), who showed that differences in HN–LN relatedness did not predict differences in mismatch effects.

In Penta and Pearlmutter’s (2015) study, HNs and LNs were matched in animacy (AA, II) within preamble versions, and rated relatedness was varied between the versions. One concern raised in the introduction to Experiment 3 was that Penta and Pearlmutter did not include animacy-mismatched preamble versions against which the animacy-matched versions could be compared, with the implication being that a potential influence of relatedness on mismatch effects had been unintentionally missed. A direct comparison of the animacy-matched preambles to the animacy-mismatched preambles in Experiment 3 suggests that an effect of relatedness was not simply going undetected in Penta and Pearlmutter. Rather, the current results corroborate Penta and Pearlmutter’s original finding.

For the purposes of controlling for differences in HN–LN relatedness while varying animacy relatedness in the Experiment 3 stimuli, animacy-match was regressed against the mean rated synonymy (representing several thousands of ratings in total) for more than 1,800 pairs of words. The residual ratings of synonymy were taken to be the amount of relatedness...
between the pairs left over when animacy-match was allowed to account for as much of the variance in the ratings as possible. Subsequent testing of the pairings used in the critical item preambles confirmed that the confound between animacy relatedness and synonymy had been effectively controlled. This supports the interpretation of the Experiment 3 results as evidence that HN–LN animacy match does not reliably increase mismatch effects when the nouns are unrelated on other dimensions. Barker et al.’s (2001) hypothesis is thus not supported by these data.

An additional concern was that differences in the measures of relatedness used in Penta and Pearlmutter (2015) and Barker et al. (2001), respectively, could have contributed to the differences in the observed effects on agreement (i.e., no increase in mismatch effects, a significant increase in mismatch effects, respectively). As presumably the same measure of relatedness was used in Experiment 3 as Barker and colleagues used in their experiments, one interpretation of these findings is that animacy-match is unlikely to be able to account for the mismatch effect patterns reported in Barker et al. One implication is that the increased mismatch effects from animacy match in Barker et al. was not actually an influence of animacy match, but rather the influence of an additional factor, such as association. While it has thus far not been possible to evaluate this claim directly, the results of Experiment 3, in combination with the results of Penta and Pearlmutter, indicate that some other factor influenced mismatch effects in the experiments of Barker et al.

The results of Experiment 3 do not rule out the possibility that differences in animacy match make some contribution to differences in mismatch effects. Despite not reaching significance, the differences in the mismatch effect patterns in Experiment 3 were consistent with the predictions of Barker et al.’s (2001) hypothesis: Mismatch effects were largest for II preambles, and smallest for AI preambles. Mismatch effects for AA and IA preambles were close in size and fell between the AI and II preambles. Separately, animate-HN preambles in
Experiment 3 tended to be less likely than inanimate-HN preambles to generate agreement errors, and they were less prone to interference from plural LNs. These findings parallel Barker et al.’s findings of a near-significant main effect of HN animacy and a significant interaction between HN animacy and local noun number.
4. General Discussion

This dissertation examined questions concerning the effects of two types of word relationships on subject-verb agreement errors during sentence production. Experiments 1–2 investigated association, which refers to the probability of two words occurring together irrespective of their meaning. Experiment 3 investigated semantic relatedness, which refers to similarity between two words based on aspects of their meaning.

Experiments 1–2: Associative Priming

Experiment 1 explored the influence of association on mismatch effects, following up on previous research by Penta and Pearlmutter (2015) that showed reliable increases in mismatch effects for preambles with associated versus unassociated HN–LN pairs. The study was anchored by a set of predictions concerning a possible priming-based mechanism for the effects of association on agreement errors. In a modified version of the standard sentence completion task, preambles were presented after sentences containing nouns that were associates of the HN, or the LN, or neither noun.

In both the direct and the indirect priming scenarios proposed in Experiment 1, association affects the timing of production of the targeted preamble noun by making information about the noun available relatively earlier during sentence planning. The key distinction between the mechanisms is that, in the direct priming scenario, associatively priming the HN does not have an effect on the timing of planning of the LN, while in the indirect scenario, priming the HN indirectly influences the rate at which the LN is planned relative to the HN. On both accounts, primed-LN preambles were predicted to generate the largest mismatch effects, with the assumption being that a primed- versus an unprimed LN would be planned with
less temporal separation from the HN. With respect to the primed-HN preambles, the direct priming scenario predicted that these would yield the smallest mismatch effects because a primed- versus an unprimed-HN should be planned with more temporal separation from the LN. In contrast, the indirect priming scenario predicted that association would increase mismatch effects for primed-HN preambles relative to preambles in the Neutral condition. Here, the assumption is that the system is more likely to plan the preamble nouns in parallel when the processing of the HN is facilitated by associative priming.

The predictions of the direct priming account were reasonably well supported by the error data, while the predictions of the indirect priming account were not supported: In addition to the main effect of local noun number (obtained in all three experiments), which replicated the head-local mismatch effect, mismatch effects were numerically smallest for primed-HN preambles, largest for primed-LN preambles, and the difference between them was significant. This showed that association between prime nouns and the preamble nouns affected the probability of agreement errors in participants’ preamble completions.

The results of Experiment 1 were notable in many respects. First, they extend the findings of Penta and Pearlmutter (2015) by providing evidence that the influence of association on mismatch effects does not require that the relationship be construed between the two nouns of a given preamble. Second, whereas association between preamble nouns corresponded to increased mismatch effects in Penta and Pearlmutter, associatively primed-HN preambles in Experiment 1 tended to elicit fewer agreement errors than preambles in which the HN was not primed. Although mismatch effects were not reliably smaller in the primed-HN condition compared to the Neutral condition, the overall pattern of mismatch effects in this experiment and in Penta and Pearlmutter implies that association is capable of both increasing and decreasing the probability of agreement errors, depending upon how the relationship is instantiated.
Experiment 2 was intended to clarify a question about whether producing the additional non-prime words in the prime sentences might have created needlessly long separations between when the primes were presented and when the preamble was presented (i.e., stimulus onset asynchronies; SOAs), which could have weakened the downstream effects of priming on the production of agreement errors. Taking into account claims that the activation levels of a primed noun decay in 2–3 seconds (Chang et al., 2006; Levelt et al., 1999), Experiment 2 extracted the prime nouns from the vignettes and presented them serially in word lists immediately before their respective preambles. The results of Experiment 2 were that mismatch effect did not differ as a function of the prime condition, suggesting that the associative relationships in the different prime conditions did not differ significantly in their ability to affect mismatch effects when the prime nouns were presented at the reduced SOAs.

Reducing the prime–target SOAs in Experiment 2 was expected to enhance the effects of associative priming, which was expected to correspond to clearer mismatch effects patterns than the ones observed in Experiment 1. It is not immediately clear why the results were inconsistent with expectations. A distinct possibility is that there may not have been enough separation between the primes and targets in Experiment 2 to support associative priming. At least one challenge to this is that the effects of associative priming on word ordering differences in participants’ picture descriptions in Bock (1986a), and separately, in Konopka and Meyer (2014) seem to have been obtained at relatively short SOAs (i.e., primes were presented immediately before their respective targets, though actual SOAs were not reported).

The main reason why priming agreement might be more sensitive to differences in SOAs than priming word order is that participants in Bock (1986a) and in Konopka and Meyer (2014) created their own subject NPs, while the participants in Experiments 1–2 were given ready-made subject NPs. The important difference is that when speakers have multiple options for ordering nouns, the decision about which noun to produce first is influenced
by when or how much information about the nouns is available at the onset of utterance formulation (cf. Bock, Irwin, & Davidson, 2004; V. Ferreira, 1996; Solomon & Pearlmutter, 2007; Wheeldon & Smith, 2003; Wheeldon et al., 2013). On the priming-based explanation for association’s effects on agreement errors, association causes relative shifts in the timing of planning for nouns whose order is given by the preamble (analogous to the mechanism proposed for semantic integration’s effects on production Gillespie, 2011; Nicol, 1995; Solomon & Pearlmutter, 2004).

The results of Experiment 2 also suggest that it may not be possible to influence mismatch effects by manipulating association between the preamble-external primes and their LN targets. This might be because the structure of NP+PP preambles results in SOAs being generally too large, or because producing a HN precedes the production of the LN, which may require encoding operations to suppress the LN’s activation. Neither of these would necessarily rule out priming the LN from within the preamble, however. There is evidence, for example, that when participants have to produce several words before a primed target, the priming effect is more reliably obtained when the prime and target are part of the same utterance (cf. Tabossi & Johnson-Laird, 1980; Traxler, Foss, Seely, Kaup, & Morris, 2000). Thus, evidence for association’s influence on mismatch effects from Penta and Pearlmutter (2015) is not inconsistent with the predictions of the priming mechanism, nor with the evidence for priming effects from prime-target pairs presented in the context of the same sentence.

Although the results of Experiments 1–2 were consistent with the predictions of an associative priming mechanism for association’s influence on agreement, the mechanism itself was not directly tested, and therefore, the experiments provide no direct evidence for or against such a mechanism. The explicit assumption made in those experiments is that priming with association affects the rate at which the HN and the LN of a preamble are planned.
relative to each other, varying as a function of the prime condition. The mismatch effect
differences predicted on the priming scenario would also be consistent with the possibility
that priming affects the degree to which the nouns are highly activated during agreement
computation without affecting the timing at which the nouns are planned.

An alternative mechanism along these lines could involve reciprocal-reinforcement, anal-
ogous to the model described in Barker et al. (2001). The relevant assumption on such a
model would be that association changes the strength of activation of the prime noun and the
target noun, of such that information about both nouns would be more highly active during
et al.’s mechanism was hypothesized not to involve associative links, and although it was
intended to account for the effects of relatedness for nouns in the same preamble, it could be
easily cast in terms of an associative network, given minor modifications.

Further testing would be needed to evaluate the possibility of a priming-based mechanism
for association’s influence on agreement. It should be possible to investigate many of the
issues identified in Experiments 1–2 by modifying the design of the experiments. For example,
SOA could be included as an additional variable in order to evaluate its role in associative
priming, and the question of whether preamble structure affects priming potentials might be
pursued through manipulations of association using the three-noun subject NP preambles
such as those used in Gillespie and Pearlmutter (2011). If the results of these follow-up
studies do not support the priming account, it would be reasonable to consider whether an
alternative mechanism that does not involve priming underlies association’s influence on
mismatch effects.

**Experiment 3: Animacy Relatedness**

Previous investigations of mismatch effect differences as a function of semantic relatedness
have yielded contradictory results: In one experiment, Barker et al. (2001) found larger mismatch effects when preamble nouns were matched in animacy and more synonymous compared to when the nouns were mismatched in animacy and less synonymous. In a second experiment, varying the amount of relatedness (rated synonymy) for nouns that were matched in animacy also yielded larger mismatch effects as a function of increased relatedness. In contrast, Penta and Pearlmutter (2015) varied relatedness across preamble nouns that were matched in animacy but also unassociated, and they found no reliable influence of relatedness. Experiment 3 sought to clarify some of the discrepancies between these studies by testing for increased mismatch effects for preambles with animacy- matched versus -mismatched nouns that were minimally similar on other dimensions.

To control for the known confound between animacy-match and relatedness, several thousands of ratings of synonymy were obtained for more than 1800 noun pairs through an online survey. These ratings were regressed against animacy match, and the residualized ratings were used to minimize the amount of relatedness between HN–LN pairings that was not attributable to differences in animacy match. The main finding of Experiment 3 was no interaction between local noun number and animacy match, suggesting that HN–LN relatedness does not reliably increase the size of the mismatch effect. This result was consistent with the lack of an interaction of local noun number and relatedness reported in Penta and Pearlmutter (2015). These findings imply that Barker et al.’s manipulations of semantic relatedness were likely to have been confounded with association, or possibly semantic integration.

The confound between animacy match and synonymy, as reported by Barker et al. and confirmed in Experiment 3, suggests that animacy match is an important component of synonymy. This is arguably also a reflection of the importance of animacy distinctions in terms of how conceptual information is organized (cf. Carey, 2009; Opfer & Gelman, 2010).
That is, in people’s valuations of similarity, animacy properties may be more heavily weighted than other semantic properties on average, being both a salient and self-relevant feature. In addition, the pervasiveness of animacy distinctions in human language (e.g., Comrie, 1989; Corbett, 1979; Dahl, 2008; Givón, 1976; Ransom, 1977; Zaenen et al., 2004), and the reliable influence of animacy on many sentence processing measures (e.g., Altmann & Kemper, 2006; Branigan et al., 2008; Carminati, van Gompel, Scheepers, & Arai, 2008; F. Ferreira, 1994; G. W. Humphreys et al., 2016; Szewczyk & Schriefers, 2011) would argue that, if HN–LN relatedness on any semantic feature is likely to have an influence on mismatch effects, it would be animacy.

Despite the several well-reasoned arguments as to why animacy match should increase the rate of agreement errors (e.g., Badecker & Kuminiak, 2007; Barker et al., 2001), however, there is little compelling evidence from the agreement literature to indicate that it does (e.g., Deutsch & Dank, 2009; Nicol & O’Donnell, 1999). Moreover, while the results of Experiment 3 were fairly consistent with Barker et al.’s (2001) claim that animate HNs reduce the probability of agreement errors, they were largely inconsistent with their hypothesis that mismatch effect are increased by animacy match. Although mismatch effects in Experiment 3 patterned similarly to those in Barker et al.’s study, the differences were not reliable.

**Association versus relatedness: An update**

Since Penta and Pearlmutter (2015), two additional agreement studies have manipulated semantic relatedness. In Brehm and Bock (2017), participants created conjoined-noun subject NP preambles where the conjoined nouns were either semantically related or unrelated (e.g., *The dish and the plates* vs. *The dish and the cats*). In that study, preambles with related nouns were a reliable predictor of slower speech onsets. In a separate study from Veenstra et al. (2015), participants described two depicted nouns using NP+PP structures.
where semantic relatedness was manipulated on the depicted nouns (e.g., *The apple next to the pear* vs. *The apple next to the car*). This study also found evidence of slower speech onsets when the descriptions had related nouns, and additionally, measures of speaker’s gazes indicated that they spent more time looking at pictures when the nouns were related. Both studies articulated the delays in processing attributed to relatedness in terms of the increased difficulty that arises when two similar nouns are active during planning (for compatible arguments, see also De Smedt, 1994; F. Ferreira & Swets, 2002; Solomon & Pearlmutter, 2007). The critical point is that, despite an apparent effect of relatedness on speech onsets and gaze durations, relatedness has little influence on mismatch effects: Neither study found evidence of an interaction between relatedness and local noun number.

These studies are consistent with the results of Experiment 3 and also with the results of Penta and Pearlmutter (2015). The picture being painted concerning relatedness is that it has measurable effects on many aspects of production (for a general overview, see Baars, 1992; Bock, 1996; for a review of the literature reporting such effects, see Abdel Rahman & Melinger, 2009), but these do not translate to an influence on subject-verb agreement errors.

Finally, Experiments 1 and 2 are still presumably the only experiments outside of Penta and Pearlmutter (2015) to have manipulated association, with the results of these studies collectively indicating a rather robust and variable influence of this factor on mismatch effects. One interesting observation is that, although association is strongly correlated with agreement errors, it does not appear to be correlated with other types of documented speech errors. Citing evidence from a corpus analysis from Hotoph (1980), Bock (1995b) notes, for example, that word substitution errors are more common for nouns that are taxonomically related than for nouns that are simply associatively related (cf. Dell, 1995). This puzzle suggests the need for a clearer theory for the role of association in affecting agreement processing specifically, and sentence production more generally.
Conclusion

What is true in life may be true in the life of a subject-verb agreement error: Some relationships matter more than others. This dissertation represented an effort to learn more about how associative and semantic relationships influence the production of agreement errors, and they offer more evidence in support of the claim that association influences mismatch effects, but relatedness does not. Agreement experiments that depend on distinguishing the effects of association from those of semantic relatedness face a number of challenges. One of the more important issues is that there are no clearly articulated accounts of agreement that incorporate models of lexical access, where questions about the relatedness and association distinctions are theory-critical (cf. Baars, 1992; Bierwisch & Schreuder, 1992; Bock, 1996; Bock & Griffin, 2000; Levelt, 1989; Roelofs, 2008; Shelton & Martin, 1992). My hope is that this dissertation will prove useful to other researchers who are interested in probing the nature of agreement and the influence of factors that do not involve number information.
References


Appendix A:

Experiment 3 Stimuli

The singular versions of the stimuli from Experiment 3 are shown below. Animate HNs are shown first. Inanimate HNs are shown second. The animate LN’s and their respective adjectives are presented first. The inanimate LN’s and their respective adjectives are shown second. The four (singular) versions of each stimulus sentence were created by combining the each HN with each LN and its adjective. The plural local noun versions were created by making the LN plural.

(1) The clerk/laptop by the sneaky student/small television
(2) The child/van behind the protective cop/protective barrier
(3) The librarian/classroom near the noisy teenager/noisy radiator
(4) The trainer/brochure by the flashy model/flashy car
(5) The photographer/banner near the respected soldier/painted statue
(6) The scientist/notebook beside the curious reporter/curious fossil
(7) The doctor/hospital near the poor villager/poor village
(8) The customer/entrance behind the pushy guard/flimsy gate
(9) The owl/trap by the wild fox/wide field
(10) The singer/painting behind the loud promoter/red curtain
(11) The mayor/park near the exhausted athlete/luxurious condo
(12) The driver/contract beside the important merchant/important parcel
(13) The biologist/microscope beside the arrogant technician/complicated manual
(14) The wolf/boulder beyond the little cow/little river
(15) The manager/poster near the famous pianist/famous gallery
(16) The painter/portrait by the generous donor/gorgeous vase
(17) The artist/sculpture beside the remarkable scholar/remarkable novel
(18) The activist/fireplace by the rugged hunter/furry rug
(19) The researcher/envelope near the controversial expert/controversial manuscript
(20) The cashier/ATM by the excited tourist/expensive display
(21) The detective/switchblade beside the suspicious juvenile/suspicious package
(22) The professor/bookcase by the wealthy patron/donated drawing
(23) The historian/dictionary beside the dedicated sophomore/detailed map
(24) The economist/refrigerator near the broke intern/broken computer
(25) The vandal/wall behind the angry cleaner/ugly monument
(26) The executive/table beside the erratic designer/electronic chalkboard
(27) The spectator/fountain near the tired cyclist/new bench
(28) The astronomer/telescope by the proud amateur/tall window
(29) The inspector/tapestry by the foreign diplomat/exotic flower
(30) The lecturer/elevator near the confused visitor/confusing plaque
(31) The actor/satchel beside the fake doctor/fake gun
(32) The scientist/sign beside the interested volunteer/interesting exhibit
(33) The contestant/advertisement behind the gourmet chef/gigantic sandwich
(34) The thief/briefcase near the gullible traveler/valuable necklace
(35) The auctioneer/catalog beside the avid collector/rare stamp
(36) The politician/folder beside the honest journalist/historic document
(37) The stylist/gown near the wealthy shopper/modern mannequin
(38) The investigator/container behind the dishonest vendor/damaged relic
(39) The dictator/bomb by the armed negotiator/armored helicopter
(40) The spy/desk behind the talkative ambassador/thick partition
Appendix B:

Synonymy Rating Instructions

In this survey, you will be presented with pairs of words. For a given pair, we would like you to rate how synonymous the words are on a scale of 1 to 7, where a 7 means that the words are very synonymous, and a 1 means that the words are not at all synonymous.

For example, you might consider the word pair \textit{basement} and \textit{cellar} to be synonyms, corresponding to a 6 or 7 rating. On the other hand, a pair like \textit{rain} and \textit{hunger} describe completely unlike things and might be rated as a 1 or 2. A pair like \textit{kitchen} and \textit{bathroom} might fall somewhere in the middle of the scale.

Finally, you might encounter a pair like \textit{mechanic} and \textit{engine}, in which the things being described could appear together in familiar situations. **Please make your judgment based NOT on this type of connection, but instead, on how synonymous the two things being described are.** Since a \textit{mechanic} and an \textit{engine} are not the same kind of thing, you should give them a fairly low rating.

Remember, we are interested in your opinion about these word pairs. There are no right or wrong answers. However, it is important that you rate each pair only on how synonymous the words are.