Accommodating presuppositions is inappropriate in implausible contexts

Raj Singh[1], Evelina Fedorenko[2], Kyle Mahowald[2] and Edward Gibson[2]

[1]Institute of Cognitive Science, Carleton University
[2]Department of Brain and Cognitive Sciences, MIT

Send correspondence to any of the following authors:

Raj Singh, Carleton University DT2207, Ottawa, ON K1S 5B6, CANADA

Evelina Fedorenko, MIT 46-3037G, Cambridge, MA 02139, USA

Edward Gibson, MIT 46-3035, Cambridge, MA 02139, USA

Email: raj.singh@carleton.ca, evelina9@mit.edu, egibson@mit.edu

Abstract

According to one view of linguistic information (Karttunen, 1974; Stalnaker, 1974), a speaker can convey contextually new information in one of two ways: (a) by *asserting* the content as new information; or (b) by *presupposing* the content as given information which would then have to be *accommodated*. This distinction predicts that it is conversationally more appropriate to assert implausible information rather than presuppose it (e.g., Heim, 1992; Stalnaker, 2002; von Fintel, 2008). A second view rejects the assumption that presuppositions are accommodated; instead, presuppositions are assimilated into asserted content and both are correspondingly open to challenge (e.g., Gazdar, 1979; van der Sandt, 1992). Under this view, we should not expect to find a difference in conversational appropriateness between asserting implausible information and presupposing it. To distinguish between these two views of linguistic information, we performed two self-paced reading experiments with an on-line stops-making-sense judgment. The results of the two experiments – using the presupposition triggers *the* and *too* – show that accommodation is inappropriate (makes less sense) relative to non-presuppositional controls when the presupposed information is implausible but not when it is plausible. These results provide support for the first view of linguistic information: the contrast in implausible contexts can only be explained if there is a presupposition-assertion distinction and accommodation is a mechanism dedicated to reasoning about presuppositions.

1. Introduction

Consider the following sentences:

(1) A bouncer argued with Bill for a while.

(2) The bouncer argued with Bill for a while.

These sentences are similar in several respects. First, they are syntactically similar, differing only in the determiner: *a* in (1) vs. *the* in (2). Second, under standard semantic treatments of (1) and (2) (e.g., Heim, 1982; van der Sandt, 1992, among others), both entail the conjunction of (3a) and (3b).

(3a) there was a bouncer *x* in the context/situation

(3b) *x* argued with Bill for a while

Accordingly, a speaker who asserts either of (1) or (2) is committed to the truth of (3a) and (3b) (Grice, 1975; Williamson, 1996). If the hearer trusts the speaker, they will also accept (3a) and (3b), and this information will then become shared knowledge, or *common ground*, between them for the remainder of the discourse, where the common ground is the set of assumptions taken for granted in the conversation by speaker and hearer (Lewis, 1969; Karttunen 1974; Stalnaker 1974, 1978, 1998, 2002; Heim 1983a; Clark & Marshall, 1981; Clark 1996).

Third, both (1) and (2) can be used to introduce a discourse referent, a bouncer *x*, which can be referred to later on by a pronoun, as in (4) (e.g., Karttunen, 1976; McCawley, 1979; Lewis, 1979; Heim, 1982, 1983b; van der Sandt, 1992; Groenendijk, Stokhof, & Veltman, 1995).

(4) A/the bouncer$_x$ argued with Bill for a while, and then he$_x$ threw Bill out.

And fourth, (1) and (2) are judged to be plausible or surprising in similar contexts. For example, in the context of Bill going to a club, as in (5), both (1) and (2) are unsurprising. But in

3

the context of Bill going to the circus, as in (6), (1) and (2) convey surprising information. (We evaluate these intuitions quantitatively in Experiment 1.)

(5) Plausible context:

Bill went to a club last night.

a. A bouncer argued with him for a while.

b. The bouncer argued with him for a while.

(6) Implausible context:

Bill went to the circus last night.

a. A bouncer argued with him for a while.

b. The bouncer argued with him for a while.

Despite these similarities, many researchers have assumed that (1) and (2) differ in a fundamental way: whereas (1) asserts (3a) as new information, (2) presupposes (3a) as background information that is not up for debate.[1] This is a special case of a more general architecture under which sentences encode two kinds of content, presuppositions and assertions, which get used during conversation in different ways: presuppositions mark what the speaker believes is or should be in the background to the conversation, and assertions mark what is new and up for debate (Karttunen, 1974; Stalnaker, 1974). We will refer to frameworks that include these two ways of adding information to a conversation as *presupposition / accommodation frameworks*.

Under one prominent formulation of the presupposition / accommodation framework, the conversational context can be thought of as a file containing file cards corresponding to individuals, with information on each card about the corresponding individual (Heim, 1982; Chierchia, 1995; Beaver, 2001).[2] Each conversational move either adds new file cards or modifies existing ones. (1)'s contribution to the file – because of the indefinite *a bouncer* – is to

---

[1]It is currently an open question in the literature whether there is also a uniqueness requirement on the presuppositional requirement on definites and, if so, what the right way to implement it is (in the semantics, pragmatics, anaphora mechanisms, domain restriction, or otherwise). We return to this in the discussion of Experiment 1.

[2] Other proposals that are similar in relevant respects include more recent pragmatic (e.g., Chemla, 2009b; Schlenker, 2008, 2009) and trivalent approaches (e.g., Beaver and Kraemer, 2001; Fox, 2008, 2012).

add a new file card, corresponding to a bouncer, and to say that this bouncer argued with Bill; (2)'s contribution – because of the definite *the bouncer* – is to find an already existing file card corresponding to a bouncer, and to say of that bouncer that (s)he argued with Bill for a while. Definite descriptions are thus anaphoric, requiring an antecedent in the context, while indefinite descriptions add new file cards. When the context is such that no antecedent for the definite is available, such as when (2) is uttered in a null context, the sentence is predicted to be uninterpretable; indeed (2) makes no sense when uttered in a null context, presumably because the search for an antecedent produces no result, whereas (1) is sensible.

The texts in (5) and (6) present an apparent problem for this analysis: note that in (5b) and (6b) the required antecedent is arguably unavailable but it is nevertheless possible to interpret the sentences. In (5b) we understand *the bouncer* to mean some contextually salient bouncer at the club, and in (6b) we understand *the bouncer* to mean some contextually salient bouncer at the circus. The interpretability of these texts can be reconciled with the predicted presupposition failure by appealing to so-called "presupposition accommodation" (Lewis, 1979), an inference that amends the context so that the presuppositional requirement of the sentence will be satisfied (Asher & Lascarides, 1998b; Simons, 2003; Beaver & Zeevat, 2007; von Fintel, 2008). In the case of *the bouncer*, the context is updated by accommodation of a discourse referent that is a bouncer, and then the sentence is interpreted with respect to this modified context.[3]

An empirical motivation for the existence of presuppositions in addition to assertions is the intuitive strangeness of (6b) in an implausible context like (6). Whereas both (5a) and (5b) are easily interpretable in the supportive context in (5), and (6a) is somewhat strange in the context of (6), (6b) seems especially inappropriate, because of the use of the definite *the bouncer* which presupposes an existing bouncer in the context of a circus. This additional strangeness of (6b) relative to (6a) can be explained by appeal to potential inappropriateness of accommodating an implausible presupposition.

Another motivation for the distinction between assertion and presupposition / accommodation concerns inference patters in complex sentences. For example, consider the polar question formed from (2): *did the bouncer argue with Bill?* The question presupposes that

---

[3] When accommodation is triggered by definite descriptions like *the bouncer* there are also so-called "bridging inferences." (Clark, 1975; Haviland and Clark, 1974; Asher and Lascarides, 1998a;, Roberts, 2003).

there was a bouncer and targets only the asserted content, asking whether the presupposed bouncer argued with Bill (for overviews see Soames, 1989; Beaver, 2001; for quantitative studies, see Chemla, 2009a; Romoli et al., 2011; Sudo et al., 2012).

Despite these motivations for a presupposition-assertion distinction, experimental or corpus evidence has not yet been provided in support of such a division. In particular, there is little quantitative evidence evaluating the conditions under which accommodation may be appropriate or not.[4] Indeed, corpus studies suggest that accommodation is frequent in naturally occurring discourse (e.g., Delin, 1992; Fraurud, 1990; Poesio & Viera, 1998; Spenader, 2002), and introspective judgments suggest that at least some accommodations are easy (e.g., (5b)). Such considerations have led some to question whether accommodation is real (e.g., Gazdar, 1979; Burton-Roberts, 1989; Gauker, 1998; Abbott, 2002; Thomason, Stone, & DeVault, 2006). Furthermore, there are proposals for dealing with the inference patterns in complex sentences alluded to above that nevertheless reject the assumption that presuppositions are distinguished as contextually given, background information (Gazdar, 1979; van der Sandt, 1992). These proposals differ in interesting ways, but they assume that when a presupposition is informative in the context, it is not distinguished in any way from asserted content. We will refer to this alternative as the *no-presupposition* framework. Critically for the purposes of this paper, the *no-presupposition* framework predicts no additional inappropriateness of (6b) compared to (6a).

These frameworks make different predictions concerning the conversational appropriateness of texts like (5) and (6). In particular, both frameworks predict (6a) and (6b) to be inappropriate (because of the implausibility of finding a bouncer in a circus context), but the presupposition/accommodation framework predicts an additional inappropriateness associated with (6b) relative to (6a), because it involves accommodation of implausible information (e.g., Heim, 1992; Stalnaker, 2002; Beaver and Zeevat, 2007; von Fintel, 2008), whereas the no-presupposition framework predicts no additional penalty associated with (6b) relative to (6a), because both involve assertion of implausible information.

Previous studies of the processing of definite noun phrases have shown various effects that might be due to accommodation, which we review here, but they leave open the question

_____

[4] Recent studies of definites have probed participants' truth-value judgments in contexts in which accommodation is not licensed because it would contradict the context. Abrusán and Szendröi (2013) argue against a preupposition/accommodation framework, and Schwarz (2014) argues in favor of such a framework. Our method compares accommodation vs. assertion when both are consistent with the context.

whether these effects would also be present with asserted content alone. Thus, they do not present evidence either for or against accommodation-specific mechanisms at work.[5]

Processing definite descriptions has been shown to be more costly in contexts without an appropriate antecedent than in contexts with an appropriate antecedent (Haviland & Clark, 1974; Seidenberg & Tanenhaus, 1981; O'Brien et al., 1988; Burkhardt, 2006, 2007; see also Schwarz, 2007; Schwarz and Tiemann, 2011; Tiemann et al., 2011; Tiemann et al., 2014) for similar evidence for the presuppositions of discourse particles like *too* and *again*). Furthermore, sentence disambiguation studies show that there is a preference for interpretations that minimize the amount of accommodation required, and hence presumably for a cost to accommodation (e.g., Crain & Steedman, 1985; Ferreira & Clifton, 1986; Altmann & Steedman, 1988; Tanenhaus et al. 1995; van Berkum, Brown & Hagoort, 1999; Nieuwland, Otten & van Berkum, 2007; Schwarz, 2007). However, it is not clear whether this cost should be attributed to accommodation of a presupposition or to assertion of new material.[6] In particular, does the cost arise because a new discourse referent is being added to the context (a cost which would also arise when a discourse referent is introduced with an indefinite noun phrase), or does the cost arise because the definite's presuppositional / anaphoric requirements are not satisfied?

One way to address this question would be to compare the processing of definite and indefinite noun phrases in contexts where no antecedent for the definite noun phrase is available, as in (5) and (6). In both cases the context needs to get updated with a discourse referent that is a bouncer; a greater cost for the definite noun phrases might then be attributable to the presuppositional/anaphoric requirements imposed by the definites. Frazier (2006) used a design similar to the one proposed here, but it did not identify any costs unique to accommodation (if there are any). A sample item is shown in (7).

(7a) Plausible, no presupposition trigger (indefinite)

My order was taken. Then a waiter rushed into the kitchen.

(7b) Plausible, presupposition trigger (definite)

---

[5] In addition to behavioral evidence, some studies have used event-related potentials (ERPs) to investigate anaphoric processes in noun phrases (e.g., Burkhardt, 2006, 2007; Schumacher, 2009; Hirotani & Schumacher, 2011), but none have been unambiguously linked to presupposition accommodation.

[6] Chemla and Bott (2013) and Schwarz and Tiemann (2011) investigate the relative costs of processing global versus local accommodation when a sentence allows for both possibilities. This ambiguity is only generated in complex sentences, and our critical items do not generate this ambiguity.

My order was taken. Then the waiter rushed into the kitchen.

(7c) Implausible, no presupposition trigger (indefinite)

My order was taken. Then a busboy rushed into the kitchen.

(7d) Implausible, presupposition trigger (definite)

My order was taken. Then the busboy rushed into the kitchen.

Frazier's results revealed a main effect of plausibility but the definite implausible condition was not read slower than the indefinite implausible condition. Instead, the fastest reading times were observed in the definite plausible condition (7b), and the slowest in the indefinite implausible condition (7c). However, because passive voice was used in the context sentence of Frazier's materials, it is not clear that any predictions can be formulated about the expected processing costs of (7a-d). In particular, a passive construction tacitly introduces an additional event participant (e.g., the order-taker in *My order was taken*; see Mauner, Tanenhaus, & Carlson, 1995; Burkhardt, 2007, for evidence that implicit arguments introduced by passives are processed). Consequently, the noun phrases that follow (*a/the waiter/busboy*) have to be interpreted either as co-referential with the order-taker or as distinct from the order-taker, sometimes necessarily so. For example, *the waiter* in (7b) following the context sentence *My order was taken* may be most naturally read as co-referential with the order-taker, whereas *the busboy* following the same context sentence seems to allow both a co-referential reading as well as a reading where the busboy is someone other than the order-taker. The indefinite continuations, on the other hand, seem to force a non-co-referential interpretation. Because of the additional event participant entailed by the use of the passive, it is unclear what the theories assumed here would predict about the relative costs associated with processing these items.

2. Experiment 1

Experiment 1 investigated materials consisting of two sentences, as in (5) and (6), repeated here as (8).

(8)
Context Sentence: Bill went to {a club / the circus} on Friday night.

Target Sentence: A / the bouncer argued with him there for a while.

The first sentence introduced a situation where some individual, introduced by a name (e.g., *Bill*), is said to have gone to some location (e.g., *a club* or *the circus*).[7] The second sentence introduced another animate entity, typically denoted by an occupation noun (e.g., *bouncer*), which interacted with the individual introduced in the first sentence in some way. The materials were constructed following a 2 x 2 design, manipulating (1) the strength of the semantic relationship between the location introduced in the first sentence and the noun in the second sentence (strong = plausible, weak = implausible), and (2) the presence or absence of a presupposition trigger preceding the critical noun in the second sentence (presupposition trigger = definite, no presupposition trigger = indefinite).

Both the presupposition/accommodation and the no-presupposition frameworks predict that the implausible versions should be more difficult to process than the plausible versions (e.g., Trueswell, Tanenhaus & Garnsey, 1994; Tanenhaus & Trueswell, 1995; Gibson & Pearlmutter, 1998). This is all that is predicted according to the no-presupposition accommodation framework. In contrast, under the presupposition accommodation framework, there should be additional inappropriateness with the implausible version when the presupposition trigger *the* is used, because it involves accommodation of implausible information.

In a pilot investigation we investigated these predictions using a self-paced reading task with a moving-window display (Just, Carpenter & Woolley, 1982) using materials like (5) and (6). Although we observed an effect of plausibility (the implausible conditions were read reliably more slowly than the plausible conditions), we did not observe any effects of definiteness. We therefore turned to a potentially more sensitive reading-time paradigm, one that includes a concurrent stops-making-sense task (Boland et al., 1990; Gibson et al., 1996). This method gives us two measures of interest: the proportion of stops-making-sense (SMS) judgments at each region and reading times (RTs) at each region. We assume that SMS judgments reflect in part the conversational appropriateness of a piece of text, and we assume that RTs reflect in part the computational complexity of processing, storing, and integrating a

---

[7]The determiner used for introducing the location – indefinite (as in *a club*) or definite (as in *the circus*) – varied across items and was chosen to sound most natural. Crucially, however, the determiner was always the same for any given location (and, as discussed below, each location was used once in the plausible condition, and once in the implausible condition), and thus a main effect of definiteness and/or an interaction between plausibility and definiteness could not be accounted for by the determiners used with the location nouns.

piece of text into the context or discourse model; the complexity of such operations is presumably increased when a text violates norms of appropriate conversation.

Methods

Participants We posted surveys for 120 workers on Amazon.com's Mechanical Turk using self-paced reading software developed by Tily & Gibson (in preparation). All workers were paid for their participation. Participants were asked to indicate their native language, but payment was not contingent on their responses.

Design and Materials

Forty items like (8) were initially constructed. The location noun (e.g., *club*, *circus*) was intended to make accessible a set of entities typically associated with that location (Altmann & Kamide, 1999; McRae, Ferretti & Amyote, 1977; Chambers et al., 2002). For example, the noun *club* should make *bouncer* accessible because clubs often have bouncers. The plausible conditions were constructed in a way such that for each location multiple entities (e.g., multiple bouncers) would be readily accessible, with one among them likely to be prominent. The implausible conditions were constructed by pairing a location with a noun denoting entities that do not typically occur in the context of this location (e.g., bouncers in the context of circuses).

Items were constructed in pairs, such that each location that was introduced in a context sentence had a strong semantic relationship with the critical noun in one item and a weak relationship with the critical noun in another item. For example, (8) above would have a pair item like (9) below. The location *club* has a strong relationship with *bouncer* in (8) and a weak relationship with *lion* in (9). Conversely, *circus* has a strong relationship with *lion* in (9) and a weak relationship with *bouncer* in (8).

(9)
Context Sentence: Adriana went to {the circus / a club} on Friday evening.
Target Sentence: A / the lion yawned at her there a few times.

This way of constructing the materials ensured that any potential differences between plausible and implausible conditions would result from the *relationships* between the locations introduced

in the context sentences and the critical nouns, rather than from some properties of the location / critical nouns.

Next, we conducted a norming study in order to ensure that the materials had desirable properties. For each of the 40 items, we asked three questions, as shown in (10a)-(10c), where X was a name, Y/W were locations (Y had a strong semantic relationship with Z, and W had a weak/no semantic relationship with Z), and Z was the critical noun:

(10)

    a. If X goes to Y, how likely is he/she to see[8] one Z there?

    b. If X goes to Y, how likely is he/she to see multiple Zs there?

    c. If X goes to W, how likely is he/she to see a Z there?

We wanted to ensure that (1) either a single or multiple entities were equally likely in the plausible context (this requires the ratings for (a) and (b) to be high and not different from each other); and (2) that there was a large difference in likelihood between the plausible conditions on the one hand and the implausible conditions on the other (this requires the ratings for (a) and (b) to be high and the rating for (c) to be low).

We posted 40 surveys on Amazon.com's Mechanical Turk using the software from Gibson, Piantadosi, and Fedorenko (2011). Participants were instructed to rate the likelihood of each event on a scale from 1 to 5 where 1 corresponded to "extremely unlikely" and 5 corresponded to "extremely likely."

Based on the results of the norming study, we first selected a set of 24 items consisting of 12 item pairs, such that the same location was used in both plausible and implausible conditions for each pair of items, as in the pair (8) and (9). For this subset, the average ratings for (a), (b) and (c) were 4.39, 4.44, and 1.63, respectively. Furthermore, we selected 8 additional items which were "unpaired." For the items that were constructed to be paired with these 8 items, one or both of the relevant criteria above was not satisfied based on the norming study results. For

---

[8] For a subset of the items in the norming study, we used *to interact primarily with* instead of *to see*, because with respect to certain occupations it is the interaction with one of the target entities that makes this entity prominent enough. For example, if you go to a restaurant, you may see multiple waiters; however, you are likely to interact only with one. Therefore, if in a context of a restaurant someone refers to *the waiter* (as in our materials: *John went to a restaurant. The waiter…*), we can make an easy inference that this must be the waiter that John interacted with or that was otherwise prominent for John.

this subset of 8 unpaired items, the average ratings for (a), (b) and (c) were 4.60, 4.39, and 1.65, respectively. We decided to include these unpaired items in order to increase the statistical power. In analyzing the data, we conducted an analysis where we excluded the 8 unpaired items, but because the results were similar to the results obtained by including all 32 items, we report the results with all 32 items included. (See Appendix A for a complete list of experimental materials used in Experiment 1.)

In addition to the 32 target items, the experiment included the 32 items from Experiment 2 (another experiment investigating the processing of presupposition accommodation; see the description in section 3 for more details) and 48 filler items. All of these materials consisted of pairs of sentences. The filler items were variable in the presence, type and location of implausibilities. In particular, in 24/48 fillers, the first sentence introduced a location, and the second sentence introduced an entity which was plausible given the context sentence (in eight of these, the second sentence began with *The …*, in another eight the second sentence began with *A/an …*, and in the remaining eight the second sentence began with a pronoun like *He/She …*). In another 10/48 fillers, the second sentence made no sense in the context of the first sentence (e.g., *Leah walked through a campsite for an hour. A socialite from Hollywood was setting fire to a tent.*). In 8/48 fillers, the second sentence involved a presupposition violation of some kind (e.g., *Few singers went to the hospital in an ambulance. The singer kept an eye on his vital signs.*; here, the use of *the singer* in the second sentence is infelicitous given that multiple singers are introduced in the first sentence). In the remaining 6 fillers, an implausibility of some sort occurred in the first sentence, and was typically followed by additional implausibilities in the later parts of the item (e.g., *Lloyd flew his wallet to a garden store last spring. He sang beautiful karaoke.*). The fillers were included in order to (a) mask the critical manipulation, thereby minimizing the possibility of the use of any conscious strategies, and (b) prevent participants from forming an expectation that a particular kind of implausibility / violation would occur at a particular point in the sentence.

Four experimental lists were created following a Latin Square design, such that each participant only saw one version of each item. A different pseudo-random order of the items within each list was created for each participant, such that critical items were always separated by at least one filler item. Before the experiment, participants were presented with several practice items. Participants took approximately 20 minutes to complete the experiment.

Procedure The experiment was run using Mechanical Turk self-paced reading (SPR) software developed by Tily & Gibson (in preparation). This method replicates effects that have been previously observed in laboratory environments, while taking advantage of crowd-sourcing in order to run more participants in a shorter time than would be possible in the lab. (Indeed, we also ran these two experiments in the lab on 32 participants, and found similar results.)

In the stops-making-sense paradigm, each trial began with a line centered on the screen, marking the length of the pair of sentences to be presented. For presentation purposes, each item was divided into nine regions, as shown in (11) below (words connected by underscores represent one region): (1) a name (e.g., *Bill* in (11)); (2) the phrase *went to*; (3) a location consisting of one to three words (usually a determiner and a noun, e.g., *a club / the circus*); (4) a temporal/spatial modifier consisting of two to four words (*on Friday night*); (5) the target noun phrase consisting of a determiner (*a / the*) and a noun (*bouncer*); (6) the main verb for the target sentence (*argued*); (7) a preposition plus object pronoun (*with him* or *with her*) or just the object pronoun (*him / her*), depending on the verb; (8) the word *there*; and (9) the ending, which was a modifier of the verb phrase and consisted of a single adverb or a prepositional phrase. The critical regions were defined as the critical noun phrase (the first region of the second sentence) and all subsequent regions.

(11) Bill went_to {a_club / the_circus} on_Friday_night. A_bouncer / the_bouncer argued with_him there for_a_while.

Participants were instructed to press the "S" key on their keyboards to reveal each region of the sentence. As each new region appeared, the preceding region was converted to a line again. They were instructed to keep pressing the "S" key as long as the sentence(s) made sense to them. They were instructed to press a different button – the "K" key – when an incoming word/phrase did not make sense in the context of the preceding words/phrases. Pressing the "K" key terminated the trial and the next trial began.

To ensure that participants read and understood the sentences, we included comprehension questions for the 48 filler trials. (Questions were not asked for the critical trials because asking about the later parts of the sentences could have biased participants to keep

reading past the point of implausibility to ensure they could answer the questions, and asking only about the earlier parts of the sentences could have biased participants to pay more attention to the earlier parts of sentences.) Participants were instructed that if they didn't press the "K" key at any time during a trial (implying that they thought the two sentences made sense), then a question would often appear. They were instructed to try their best to answer these questions correctly. They received feedback when they answered incorrectly. Participants were advised that because some Mechanical Turk workers sometimes do not pay close attention, we would evaluate their performance using three criteria: (1) participants should always press the "K" key when a sentence is weird in some way; (2) participants should not press the "K" key for sentences that make sense; and (3) participants should answer comprehension questions correctly. Participants were warned that they would not be paid if they made too many errors.

There are two key dependent measures in the stops-making-sense self-paced reading paradigm: (a) the reading time for each region; and (b) the proportion of trials on which participants choose to continue reading the item (i.e., the proportion of "continue"-responses) at each region.

Results

Although we posted surveys for 120 participants, only 114 completed the task. Of these 114 participants, we considered data from participants who identified themselves as native English speakers from the United States, thus excluding 5 more participants. Furthermore, data from participants with less than a 65% accuracy rate for the comprehension questions were excluded, leaving 105 participants to be included in the analyses. As a result of this trimming, 2.5% of the data were excluded.

Results: Stops-Making-Sense Task

Analyses reported here were conducted with the lme4 package (Bates et al., 2008) for the statistical language R (R Core Development Team, 2008). The proportions of "continue" responses for the four conditions across the eight regions are presented in Figure 1. Note that these responses are cumulative, such that each data point represents the proportion of total participants (including those who ended the sentence earlier) who have dropped out at that point.

Because of a lack of independence in the SMS responses among the regions, we focus on the initial region containing the presupposition trigger and then the final region.
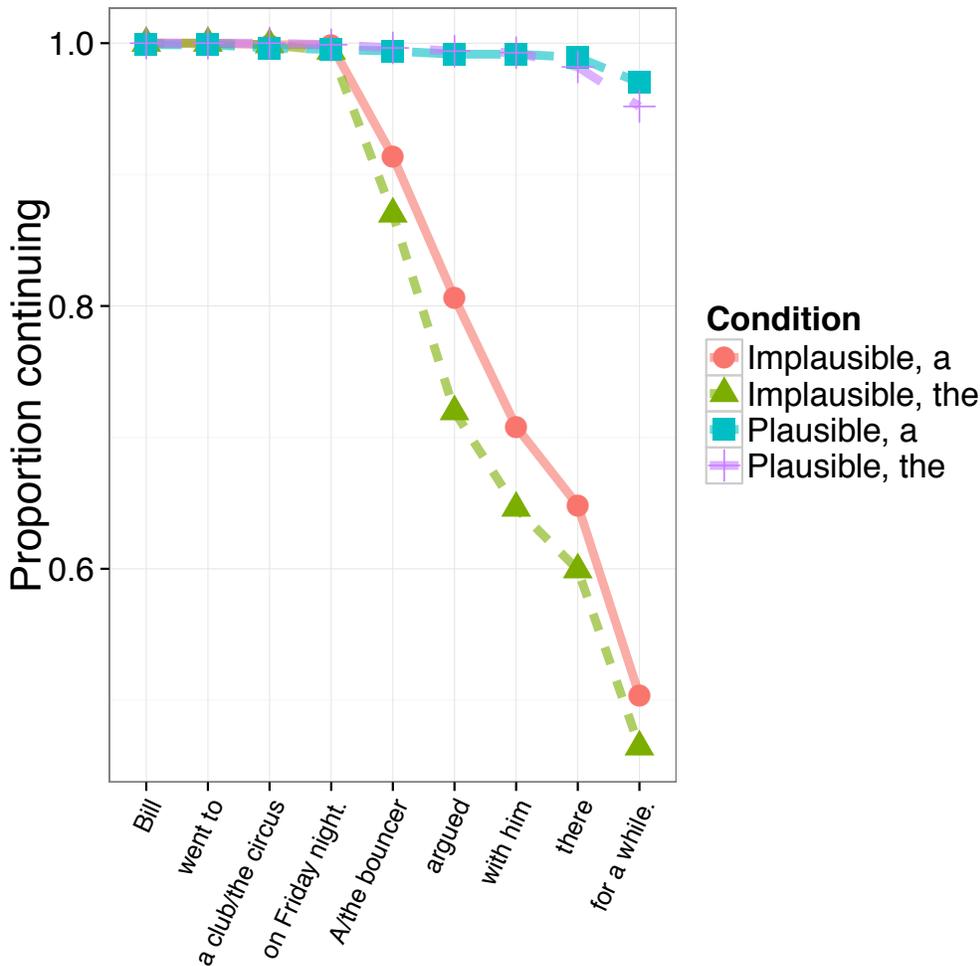
Figure 1: The proportions of "continue"-responses (out of total number of participants in the experiment, including those who already dropped out) for the four conditions across all the regions in Experiment 1.

First, we observed a large effect of plausibility that began at region 5 (e.g., the noun phrase *A/the bouncer* in (11)) and persisted until the end of the sentence. This effect was highly reliable in a mixed effect logistic regression with a maximal random effect structure for participants and items (with no interaction slopes in order to make the model converge), beginning in region 5 (p < .05) and becoming large by the end of the sentence (48% acceptance

in the implausible conditions vs. 96% acceptance in the plausible conditions; β = .38, z = 2.58, p

< .01). Participants were near ceiling in accepting plausible conditions (i.e., conditions with a strong semantic fit between the critical noun and the location introduced in the first sentence), independent of the determiner. Even at the last region of the sentence, the acceptance rate was 97% in the *a* condition and 95% in the *the* condition.

Because it is difficult to meaningfully compare percentages that are close to zero or one, we examined the implausible conditions separately from the plausible conditions (whose acceptance rates were close to one). For the implausible conditions, participants showed differential acceptability for the *a* and *the* conditions, with the *the* condition being terminated more often. This difference began numerically at region 5 (the region introducing the definite/indefinite contrast: implausible *a*: 91%; implausible *the*: 87%; β = .87, z = 1.56, p =.11), although it was not significant according to a maximal mixed effect model with random intercepts and slopes for participant and item. By the end of the sentence, summing over all participants, significantly more people dropped out in the *the* condition than in the *a* condition (46% vs. 50%). This difference was significant by a mixed effect logistic regression with random intercepts and slopes for participant and item (β = .37, z = 2.74, p < .01).
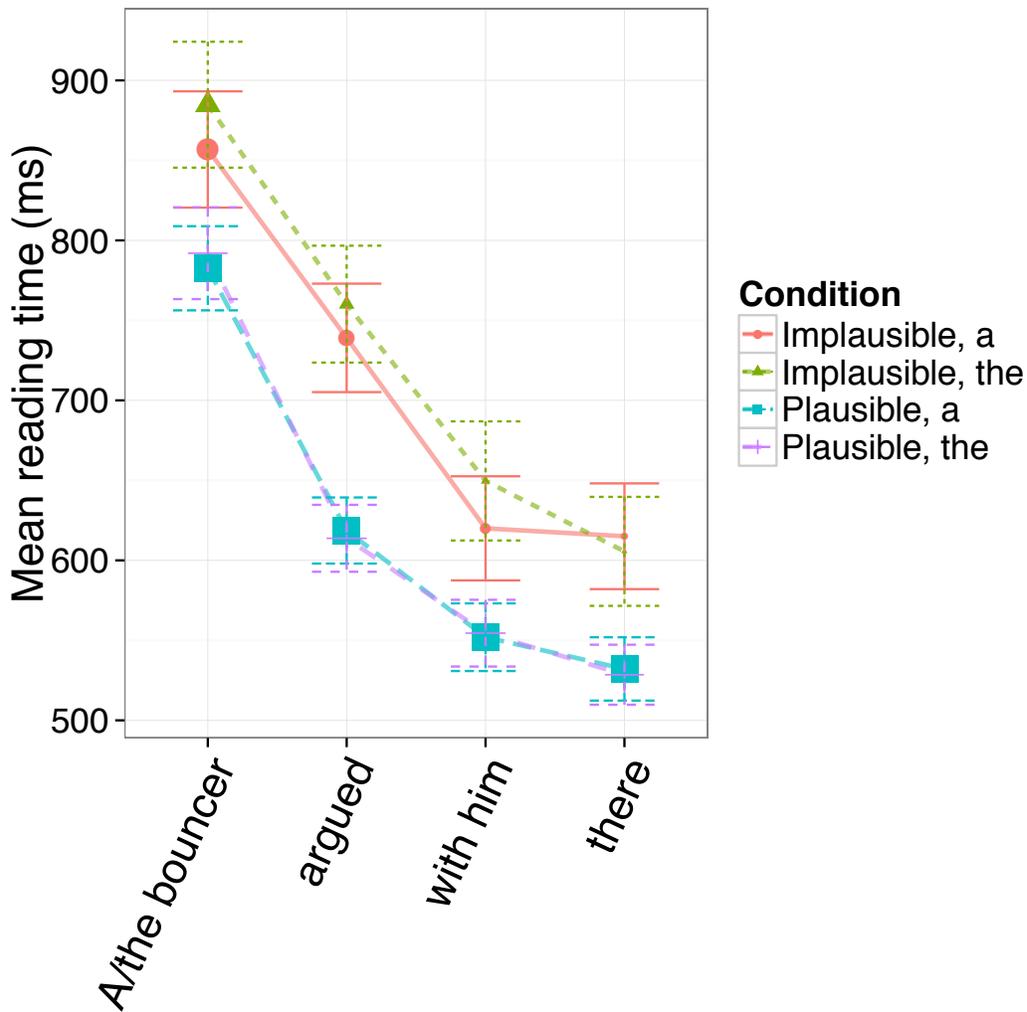
Figure 2: Raw reading times in each region in Experiment 1. The size of the point reflects the number of data points remaining at the region.

Results: Reading times

We analyzed only those reading times in which participants hit the "continue" button since there is an obvious time cost in switching to hitting the "stop" button. We do not present residual reading times here since the design involved participants dropping off at different points, and this lack of consistency across participants makes it difficult to accurately calculate residuals. The critical comparison is the difference between the definite condition and the indefinite condition in the implausible sentences compared to the difference between the definite and indefinite condition in the plausible sentences. That is, we focus on the difference between

the differences. We looked for this critical effect between "a" and "the" in the triggering region (Region 5) and the 2 regions immediately after. To trim extreme data points, we included only reading times between 100 ms and 3000 ms. The significance of the reading times was assessed by a mixed effect model predicting reading time from condition (plausibility x presupposition) with random intercepts for subject and item and with slopes by condition (i.e., a maximal mixed effect model as recommended by Barr et al. 2013). In all three regions, there was a main effect of plausibility such that plausible sentences were read more quickly (ps < .001). In all three regions, there were trends for presupposition sentences to be read more slowly, although this was not significant at p < .05. And there was no significant interaction between plausibility and presupposition, although there were trends for the implausible *the* sentences to be read slower than the implausible *a* sentences (a difference between conditions of 28 ms in Region 5; 21 ms in Region 6; and 30 ms in Region 7). There was no such trend in the plausible sentences, where the difference between the "the" condition and the "a" condition was near 0.[9] See Figure 2, which shows the raw reading times for each region after and including the presupposition trigger (Region 5), along with a 95% confidence interval. The size of the point reflects how many data points are in that particular region. Note that, as the sentence stops making sense, more people drop out—especially in the implausible conditions.[10]

Discussion

The main effect of plausibility in both the stop-making-sense proportions and the reading times replicates the observation that processing of implausible information is generally difficult (e.g., Trueswell, Tanenhaus & Garnsey, 1994; Tanenhaus & Trueswell, 1995; Gibson & Pearlmutter, 1998). The novel observation in Experiment 1 is a quantitative measure of how much more inappropriate it is to accept a discourse referent through accommodation than through assertion when the information is implausible (e.g., Heim, 1982, 1992; Soames, 1989; Stalnaker, 2002; von Fintel, 2008). This observation supports the existence of accommodation as

---

[9] Because of the difference in length between "a" and "the" (which may or may not affect the reading time of those words), we cannot directly interpret the main effects of "a" vs. "the" here. But, since the critical interaction depends only on the difference between "a" and "the" across conditions, we can do that analysis using raw reading times.
[10] See Schwarz (2014) for evidence that processing of definite descriptions when their existence presupposition is not satisfied by a visual context *does* lead to increased RTs relative to non-presuppositional indefinite controls.

a distinct, presupposition-specific inference.[11]

We did not detect any inappropriateness of accommodation compared with assertion when the information was plausible in the context. The absence of an effect of definiteness in the plausible conditions suggests that incorporation of a new discourse referent through bridging to already existing referents is as easy as incorporating a referent introduced via assertion. The costs of accommodating discourse referents found in previous experimental studies (e.g., Haviland & Clark, 1974; Seidenberg & Tanenhaus, 1981; O'Brien et al., 1988; Burkhardt, 2006, 2007) might therefore be interpreted simply as the cost of introducing new discourse referents, and not of accommodation itself. This result is somewhat surprising for any theory that assumes definites have anaphoric/presuppositional requirements that indefinites do not (e.g., Heim, 1982; van der Sandt, 1992; Chierchia, 1995; Beaver, 2001), because it suggests that i) the search for an antecedent (initiated by the definite noun phrase), ii) the failure to find one, and iii) the subsequent introduction of a new referent along with iv) the construction of a bridge between this referent and already existing ones are no more costly than simply introducing a new referent (as in the case of indefinite noun phrases).

Definite descriptions like *the bouncer* have sometimes been argued to entail – either through presupposition or assertion – that there is a unique bouncer in the context (e.g., Heim and Kratzer, 1998). Two anonymous reviewers asked whether our effect might not be due to a uniqueness entailment rather than the presupposition/assertion distinction itself. We do not think this is likely. First, Experiment 2 uses a different presupposition trigger, *too*, and we will see that the results are similar to those of Experiment 1. This suggests that the results are more likely about presupposition accommodation, rather than the particular semantics of the definite/indefinite contrast.[12] Second, it is not even clear that there is a uniqueness entailment. Note that the uniqueness requirement can be waived without contradiction: *John went to a club last night. The bouncer wouldn't let him in, but then another bouncer who works at the club came over and let John in*. This text is consistent and is relatively felicitous. The first bouncer is understood as the bouncer who initially encountered John, but this is consistent with there being

---

[11] A reviewer points out that this result is consistent with several ways that participants might approach the task. For example, they might take themselves to be the addressee of the text, or as someone who is not party to the conversation but is nevertheless assessing it for appropriateness. In either case, participants seem to be using the presupposition-assertion distinction in their assessment of the text.

[12] Another way to overcome debates about uniqueness in similar materials would be to use plural (in)definites like *some/the bouncers*.

other bouncers, and we are later introduced to one of these other bouncers. However, when we modify the relevant sentences so that uniqueness entailments are explicit, the text is contradictory, and hence odd: #*John went to a club last night that had exactly one bouncer working there. The bouncer didn't let John in, but then another bouncer who was working there came over and let John in*.

3. Experiment 2

In order to investigate whether the results of Experiment 1 are not specific to the presupposition trigger *the,* but are indicative of accommodation more generally, we investigated the presupposition trigger *too* in Experiment 2. A sentence like *Peter will go swimming too*, with a pitch accent on *Peter*, triggers the presupposition that some salient individual other than Peter will go swimming (e.g., Kripke, 1990; Heim, 1992).

Methods
Participants The same 120 participants from Mechanical Turk who took part in Experiment 1 also took part in Experiment 2.
Design and Materials The experiment used materials consisting of two sentences, as shown in (12).

(12)
a. John will go to {the pool / the mall} this morning.
b. Peter will go swimming {tomorrow / too} after he gets back from school.

The first sentence introduced an individual (e.g., *John* in (12a)) and stated that this individual would go to some location (e.g., *the pool* / *the mall* in (12a)). The second sentence introduced another individual and stated that this individual would perform some action (e.g., *go swimming* in (12b)). The materials were constructed following a 2 x 2 design, manipulating (1) the strength of the semantic relationship between the location introduced in the first sentence and the action in the second sentence (strong = plausible, weak = implausible), and (2) the presence or absence of a presupposition trigger in the post-verbal position in the second sentence. This position was

occupied either by the presupposition trigger *too* or by a temporal modifier like *tomorrow* indicating some time in the future.

The location noun in the first sentence was intended to make accessible a set of actions/events typically associated with that location. For example, the noun *pool* should make *swimming* accessible (possibly along with other actions like *tanning*) because swimming is common in the context of pools. The implausible conditions were constructed by pairing a location with an action that doesn't typically occur in the context of the location (e.g., swimming in the context of a mall). The temporal modifiers (e.g., *tomorrow* in (12b)) served as control modifiers with no presupposed content and were intended to be interpreted as plausibly contrastive with the time that was set up in the first sentence. For example, in (12b) above, *tomorrow* in the second sentences is intended to be contrasted with *this morning* in the first sentence.

Forty-four items like (12) were initially constructed. Items were constructed in pairs, such that each location that was introduced in a context sentence had a strong semantic relationship with the critical action in one item and a weak relationship with the critical action in another item. For example, (12) above would have a pair item, like (13) below. The location *pool* has a strong semantic relationship with *swimming* in (12), and a weak relationship with *shopping* in (13). Conversely, *mall* has a strong relationship with *shopping* in (13), and a weak relationship with *swimming* in (12).

(13)
a. Luke will go to {the mall / the pool} on Friday.
b. Ann will go shopping {on Saturday / too} once she comes back from cello practice.

Similar to Experiment 1, this way of constructing the materials ensured that any potential differences between plausible and implausible conditions would result from the *relationships* between the locations introduced in the context sentences and the critical verbs denoting the actions, rather than from some properties of the location nouns / critical verbs.

Next, we conducted a norming study in order to ensure that the materials had desirable properties. For each of the 44 items (22 pairs, each with a likely and unlikely location for the relevant action) we asked a question, as exemplified in (14):

(14) If John goes to the pool, how likely is he to go swimming there?

We wanted to ensure that that there was a large difference in likelihood between the plausible and implausible conditions.

We posted 40 surveys on Amazon.com's Mechanical Turk using the software from Gibson, Piantadosi & Fedorenko (2011). Participants were instructed to rate the likelihood of each event on a scale from 1 to 5 where 1 corresponded to "extremely unlikely" and 5 corresponded to "extremely likely."

Based on the results of the norming study, we selected a set of 32 items (i.e., 16 item pairs, such that the same location was used in both plausible and implausible conditions across items, as in Experiment 1). For this subset, the average rating for the likely location was 4.45, and the average rating for the unlikely location was 1.70. See Appendix C for a complete list of experimental materials used in Experiment 2.

In addition to the 32 target items, the experiment included the 32 items from Experiment 1, and 48 fillers, as described above.

Procedure The experiment was run using the Mechanical Turk self-paced reading software discussed above, concurrently with Experiment 1. For presentation purposes, each item was divided into eleven regions, as shown in (15) below (words connected by underscores represent one region): (1) a name; (2) the phrase *will go to*; (3) a location consisting of two or three words (a determiner and a noun phrase); (4) a temporal modifier consisting of one to four words; (5) another name; (6) the phrase *will go*; (7) the critical verb-form, which was either a gerund (e.g., *swimming*) or an infinitival form (e.g., *to meditate*); (8) a temporal modifier consisting of one to four words or the word *too*; (9) the word *after*, *before*, *when*, or *once*; and (10)-(11) the sentence ending consisting of a subordinate clause divided into two regions according to natural constituent groupings.

(15) John will_go_to {the_pool / the_mall} this_morning. Peter will_go swimming {tomorrow / too} after he_gets_back from_school.

Results: Stops-Making-Sense Task

The same 120 participants from Experiment 1 also participated in Experiment 2. As discussed in Experiment 1, 15 of these participants either did not identify themselves as native English speakers, or answered fewer than 65% of the comprehension questions correctly, leaving 105 who were analyzed. As a result of this trimming, 1.8% of the data were excluded. The proportions of "continue"-responses for the four conditions across the eleven regions are presented in Figure 3.
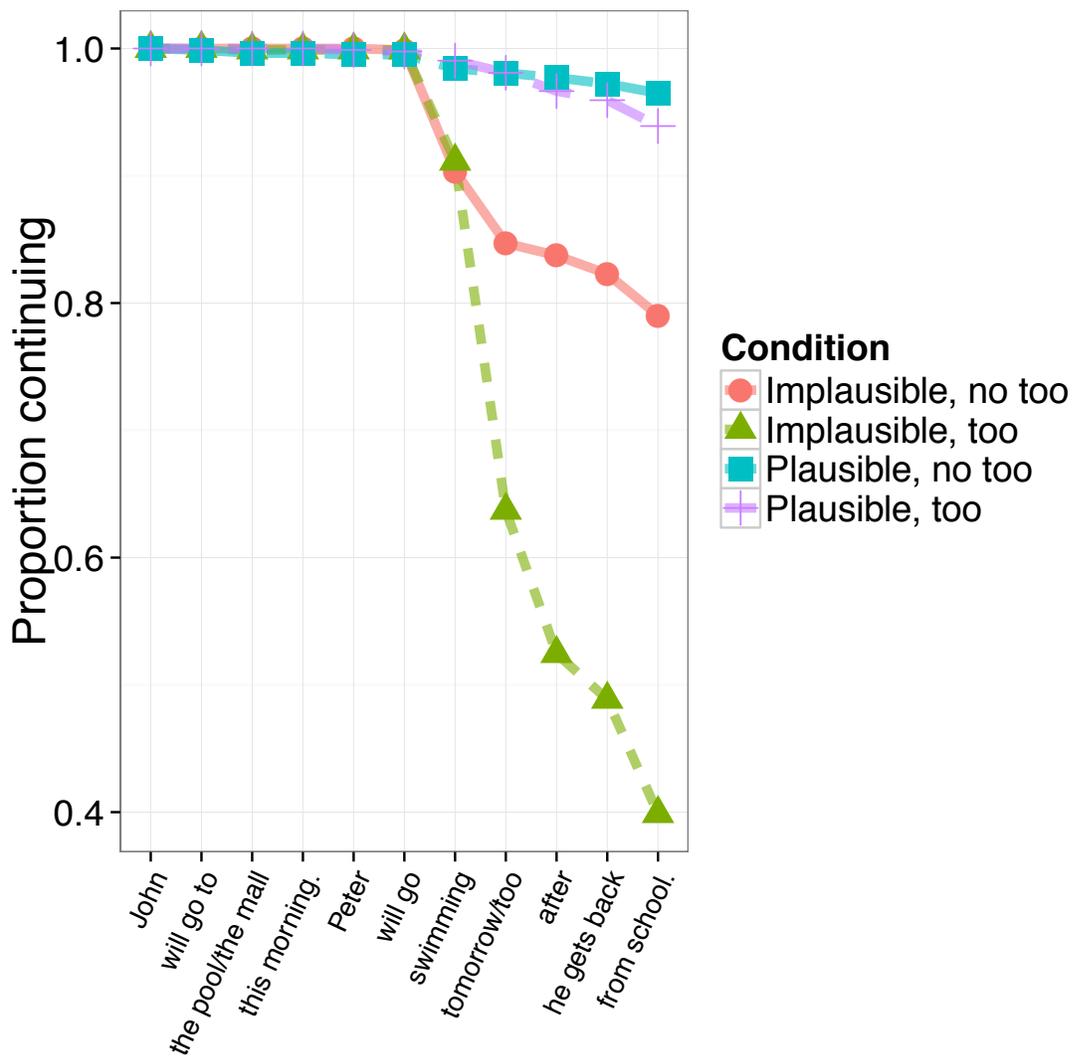
Figure 3: The proportions of "continue"-responses for the four conditions across all regions.

The pattern of results was similar to that observed in Experiment 1. First, we observed a large effect of plausibility that begins at Region 7 (i.e., the appearance of the unrelated verb) and persists until the end of the sentence. This effect was reliable in a mixed effect logistic regression with maximal random intercepts and slopes by participant and item, excluding interaction slopes ($\beta=2.92$, $z=1.95$, $p=.05$) for Region 7 and becomes larger in later regions. Furthermore, participants were near ceiling in accepting plausible conditions independent of the presence of the presupposition trigger *too*. In fact, no difference was observed between the *too* and temporal-modifier conditions for the plausible sentences in any region. Even at the last region of the sentence, the acceptance rate was 96% (97% in the temporal-modifier condition; 94% in the *too* condition).

Because it is difficult to meaningfully compare percentages that are close to zero or one, we examined the implausible conditions separately from the plausible conditions (whose acceptance rates were close to 1). For the implausible conditions, participants showed differential acceptability for the *too* and the adverbial conditions, with the *too* condition being terminated more often. This difference began at region 8 (at region 8: implausible adverbial: 85.0%; implausible *too*: 64%; $\beta = 2.47$; $z = 4.75$; $p < .0001$). By the end of the sentence, the proportion of participants remaining in the implausible adverbial condition was 79% compared to just 40% in the implausible *too* condition.

Results: Reading times

Reading times were calculated for each region exactly as in Experiment 1 (with data points restricted to those between 100ms and 3s). Figure 4 shows the mean raw reading time for each region after and including the triggering region (Region 8), along with a 95% confidence interval. The size of the point reflects how many data points are in that particular region. Note that, because the violation condition uses a much shorter word (*too*) than the words used in the non-violation condition (*tomorrow*, etc.), we cannot directly interpret the main effect of violation. But the critical comparison is the difference between *too* and *tomorrow* in the plausible condition compared to the difference between *too* and *tomorrow* in the implausible condition. This comparison is not affected by the difference in length. We fit a linear mixed effect model for each region predicting RT with fixed effects of plausibility, determiner, and their interaction and a maximal random effect structure (random intercepts for participant and item with random

slopes for determiner, plausibility, and their interaction). Critically, we found an interaction in region 9 (one region after the triggering region) between plausibility and presupposition violation. The difference between the violation condition and non-violation condition is larger in the implausible case (676 ms – 587 ms = 89 ms), whereas the difference between the two conditions in the plausible case is actually in the opposite direction (578 ms – 604 ms = -26 ms). To assess the significance of this interaction, we ran a likelihood ratio test comparing the full linear mixed effect model to a null model with the same random effect structure and fixed effects of presupposition violation and plausibility but no interaction. The result of the likelihood ratio test tells us that including the interaction significantly improves the model ($\beta=202$, $\chi2(1)=$ 27.00, p < .0001). Under any reasonable correction for multiple comparisons across these three regions, this result is highly significant. There was no significant interaction in the triggering

region (Region 8) or in Region 10.
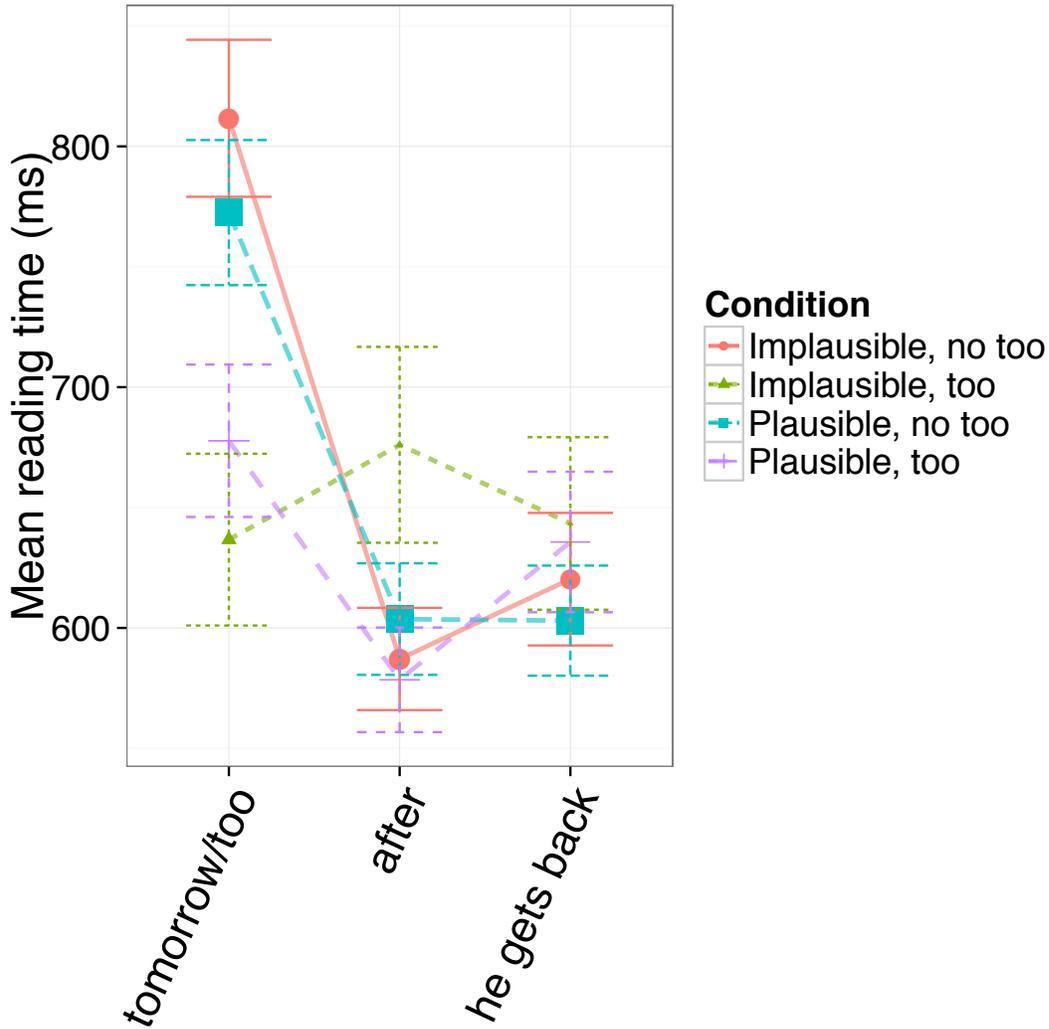
Figure 4: The plot shows reading times for Regions 8, 9, and 10. The size of the dot represents the number of data points in that condition and region. The error bars are 95% confidence intervals. There is an interaction in region 9 between the presupposition violation and plausibility, as shown by the green bar being significantly higher than the other three in that region.

Discussion

The pattern of results in Experiment 2 was similar to that observed in Experiment 1. In particular, accommodation of the presupposition triggered by *too* was inappropriate when it was

implausible but not when it was plausible, as shown by a greater proportion of SMS judgments and longer RTs. The increased RT difficulty of processing implausible information when it needs to be accommodated in Experiment 2 provides further evidence for the presupposition / accommodation framework over the no-presupposition framework.[13] The extra cost follows from the assumption that, all else being equal, processing complexity should vary with inappropriateness.

There was one way in which the results of Experiment 2 were different from those of Experiment 1: in Experiment 2, the inappropriateness of implausible accommodations was reflected not only in the SMS data, but also in the reading times. This difference across the two experiments might be because *too* may be more resistant to accommodation than *the* is, as has been suggested by several researchers (e.g., Geurts and van der Sandt, 2001, 2004; Simons, 2001; Abusch, 2010; Beaver and Zeevat, 2007). However, it has also been noted that *too*'s resistance to accommodation can be overcome if the presupposition is contextually plausible (e.g., von Fintel, 2008; Chemla and Schlenker, 2012); the lack of difficulty in the plausible conditions supports this assumption. Furthermore, the differences between *too* and *the* in implausible contexts is only suggestive, because the materials are completely different across the two experiments, which makes interpreting such differences in effect sizes problematic.

4. Conclusions

Our results provide clear evidence in favor of the presupposition / accommodation framework over the no- presupposition framework. In both experiments, we found that it is more inappropriate to integrate implausible information when it is presupposed than when it is asserted. This follows from the assumption that presuppositions are *shared* by speaker and hearer; it is inappropriate for a speaker to use a sentence with a contextually implausible presupposition because – unlike with assertions – the hearer is expected to accept this presupposition without discussion. The results therefore suggest that whatever the nature of the cognitive systems that decide how much credence to assign to information, they are sensitive to

---

[13] Two reviewers point out that our critical sentences can be interpreted as making sense in these contexts, say, as just listing events that have happened or will happen. The fact that so many participants judge the texts as nonsensical suggests that they are interpreting them as we had intended, with narrow focus on the subject and hence not in this charitable way.

whether the information has been presented as a presupposition or not.

Our results fit with other observations from the linguistics literature that suggest that presuppositions are components of meaning that are distinct from assertions. For example, as noted in the introduction, presuppositions project out of embeddings differently from assertions. For example, the question *At the club last night, did you see the bouncer?* presupposes that there was a bouncer there (both *yes* and *no* answers are committed to the existence of a bouncer), but the question "At the club last night, did you see a bouncer"? does not: the answer "no" is consistent with there being no bouncer at the club. Presupposition-assertion distinctions are also manifested in various discourse interactions. For example, in response to "The bouncer argued with Bill", the objection "No, that's false," cannot be used to object to the presupposition that there was a bouncer; rather, the presupposition is accepted, and the objection is to the assertion that this bouncer argued with Bill. Objections to the presupposed component of meaning are always more complicated (von Fintel, 2004; Cummins, Amaral, and Katsos, 2012).

Acknowledgements

References

Abrusán, M., & Szendröi, K. (2013). Experimenting with the king of France: Topics, verifability, and definite descriptions. *Semantics and Pragmatics*, 6(10), 1-43.

Abusch, D. (2010). Presupposition triggering from alternatives. *Journal of Semantics*, 27(1), 37-80.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

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

Asher, N., & Lascarides, A. (1998a). Bridging. *Journal of Semantics*, 15, 83-113.

Asher, N., & Lascarides, A. (1998b). The semantics and pragmatics of presupposition. *Journal of Semantics*, 15, 239-299.

Beaver, D. (2001). *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.

Beaver. D., & Krahmer, E. (2001). Presupposition and partiality: Back to the future? *Journal of Logic, Language, and Information*, 10, 147-182.

Beaver, D., & Zeevat, H. (2007). Accommodation. In G. Ramchand and C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces* (pp. 503-538). Oxford: Oxford University Press.

van Berkum, J.J.A., Brown, C.M., & Hagoort, P. (1999). Early context effects in sentence processing: Evidence from event-related potentials. *Journal of Memory*

and Language, 41, 147-182.

Boland, J.E., Tanenhaus, M.K., & Garnsey, S.M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, 29, 413-432.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related potentials. *Brain and Language*, 98(2), 159-168.

Burkhardt, P. (2007). The P600 reflects costs of new information in discourse memory. *Neuroreport*, 18(17), 1851-1854.

Chambers, C.G., Tanenhaus, M.K., Eberhard, K.M., Filip, H., & Carlson, G. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language*, 47, 30-49.

Chemla, E. (2009a). Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics*, 17(4), 299-340.

Chemla, E. (2009b). Similarity: Towards a unified account of scalar implicature, free choice permission, and presupposition projection. Under revision for *Semantics & Pragmatics*.

Chemla, E., & Bott, L. (2013). Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and Cognitive Processes*, 28(3), 241-260.

Chemla, E., & Schlenker, P. Incremental vs. symmetric accounts of presupposition projection: An experimental approach. *Natural Language Semantics*, 20(2), 177-226.

Chierchia, G. (1995). *Dynamics of Meaning*. Chicago: University of Chicago Press.

Clark, H.H. (1975). Bridging. In R.C. Schank & B.L. Nash-Webber (Eds.), *Theoretical Issues in Natural Language Processing* (pp. 169-174). New York: Association for Computing Machinery.

Clark, H.H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H.H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. Joshi and B.L. Webber (Eds.), *Elements of Discourse Understanding* (pp. 10-63). Cambridge: Cambridge University Press.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D.R.Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural Language Parsing* (pp. 320-358). Cambridge: Cambridge University Press.

Cummins, C., Amaral, P., and Katsos, N. (2012). Experimental investigations into the typology of presupposition triggers. *Humana Mente*, 23, 1-16.

Delin, J. (1992). Properties of *it*-cleft presupposition. Journal of Semantics, 9, 289-306.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348-368.

von Fintel, K. (2004). Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In A. Bezuidenhout & M. Reimer (Eds.), *Descriptions and Beyond* (pp. 315-341). Oxford: Oxford University Press.

von Fintel, K. (2008). What is presupposition accommodation, again? *Philosophical Perspectives*, 22(1), 137-170.

Fox, D. (2008). Two short notes on Schlenker's theory of presupposition projection. *Theoretical Linguistics*, 34, 237-252.

Fox, D. (2012). Presupposition projection from quantificational sentences: trivalence, local accommodation, and presupposition strengthening. Manuscript, MIT, Cambridge, MA.

Fraurud, K. (1990). Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics*, 7, 395-433.

Frazier, L. (2006). The big fish in a small pond: Accommodation and the processing of novel definites. Manuscript, UMass Amherst, Amherst, MA.

Geurts, B., & van der Sandt, R. (2001). Too. In R. van Rooij & M. Stokhof (Eds.), *Proceedings of the 13th Amsterdam Colloquium* (pp. 180-185), Amsterdam.

Geurts, B., & van der Sandt, R. (2004). Interpreting focus. *Theoretical Linguistics*, 30, 1-44.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human processing system. *Cognition*, 59, 23-59.

Gibson, E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science*, 2, 262-268.

Gibson, E., Piantadosi, S., & Fedorenko, K. (2011). Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass*, 5(8), 509-524.

Grice, H.P. (1975). Logic and Conversation. In P. Cole & J.L. Morgan (Eds,), *Syntax and Semantics, Volume 3: Speech Acts* (pp. 41-58), New York, NY: Academic Press.

Groenendijk, J., Stokhof, M., & Veltman, F. (1995). Coreference and contextually restricted quantification. In M. Simons & T. Galloway (Eds.), *Proceedings of the Fifth Conference on Semantics and Linguistic Theory* (pp. 112-129). Ithaca, NY: CLC Publications.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Heim, I. (1983a). On the projection problem for presuppositions. In M. Barlow, D. Flickinger, & N. Wiegand (Eds.), *Proceedings of the Second West Coast Conference on Formal Linguistics* (pp. 114-125). Stanford, CA: Stanford University Press.

Heim, I. (1983b). File-change semantics and the familiarity theory of definiteness. In R. Baürle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Natural Language* (pp. 164-189). Berlin: Walter de Gruyter.

Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9, 183-221.

Hirotani, M., & Schumacher, P. (2011). Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics*, 24(3), 276-292.

Just, M.A., Carpenter, P.A., & Wooley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228-238.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1,

181-194.

Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground* (pp. 363-385). New York, NY: Academic Press.

Kripke, S. (1990). Presupposition and anaphora: Remarks on the formulation of the projection problem. Manuscript, Princeton University, Princeton, NJ.

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339-359.

Mauner, G., Tanenhaus, M.K., & Carlson, G.N. (1995). Implicit arguments in sentence processing. *Journal of Memory and Language*, 34, 357-382.

McCawley, J. (1979). Presupposition and discourse structure. In C.K. Oh & D. Dineen (Eds.), *Syntax and Semantics 11: Presupposition* (pp. 1-56). New York, NY: Academic Press.

McRae, K., Ferretti, T.R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, 12, 137-176.

Nieuwland, M.S., Otten, M., & van Berkum J.J.A. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related potentials. *Journal of Cognitive Neuroscience*, 19, 228-236.

O'Brien, E.J., Shank, D.M., Myers, J.L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 410-420.

Poesio, M. & R. Viera. (1998). A corpus-based investigation of definite description description use. *Computational Linguistics*, 24(2), 183-216.

Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, 26, 287-350.

Romoli, J., Sudo, Y., & Snedeker, J. (2011). An experimental investigation of presupposition projection in conditional sentences. In *Proceedings of SALT* 21 (pp. 598-608).

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4), 333-377.

Schlenker, P. (2008). Be Articulate! A pragmatic theory of presupposition projection. *Theoretical Linguistics*, 34(3), 157-212.

Schlenker, P. (2009). Local contexts. *Semantics and Pragmatics*, 2(3), 1-78.

Schumacher, P. (2009). Definiteness marking shows late effects during discourse processing: Evidence from ERPs. In S. Lalitha Devi, A. Branco, & R. Mitkov (Eds.), *Lecture Notes in Artificial Intelligence, Vol. 5847: Anaphora Processing and Applications* (pp. 91-106). Heidelberg: Springer.

Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics*, 24, 373-416.

Schwarz, F., & Tiemann, S. (2011). Presupposition processing: The case of German *wieder*. In M. Aloni, V. Kimmelman, F. Roelofsen, G.W. Sassoon, K. Schulz, M. Westera (Eds.), *Lecture Notes in Computer Science: Logic, Language, and Meaning (*pp. 200-209*)*. Berlin: Springer.

Schwarz, F. (2014). False but slow: Evaluating statements with non-referring definites. Manuscrupt, University of Pennsylvania.

Simons, M. (2001). On the conversational basis of some presuppositions. In R. Hastings, B. Jackson, & Z. Zvolensky (Eds.), *Proceedings of Semantics and Linguistic Theory 11* (pp. 431-448). Ithaca, NY: CLC Publications.

Simons, M. (2003). Presupposition and accommodation: Understanding the Stalnakerian picture. *Philosophical Studies*, 112(3), 251-278.

Spenader, J. (2002). Presuppositions in spoken discourse. Doctoral Dissertation, Stockholm University, Stockholm.

Stalnaker, R.C. (1974). Pragmatic presupposition. In M. Munitz & P. Unger (Eds.), *Semantics and Philosophy* (pp. 197-213). New York, NY: NYU Press.

Stalnaker, R.C. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 315-332). New York, NY: Academix Press.

Stalnaker, R.C. (1998). On the representation of context. *Journal of Logic, Language, and Information*, 7, 3-19.

Stalnaker, R.C. (2002). Common ground. *Linguistics and Philosophy*, 25, 701-721.

Sudo, Y., Romoli, J., Hackl, M., & Fox, D. (2012). Presupposition projection out of quantified sentences: Strengthening, local accommodation, and inter-speaker variation. In M. Aloni, V. Kimmelman, F. Roelofsen, G.W. Sassoon, K. Schulz, M. Westera (Eds.), *Lecture Notes in Computer Science: Logic, Language, and Meaning (*pp. 210-219)*. Berlin: Springer.

Tanenhaus, M.K., & Seidenberg, M. (1981). Discourse context and sentence perception.

Discourse Processes, 4, 197-220.

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

Thomason, R.H., Stone, M., & DeVault, D. (2006). Enlightened update: Informative presupposition in a computational architecture for pragmatic reasoning. Manuscript, Workshop on Presupposition Accommodation at the Ohio State University.

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., Knapp, J., & Beck, S. (2011). Psycholinguistic evidence for presuppositions: On-line and off-line data. In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn und Bedeutung 15* (pp. 581-595). Saarbrücken: Universaar-Sarland University Press.

Tiemann, ., Kirsten, M., Beck, S., Hertrich, I., & Rolke, B. (2014). Presupposition processing and accommodation: An experiment on wieder ('again') and consequences for other triggers. In F. Schwarz (Ed.), *Experimental Perspectives on Presuppositions*. Dordrecht: Springer.

Tily, H. & Gibson, E. (in preparation). Self-paced reading over Mechanical Turk.

Trueswell, J.C., Tanenhaus, M.K., and Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, 33, 285-318.

Williamson, T. (1996). Knowing and asserting. *The Philosophical Review*, 105(4), 489-523.

Appendix A: Materials for Experiment 1

1. Mike went to {a boathouse / a theater} two days ago. A / the boatman chatted with him there for a long time.

2. Ashley went to {a theater / a boathouse} in the spring. A / the ticket man talked to her there for a long time.

3. Seth went to {jail / a cafe} on Saturday night. A / the guard spoke to him there for a while.

4. Kristen went to {a cafe / jail} in the morning. A / the waiter served her there quickly.

5. Sue went to a {hair salon / a French restaurant} in the fall. A / the hairdresser addressed her there politely.

6. Carrie went to {a French restaurant / a hair salon} on Christmas eve. A / the chef welcomed her there warmly.

7. Adriana went to {the circus / a club} on Friday evening. A / the lion yawned at her there a few times.

8. Bill went to {a club / the circus} on Friday night. A / the bouncer argued with him there for a while.

9. John went to {school / a concert} on Monday afternoon. A / the substitute teacher spoke to him there briefly.

10. Gabriella went to {a concert / school} two weeks ago. A / the guitarist winked at her there flirtatiously.

11. Fred went to {a dance show / a clinic} in the evening. A / the choreographer smiled at him there from a distance.

12. Ben went to {a clinic / a dance show} on Wednesday evening. A / the doctor greeted him there calmly.

13. Cole went to {a pub / the gym} three days ago. A / the bartender recognized him there immediately.

14. Jane went to {the gym / a pub} in the morning. A / the trainer talked to her there loudly.

15. Mark went to {the courthouse / the post office} yesterday afternoon. A / the witness noticed him there from across the room.

16. Rachel went to {the post office / the courthouse} in the afternoon. A / the postal worker assisted her there in a friendly way.

17. Chase went to {the office / a beach} yesterday after lunch. A / the secretary helped him there for an hour.

18. Mary went to {a beach / the office} a few hours ago. A / the lifeguard warned her there about the weather.

19. Louise went to {the local art gallery / a spa} on Thursday night. An / the art dealer intrigued her there with a comment.

20. Katherine went to {a spa / the local art gallery} during the weekend. A / the masseuse complimented her there about her appearance.

21. Philip went to {a pool / a laboratory} on Tuesday evening. A / the swim instructor insulted him there rudely.

22. Janet went to {a laboratory / a pool} in the afternoon. A / the senior researcher inspired her there about science.

23. Dolores went to {a pharmacy / a zoo} on Friday afternoon. A / the pharmacist harassed her there for five minutes.

24. Todd went to {a zoo / a pharmacy} on Monday morning. A / the zookeeper shouted at him there unexpectedly.

25. Jeff went to {a car repair shop / a mansion} yesterday at noon. A / the mechanic consulted with him there briefly.

26. Tina went to {a diner / church} on Monday evening. A / the waitress offended her there with a swear word.

27. Paige went to {Wal-Mart / a funeral} on Tuesday afternoon. A / the cashier watched her there intently.

28. Nora went to {a farm / group counseling} Thursday before dinner. A / the farmer informed her there about the community.

29. Irma went to {a hotel / a show} on Wednesday night. A / the bellboy flirted with her there shyly.

30. Jim went to {a recording studio / a demonstration} yesterday before dinner. A / the producer ignored him there the whole time.

31. Andrew went to {the train station / a nightclub} in late September. A / the train conductor conversed with him there for a short time.

32. Harold went to {court / a preschool} on Wednesday morning. A / the judge disagreed with him there openly.

Appendix C: Materials for Experiment 2

1. John will go to {the pool / the mall} this morning. Peter will go swimming {tomorrow / too} after he gets back from school.

2. Luke will go to {the mall / the pool} on Friday. Ann will go shopping {on Saturday / too} once she comes back from cello practice.

3. Frank will go to {the gym / the pub} on Monday afternoon. Lily will go to exercise {next week / too} when she feels better.

4. Lisa will go to {the pub / the gym} on Saturday night. Rachel will go drinking {next Saturday / too} after she gets out from work.

5. Fiona will go to {the pond / the club} on Monday evening. Monica will go fishing {on Saturday / too} when she won't have to work.

6. Tracy will go to {the club / the pond} on Tuesday afternoon. Francisco will go dancing {on Sunday / too} after his parents leave.

7. Amelia will go to {the park / the karaoke bar} on Tuesday evening. Josh will go for a walk {on Tuesday night / too} after he finishes his homework.

8. Dave will go to {the karaoke bar / the park} on Tuesday evening. Katherine will go singing {on Tuesday evening / too} before it gets dark outside.

9. Edward will go to {the mountains / the church} on Tuesday morning. Noah will go hiking {on Wednesday / too} before he goes to visit his parents.

10. Lenny will go to {the church / the mountains} on Tuesday afternoon. Jake will go to pray {on Friday / too} when he will have the whole day off.

11. Anthony will go to {the library / the rink} on Wednesday morning. Chris will go to study {in the afternoon / too} after he comes back from the mall.

12. William will go to {the rink / the library} on Wednesday morning. Loraine will go ice-skating {in the afternoon / too} before she goes to her yoga class.

13. Sophie will go to {the ATM / the barbershop} after breakfast. Adele will go to get money {later on / too} once she gets back from jogging.

14. Matthew will go to {the barbershop / the ATM} after breakfast. Mary will go to get a haircut {in the evening / too} after she finishes ironing her clothes.

15. Henry will go to {the beach / the casino} on Monday. Janet will go sunbathing {on Thursday / too} when she has more free time.

16. Donovan will go to {the casino / the beach} on Monday. Karla will go gambling {on Wednesday / too} before her vacation ends.

17. Troy will go to {the slopes / the archaeological site} on Thursday. Raul will go skiing {next week / too} after he recovers from the flu.

18. Lisa will go to {the archaeological site / the slopes} on Thursday. Bella will go to dig {on Tuesday morning / too} once she finishes the training.

19. David will go to {the boathouse / the demonstration} on Thursday afternoon. Patricia will go sailing {on Thursday evening / too} before she takes the train back home.

20. Thomas will go to {the demonstration / the boathouse} on Thursday afternoon. Linda will go protesting {on Thursday night / too} after her shift is over.

21. Eric will go to {the political rally / the golf course} on Friday. Barbara will go campaigning {on Monday / too} once her kids go to day camp.

22. Amy will go to {the golf course / the political rally} on Friday. Randy will go golfing {on Saturday / too} when his boss leaves for Hawaii.

23. Wayne will go to {the polling station / an internet cafe} in half an hour. Paula will go to vote {later on / too} when it stops raining.

24. Russell will go to {an internet cafe / the polling station} in half an hour. Lauren will go to check her email {in fifteen minutes / too} before she goes to class.

25. Tracy will go to {the art studio / the observatory} on Friday night. Beverly will go to paint {on Thursday night / too} once she comes back from the gym.

26. Alan will go to {the observatory / the art studio} on Friday night. Valerie will go star-gazing {on Monday night / too} when her new contact lenses will be ready.

27. Brian will go to {the winery / the skate park} on Sunday evening. Avril will go wine-tasting {next Friday / too} once her guests arrive.

28. Carlos will go to {the skate park / the winery} on Sunday evening. Angela will go skateboarding {next Saturday / too} before her brother comes home from camp.

29. Danny will go to {the spa / the office} on Monday morning. Chris will go to relax {on Tuesday / too} once he comes back from the city.

30. Vincent will go to {the office / the spa} on Monday morning. Leslie will go to work {on Monday afternoon / too} after she goes to the dentist.

31. Andrea will go to {the locker room / the temple} now. Elizabeth will go to change {later / too} before she goes home.

32. Ayden will go to {the temple / the locker room} now. Amanda will go to meditate {in half an hour / too} before the sun sets.