

Language production in a shared task: Cumulative Semantic Interference from self- and other-produced context words[☆]



Renske S. Hoedemaker^{a,*}, Jessica Ernst^b, Antje S. Meyer^a, Eva Belke^b

^a Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

^b Ruhr-Universität Bochum, Bochum, Germany

ARTICLE INFO

Article history:

Received 25 May 2016

Received in revised form 17 October 2016

Accepted 15 November 2016

Available online xxxx

Keywords:

Cumulative Semantic Interference

Language production

Language comprehension

Joint picture naming

Dialogue

Lexical selection

ABSTRACT

This study assessed the effects of semantic context in the form of self-produced and other-produced words on subsequent language production. Pairs of participants performed a joint picture naming task, taking turns while naming a continuous series of pictures. In the single-speaker version of this paradigm, naming latencies have been found to increase for successive presentations of exemplars from the same category, a phenomenon known as Cumulative Semantic Interference (CSI). As expected, the joint-naming task showed a within-speaker CSI effect, such that naming latencies increased as a function of the number of category exemplars named previously by the participant (self-produced items). Crucially, we also observed an across-speaker CSI effect, such that naming latencies slowed as a function of the number of category members named by the participant's task partner (other-produced items). The magnitude of the across-speaker CSI effect did not vary as a function of whether or not the listening participant could see the pictures their partner was naming. The observation of across-speaker CSI suggests that the effect originates at the conceptual level of the language system, as proposed by Belke's (2013) Conceptual Accumulation account. Whereas self-produced and other-produced words both resulted in a CSI effect on naming latencies, post-experiment free recall rates were higher for self-produced than other-produced items. Together, these results suggest that both speaking and listening result in implicit learning at the conceptual level of the language system but that these effects are independent of explicit learning as indicated by item recall.

© 2016 Elsevier B.V. All rights reserved.

Conversations typically center on a specific topic. As a result, speakers may repeatedly refer to different exemplars of a semantic category. For example, a conversation about a past vacation may include mentions of multiple types of vehicles, buildings or geographical features. Even though the repeated retrieval of same-category exemplars is likely to occur regularly during normal language use, it appears to be quite challenging from a procedural point of view. When speakers are asked to name a series of pictures containing multiple items from the same category, naming latencies slow down for each additional item that belongs to the same category as a previously named item, a phenomenon known as the Cumulative Semantic Interference effect (CSI; Brown, 1981; Howard, Nickels, Coltheart, & Cole-Virtue, 2006; see also Alario & del Prado Martín, 2010; Navarrete, Mahon, & Caramazza, 2010; Oppenheim, Dell, & Schwartz, 2010).

In the typical continuous CSI paradigm (Howard et al., 2006), participants name a series of pictures containing five items each from 24 different semantic categories plus a number of fillers. In this paradigm the CSI effect manifests as a 20 to 30 ms increase in naming latency for each additional member of a category that is named (Alario & del Prado Martín, 2010; Belke, 2013; Howard et al., 2006; Oppenheim et al., 2010; Navarrete et al., 2010; see also Rose & Abdel Rahman, 2016). Thus far the CSI effect has been found to emerge only in the context of word production tasks requiring name retrieval. When speakers name written words instead of pictures, the effect is absent (Navarrete et al., 2010; Belke, 2013). Moreover, when pictures are semantically categorized (e.g., as natural or man-made), repeated access to a semantic category results in cumulative *facilitation*, such that categorization times decrease each time an additional exemplar from the same category is presented (Belke, 2013; Riley, McMahan, & de Zubicaray, 2015).

The CSI effect demonstrates that language production is sensitive to linguistic context in the form of previously produced words. More specifically, the effect suggests that name retrieval for the purpose of language production leads to implicit learning within the language system, which affects subsequent retrieval operations (Belke, 2013; Howard et al., 2006; Oppenheim et al., 2010). Models developed to account for the CSI effect differ with regards to the level of language processing at which this retrieval-based implicit learning is proposed to

[☆] Author note: We thank Annika Hitschler and Thomas Dornbusch for programming the experiment setup and Martin Kitzinger for testing participants. We also thank Amie Fairs, Suzanne Jongman, Shiri Lev-Ari, Annelies van Wijngaarden, and members of the Psychology of Language department at the Max Planck Institute for Psycholinguistics for their help with this project.

* Corresponding author at: Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands.

E-mail address: renske.hoedemaker@mpi.nl (R.S. Hoedemaker).

occur. In the influential models by Howard et al. (2006) and Oppenheim et al. (2010), CSI emerges as a result of incremental changes in the connections between the conceptual and lexical levels of representation. In both models, target processing during picture naming results in conceptual activation of the target as well as semantically related items (Collins & Loftus, 1975; Levelt, Roelofs, & Meyer, 1999; McClelland & Rogers, 2003). In the model by Howard et al., naming a picture strengthens the conceptual-to-lexical connections for that item. As a result, previously named category coordinates compete more strongly for selection with a subsequent target name than non-previously named category coordinates, slowing down lexical selection. In the Oppenheim et al. model, naming a picture does not only strengthen the conceptual-to-lexical connections for that item but also weakens these connections for its (co-activated) category coordinates, which results in slower naming of these items on subsequent trials.

In contrast, Belke's (2013) Conceptual Accumulation account does not identify the CSI effect as originating in the conceptual-to-lexical connections. Instead, the account proposes that the incremental changes that give rise to the CSI effect are contained within the conceptual level of representation. The Conceptual Accumulation account is based on the Featural and Unitary Semantic Space (FUSS) model of lexical-semantic memory (Vigliocco, Vinson, Lewis, & Garret, 2004). The FUSS model consists of two distinct layers: a layer of conceptual features (e.g., has wings, can swim) and a layer of amodal, unitary "lexico-semantic representations" (e.g., duck; see Vigliocco et al., p. 433). The conceptual features are connected to the lexico-semantic representations in a many-to-one fashion, such that semantically similar concepts (such as items from the same category) share more features than semantically dissimilar concepts. The lexico-semantic representations are akin to the lexical concepts implemented in non-decompositional models of lexical-semantic encoding, such as WEAVER++ (Levelt et al., 1999; Roelofs, 1998; Roelofs, 2014). Belke (2013) drew upon the FUSS model and the WEAVER++ model in order to establish a working model of lexical-semantic encoding, including a fully specified level of conceptual representations and a mechanism for lexical selection. This model includes a layer of conceptual features, a layer of lexical concepts and a layer of lexical representations (Belke, 2013). Picture naming results in the co-activation of the target lexical concept and the lexical concepts of related items by virtue of their shared conceptual features (consistent with the characterization of shared activation in Oppenheim et al., 2010). According to the Conceptual Accumulation account (Belke, 2013), repeated access to a semantic category strengthens the connections between the lexical concepts that are being retrieved and their shared conceptual features, which results in the durable accumulation of activation within the conceptual level (Kroll & Stewart, 1994; Belke, Meyer, & Damian, 2005). For instance, when naming a duck, the links between its conceptual features and its lexical concept are strengthened. As a result, the lexical concept of duck and its associated lexical representation will be co-activated more easily when another item from the same semantic category is retrieved shortly afterwards. The same holds for other exemplars from the same category, so retrieving several items from the same semantic category results in high levels of activation within that category's section of the conceptual network. The observed effect of this increased conceptual activation varies depending on the task. During semantic categorization, increased activation in the appropriate section of the conceptual network will facilitate categorical (e.g., natural or man-made) responses, resulting in a cumulative facilitation effect as observed by Belke (2013) and Riley et al. (2015). By contrast, item naming requires the selection of a single entry at the level of lexical representations. Increased conceptual activation among a target and its category coordinates spreads down to the lexical level where it results in increased competition, causing the observed cumulative slowing of naming latencies.

Thus far, most semantic context effects including CSI have been studied within the domain of either comprehension or production, assessing the effects of prior production on subsequent production, or

of prior comprehension on subsequent comprehension. However, a sizeable portion of everyday language use occurs in the context of conversation, which involves interlocutors rapidly alternating between the production and comprehension of utterances (Pickering & Garrod, 2004). Consequently, the linguistic context preceding many spontaneous utterances comprises a mixture of self-produced and other-produced (i.e. perceived) speech. A small number of studies have investigated the possible transfer of CSI between comprehension and production tasks using printed word naming as the comprehension task. This approach has yielded mixed results. Vitkovitch and colleagues found that word naming interfered with subsequent naming of semantically related pictures (Vitkovitch & Cooper, 2012; Vitkovitch, Cooper-Pye, & Ali, 2010). In contrast, Navarrete and colleagues found that picture naming interfered with subsequent printed word naming, but not vice versa (Navarrete et al., 2010; Navarrete, Mahon, Lorenzoni, & Peressotti, 2016; see also Belke, 2013). Studying the role of the presence of a partner in a shared naming task, Kuhlen and Abdel Rahman (subm.) presented evidence demonstrating that CSI can operate across speakers. Specifically, they showed that the CSI effect was greater when task partners alternated naming category exemplars compared to when a single participant named every other exemplar in a category and the remaining exemplars were shown but not named. These findings suggest that listening to a partner naming objects affects language production in a similar fashion as naming the objects oneself. However, in Kuhlen's and Abdel Rahmans' study, all pictures were shown to both participants, who may have covertly named all objects. Thus, there are remaining questions regarding when and how implicit learning in the language system may occur across domains.

Although neither the models by Howard et al. (2006) or Oppenheim et al. (2010) nor the Conceptual Accumulation account (Belke, 2013) were developed to consider potential cross-domain CSI effects of prior comprehension on subsequent production, predictions can be derived from each model. According to the models by Howard et al. (2006) and Oppenheim et al. (2010), the CSI effect originates in the semantic-to-lexical connections. Specifically, the effect is modelled to arise as a consequence of the process of selecting a specific lexical representation among all the lexical entries co-activated by shared conceptual activation. Thus, the incremental changes in the production system that give rise to the CSI effect are a direct consequence of the challenges posed by the conceptual-to-lexical mapping requirement of language production. As a result, the effect is considered to arise only following a process of conceptual-to-lexical mapping, as is required for picture naming but not passive listening (Oppenheim et al., 2010). In contrast, the Conceptual Accumulation account proposes that the CSI effect results from changes in the connections within the conceptual level. As it is typically assumed that conceptual representations are shared between the production and comprehension systems (Levelt et al., 1999; Roelofs, 1992), this account predicts that prior comprehension and production of a picture name will result in a similar buildup of activation among the conceptual features and lexical concepts of the appropriate sections of the conceptual network, as also suggested by the cumulative facilitation observed for semantic categorization tasks (Belke, 2013). As a result, both comprehension and production are predicted to result in CSI during production of subsequent exemplars from the same category.

The current study investigated the effects of self- and other-produced semantic context words on subsequent picture naming using a joint (two-person) version of the continuous CSI paradigm. Pairs of participants named a series of pictures while seated side by side, alternating naming and listening turns within each category (see Fig. 1). We assessed the cumulative effects of self-produced context words (within-speaker interference) and other-produced context words (across-speaker interference). According to the model of CSI by Howard et al. (2006) and Oppenheim et al. (2010), other-produced exemplars do not involve semantic-to-lexical mapping for selection and therefore should not interfere with a speaker's subsequent naming of items from the same category,

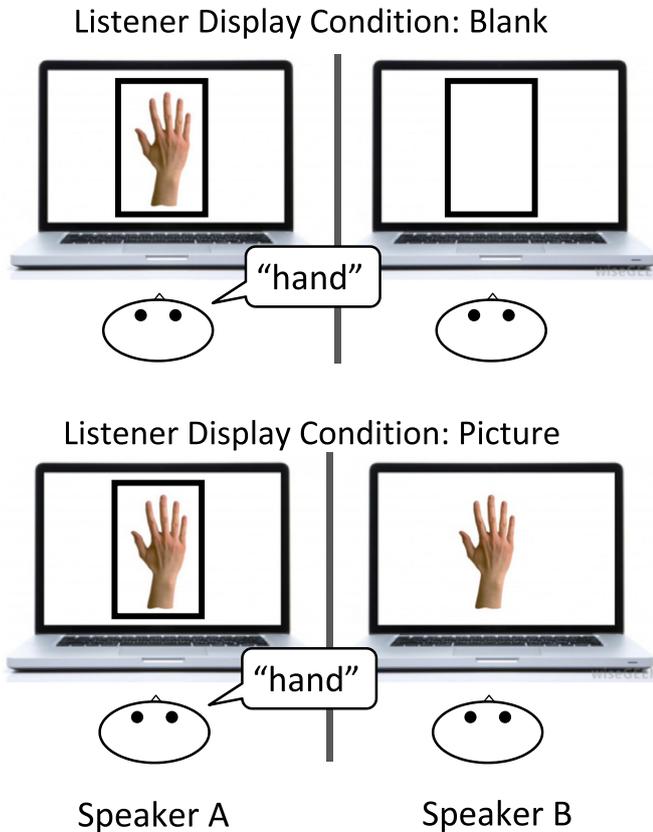


Fig. 1. Illustration of the experimental setup and listener-display conditions. During the experiment, pairs of participants were separated by a screen so that they could see and hear each other, but they could not see each other's display. In the Blank Listener Display condition (top panel), each picture was presented only to the participant whose turn it was to name the item, while the listening participant was shown a blank rectangle. In the Picture Listener Display condition (bottom panel), each picture was shown to both participants, and a black rectangle around the picture indicated whose turn it was to name the item.

yielding within-speaker CSI but not across-speaker CSI. In contrast, the Conceptual Accumulation account (Belke, 2013) predicts both within- and across-speaker CSI, as the effect is thought to originate within the conceptual level of representation. Of course, under some circumstances listeners might engage in the same type of conceptual-to-lexical mapping as the speaker by predicting their partner's utterance or even covertly naming their partner's item. To investigate the importance of these processes for potential across-speaker CSI we manipulated whether participants could see the picture their partner was about to name. If across-speaker CSI depends on prediction or covert naming we should only observe the effect when the to-be-named picture is presented to both the speaker and the listener. In contrast, if across-speaker CSI can arise from listening alone, as predicted by the Conceptual Accumulation account (Belke, 2013), we should observe across-speaker CSI even when the to-be-named pictures are presented only to the speaker and not to the listener, making it impossible for the listener to predict their partner's upcoming utterance or covertly name their partner's item. In addition, we explored the effects of producing a name and listening to a name on subsequent recall memory. Pictures have a well-established memory advantage over words (Paivio, 1971) and naming words or pictures out loud typically results in better memory for these items than for items that were read silently or named by a partner (e.g., MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; MacLeod,

2011). Therefore, we expect to find better recall memory for self-compared to other-produced items and better memory for other-produced items when listeners could also see the to-be-named picture compared to when they could not see them. Importantly, the memory task will allow us to compare the extent to which these factors affect implicit learning as indicated by the CSI effect to their effect on explicit learning as indicated by post-experiment free recall.

1. Method

1.1. Participants

Twenty-four participant pairs (native speakers of German, mean age: 23.16 years, $SD = 2.44$ years) were tested. They were paid for their participation. One pair was replaced due to technical difficulties. Each pair consisted of one male and one female participant.

1.2. Materials and design

165 color photographs of objects, animals and people that were also used in Belke (2013) were used, comprising 24 semantic categories with five items each and 45 filler items. The items were arranged in twenty-four lists according to the design by Howard et al. (2006) and identical to those used in Belke (2013). All items appeared once in each list. Category exemplars were presented with a lag of 2, 4, 6, or 8 intervening items and all four possible lags were used within each category. Within a list, all 24 possible within-category lag orders appeared once (2-4-6-8, 4-2-6-8, etc.) and across lists all categories appeared once with each lag order. Once the order of lags and categories in each list was determined, the exemplars were inserted into their respective category's 'slots' in the list in a pseudorandom order, such that, across lists, they appeared at least four and at most five times in each position.

The naming task was assigned to each pair of participants such that for half of the categories, one member of the participant pair (participant A) named the first, third and fifth category exemplar presented whereas the other member (participant B) named the second and fourth category exemplar. For the other half of the categories, the assignment was reversed so that participant B named the first, third and fifth exemplar and participant A named the second and fourth exemplar. When participant A named the item, participant B simply listened and vice versa. Participant pairs thus alternated between self-named and other-named items within a category, but due to the varying number of intervening items between exemplars, the order of turns in the sequence of the list appeared to be random. Individual categories always appeared with the same naming assignment (ABABA vs BABAB). Because individual items were assigned pseudorandomly to category slots, the specific category exemplars named by participants A and B varied across lists. Filler items were consistently assigned so that each member of the pair named half of the filler items.

1.3. Apparatus and procedure

The experimental setup is depicted in Fig. 1. Pairs of participants were tested together. Stimuli were presented on two SR EyeLink 1000 systems connected by parallel printer cable. The stimulus presentation and communication between the two systems were controlled by Experiment Builder (SR Research). During the experiment Participant A and B sat side by side, each viewing their own ViewSonic G22f 19" display monitor. The pair was separated by a screen so participants could see and hear their task partner, but they could not see the other person's display. The stimuli were sized to 500×500 pixels, corresponding to 14.7×14.7 cm or $24^\circ \times 24^\circ$ of visual angle, and presented at the center of the display. Vocal responses were recorded using a Sony ECM-MS907 microphone.

Pairs participated in four experimental sessions administered on two days, allowing us to present all manipulations of interest in a within-subject design but ensuring the duration of each experimental sitting remained manageable.¹ Throughout all four of the sessions, assignment of participants as A or B was kept the same within the pair. During each session, a pair ran through one experimental list, meaning each picture was presented once. On each day, a pair completed two sessions, each with a different list. On the second day, the same two lists were used as on the first day. For the first session of the first day, participants were instructed simply to name the object that was presented during their turn. Before beginning the second session of the day, they were informed that the experiment would be followed by a recall memory test, and instructed to pay careful attention to all items, regardless of whether they or their partner were asked to name the item. After finishing the second session of the day, participants were asked to write down all object names they could remember during a recall memory task. At the end of Day 1, no preview was given to participants on the tasks awaiting them on Day 2. When they returned on Day 2, they were again first asked to name the object that was presented during their turn. After that, the memory task was re-introduced, and participants continued with the fourth session, which was followed by the recall memory task.

It is possible that informing participants of the upcoming memory task before the second session of each day motivated them to attend more carefully to their partner's utterances. If the cross-speaker CSI effect depended on the listener's attention to the other-produced items, we would expect to find greater across-speaker interference in the second as compared to the first session of each day (i.e., after the participants were made aware or reminded of the upcoming memory task). Of course, based on their experience on Day 1, participants may have anticipated the upcoming memory task during both sessions on Day 2. Regardless, as will be described below, the naming latency patterns were very similar across the four sessions and showed no evidence that anticipation of the memory test affected how the participants processed their partners' utterances.

Over the course of the two days, each pair completed two different versions of the experiment with regards to the listener's display (see Fig. 1). In the Blank Listener Display version, each picture was presented only to the person whose turn it was to name the item, while the other participant saw an empty black frame. In the Picture Listener Display version, each picture was presented to both participants, and turns were indicated by a black frame around the picture of the speaking participant on that trial. On a testing day, both sessions were presented in the same listener-display version. Half the participant pairs performed the Blank Listener Display version on Day 1 and the Picture Listener Display version on Day 2, and for the other participant pairs this order was reversed.

1.4. Data analysis

Trials on which none or both of the speakers responded (2.2%) and trials that contained hesitations or false starts (1.8%) were coded as errors and excluded from the analysis of naming latencies, as were trials on which the speaker responded with an incorrect name (3.5%). Trials on which participants used an acceptable alternative name (e.g., 'Schloss' (castle) instead of 'Burg' (fortress)) were included in the analyses (cf. Howard et al., 2006). Response latencies on correct critical items were measured from picture onset using Praat (Boersma & Weenink, 2016). Valid response times that deviated more than two standard deviations from the mean response latency for a particular position in each listener-display version (4.4% of correct critical trials) were excluded from the data. In accordance with previous studies

¹ Prior research (e.g. Navarrete et al., 2010) has shown that in a single-speaker setup, administering the continuous paradigm with the same items repeatedly in short succession yields substantial repetition priming effects but these do not interact with the CSI effect.

using this design, analyses were conducted with subjects and categories (rather than items) as random factors.

To facilitate comparison of the CSI effect in the current two-speaker situation to a single-speaker situation, our initial analyses mirror those reported by Howard et al. (2006). A repeated-measures ANOVA assessed the effects of ordinal position and lag over positions two through five (position one is excluded from this analysis because 'lag' does not apply to the first-presented exemplar of a category). If there is a CSI effect, we should see a continuous increase in naming latencies for each additional item in a category. As the CSI effect has been shown to be long-lasting, we should see no interaction of position and lag and lag should not have an effect on naming times. Accordingly, the CSI should remain stable when collapsing across lags and including all five positions in the analyses of the effect of ordinal position.

To assess the effects of within- and across-speaker interference separately, we first coded the items in each category with respect to the number of items produced by the speaker (round). Within a given category, one speaker named three objects (ordinal positions 1, 3, and 5) and the other speaker named two objects (ordinal positions 2 and 4). For reasons that we will introduce shortly, we excluded ordinal position 5 from the analyses, which left us with two rounds per speaker to name an object. Within each round, we coded whether a speaker was the first or the second to speak (turn). Accordingly, within-speaker interference was operationalized as the extent to which the average naming latency slowed down across rounds. Across-speaker interference was operationalized as the extent to which naming latencies slowed down across turns within a round. This way, the effect of turn is measured while controlling for round, allowing us to assess within- and across-speaker CSI as separate effects. Because only one item per category functions as the third self-named exemplar (ordinal position 5), we could not assess the across-speaker interference effect for this position, and items in this position were excluded from the analyses below. The full coding scheme is presented in Table 1.

2. Results

Fig. 2 displays the average naming latencies for correct responses to category exemplars by position. On average, response latency slowed down by 10 ms with every category exemplar that was named in a session. The initial analysis across ordinal positions two to five (cf. Howard et al., 2006) showed a linear effect of position, with responses slowing down as a function of position, $F_1(1,47) = 37.12, p < 0.001, \eta_p^2 = 0.44$, $F_2(1,23) = 23.03, p < 0.001, \eta_p^2 = 0.50$. There was no effect of lag, $F_s < 1$, and no position by lag interaction, $F_s < 1$. A separate ANOVA collapsing across lag and assessing the effect of ordinal position across all five positions also showed a significant linear increase of response latency as a function of position, $F_1(1,47) = 88.90, p < 0.001, \eta_p^2 = 0.65$,

Table 1

Coding of rounds and turns as a function of ordinal position across category types (ABABA vs BABAB). Within-speaker interference is measured as the effect of round on naming latencies. To assess across-speaker interference, turns were coded within rounds. The contrast coding represents the coding scheme in the statistical models.

	Ordinal position	1	2	3	4	5
ABABA categories	Speaker	A	B	A	B	A
	Round (within-speaker CSI)	1	1	2	2	(3)
	Contrast coding	-1/2	-1/2	1/2	1/2	-
	Turn (across-speaker CSI)	1	2	1	2	-
	Contrast coding	-1/2	1/2	-1/2	1/2	-
BABAB categories	Speaker	B	A	B	A	B
	Round (within-speaker CSI)	1	1	2	2	(3)
	Contrast coding	-1/2	-1/2	1/2	1/2	-
	Turn (across-speaker CSI)	1	2	1	2	-
	Contrast coding	-1/2	1/2	-1/2	1/2	-

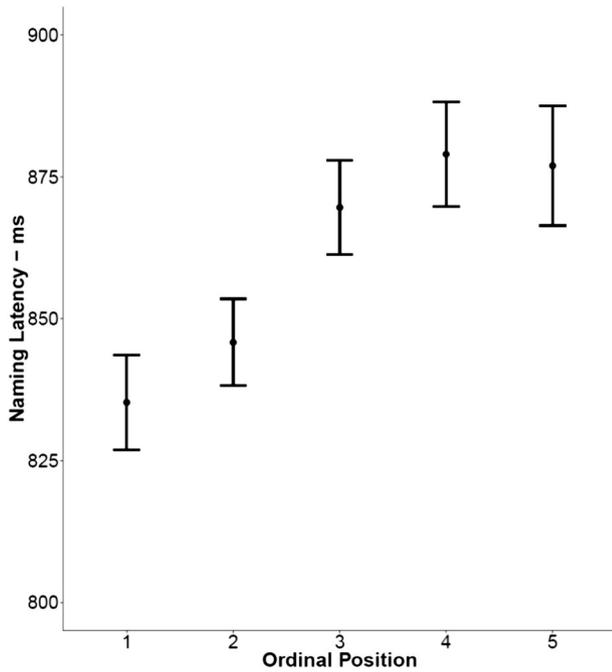


Fig. 2. Mean naming latencies across position, collapsed across listener-display version and lag. Error bars represent 95% within-subject confidence intervals (Morey, 2008, see also Cousineau, 2005; Loftus & Masson, 1994).

$F_2(1,23) = 49.4, p < 0.001, \eta_p^2 = 0.68$. These findings correspond to what Howard et al. had reported for the single-speaker case.

Average error rate per participant per list was 7%. Consistent with previous studies of semantic interference in the continuous naming paradigm, error rates across positions 2 to 5 did not vary as a function of position, $F_1(1,47) = 1.3, p = 0.26, F_2(1,23) = 2.7, p = 0.12$ or lag, $F_s < 1$, and there was no position by lag interaction, $F_s < 1$. Collapsed across lag, there was no effect of position on error rates, $F_s < 1$.

2.1. Within- and across-speaker CSI: naming latency

Consistent with previous studies (Belke, 2013; Howard et al., 2006; Oppenheim et al., 2010) we did not observe an effect of lag on naming latencies. Therefore, the data were collapsed across lag for all following analyses. Naming latencies by listener-display version and ordinal position (labelled by round and turn) are presented in Fig. 3. Linear mixed-effects models (Baayen, Davidson, & Bates, 2008) with subjects and categories as crossed random effects as implemented in the lme4 library (Bates, Maechler, Bolker, & Walker, 2015) for R (R Development Core Team, 2012) were used to predict log-transformed naming latencies. The maximal random effect structure supported by the data across all models included random intercepts for subjects and categories, random by-subject slopes for the effects of round (within-speaker CSI), turn (across-speaker CSI) and listener-display version, and random by-category slopes for the effects of turn (across-speaker CSI) and listener-display version. Fixed effects were considered significant if $|t| \geq 2$. We first ran a control model with the following fixed effect predictors: trial number within a list (mean centered), participant role (A vs B), category type (ABABA vs BABAB) and session of the experiment within a participant (coded as an integer from zero to three to represent the 1st through 4th session completed by each participant). Only session emerged as a significant fixed effect, $B = -0.03, SE = 0.003, t = -11.2$, indicating an overall decrease in response latencies as participants completed more sessions (see Navarrete et al., 2010). In all following models, the main effect of session was included as a fixed effect.

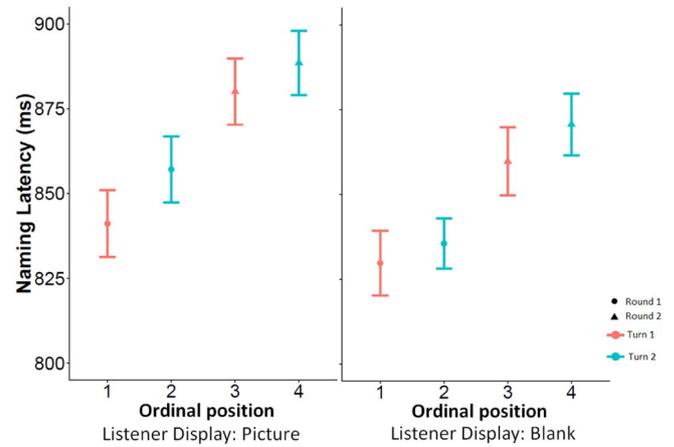


Fig. 3. Naming latencies as a function of ordinal position within the category and listener display. Round refers to the number of exemplars named within each round. Within-speaker interference was operationalized as the effect of round, i.e. comparing the average naming latency in ordinal positions One and Two to the average naming latency in positions Three and Four. Across-speaker interference was operationalized as the effect of turn, i.e., the naming latency difference between ordinal positions One and Two and between ordinal positions Three and Four. Data are collapsed across categories, so that each position contains data from categories assigned in the ABABA and BABAB naming order. Error bars represent 95% within-subject confidence intervals, normalized within each listener display condition (Morey, 2008, see also Cousineau, 2005; Loftus & Masson, 1994).

In these models, the effect of session was significant but did not interact significantly with other variables.

The full model included the following fixed-effect predictors of interest: round (within-speaker CSI), turn (across-speaker CSI), listener-display version, and the interactions of listener-display version with both within- and across-speaker CSI. All three fixed effect predictors of interest were treated as nominal variables with two levels, and were contrast coded (see Table 1). The results of the model's fixed effects are displayed in Table 2.

The results show that naming latencies slowed down significantly as a function of round. Responses to items in round 2 (ordinal positions Three and Four) were an average of 34 ms slower compared to responses in round 1 (ordinal positions One and Two), indicating a within-speaker CSI effect. In addition, naming within a round slowed down significantly as a function of turn. On average, latencies for items in ordinal position Two and Four were 10 ms slower than latencies for items in ordinal position One and Three, indicating an across-speaker CSI effect. There was a significant effect of listener display, such that average naming latencies were 18 ms slower when the listener was shown the speaking participant's picture compared to when the listener's display was blank. However, the magnitude of the within- and

Table 2

Results of the mixed-effects model assessing the effects of round (within-speaker CSI), turn within round (across-speaker CSI) and listener display for log-transformed picture naming latencies. Significant effects are indicated with an asterisk.

Fixed effect	Estimate	Standard error	t value
Intercept	6.735	0.017	385.0
Round (within-speaker CSI)	0.037	0.004	9.2*
Turn (across-speaker CSI)	0.013	0.005	2.4*
Listener display (blank vs picture)	0.020	0.008	2.6*
Session	-0.031	0.003	-11.5*
Listener display * within-speaker CSI	0.003	0.008	0.3
Listener display * across-speaker CSI	0.005	0.008	0.6

across-speaker interference effects did not vary as a function of listener display. When the fixed effect of trial number was added to the model, its effect was not significant and none of the other fixed effects changed indicating that the CSI effect cannot be explained as resulting from fatigue or other types of general slowing over the course of the experiment.

Supplementary analyses assessed the effect of memory task on naming latencies. We observed a main effect of memory task, such that naming latencies were faster during the second compared to the first session of each day, $B = -0.037$, $SE = 0.004$, $t = -9.7$. Note, however, that the memory task was always introduced after the first and before the second session of a day, meaning that memory task was directly correlated with session number. Neither memory task (two levels, coded within each testing day) nor session number (four levels, coded across testing days, see above) interacted significantly with the within- or the across-speaker CSI effect on naming latencies, $|ts| < 1.5$.

2.2. Analysis of within- and across-speaker CSI: errors

Response accuracy was analyzed using a multi-level binary logistic regression model as implemented in the lme4 package (Bates et al., 2015) in R (using the BOBYQA optimizer, Powell, 2009). The full model supported by the design included random intercepts for subjects and categories, and random slopes for the effects of round (within-speaker CSI), turn (across-speaker CSI) and listener display (blank vs picture) by subjects and by categories. All effects were contrast coded according to the same coding scheme as the analysis of naming latencies. Results of the model are presented in Table 3. Accuracy did not vary as a function of round or turn. The effects of the other predictors were consistent with the naming latency results. Error rates decreased significantly across sessions (from 10.4% in session 1 to 6.3%, 5.6% and 4.8% in sessions 2 to 4, respectively) and there were more errors when the listener display contained a picture ($M = 7.8\%$) compared to when it was blank ($M = 5.8\%$). This occurred because listeners in the picture display condition occasionally initiated a naming response for items assigned to be named by their partner. Such an error could only occur in the picture display condition but not in the blank display condition, accounting for the slightly higher error rate in this condition.

2.3. Free recall performance

As described, all category exemplars appeared once per session and the memory test was administered after the two sessions of that day had been completed. Exemplars were distributed pseudorandomly within the slots assigned to each category within a list, so that on a particular testing day a participant may have self-named a particular exemplar once (in either the first or second session, here referred to as self-other and other-self items respectively), twice (self-self items) or never (other-other items).

Table 3

Results of the multi-level binary logistic regression model assessing the effects of round (within-speaker CSI), turn (across-speaker CSI) and listener display on response accuracy. Significant effects are indicated with an asterisk.

Fixed effect	Estimate	Standard error	z value
Intercept	-3.030	0.167	-18.12
Round (within-speaker CSI)	-0.061	0.133	-0.46
Turn (across-speaker CSI)	0.002	0.135	0.01
Listener display (blank vs picture)	0.367	0.124	2.97*
Session	-0.301	0.043	-6.96*
Listener display * within-speaker CSI	0.068	0.179	0.38
Listener display * across-speaker CSI	-0.123	0.181	-0.68

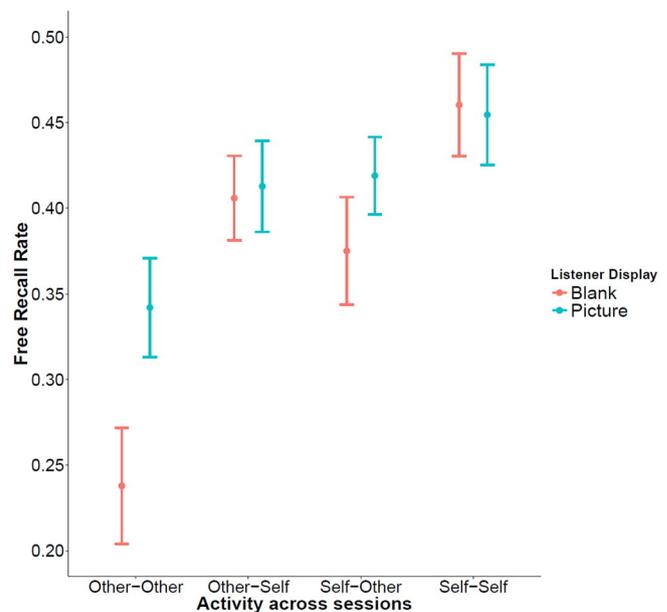


Fig. 4. Average recall rate across subjects as a function of listener display and the type of activity performed on an item on the first and second iteration within the experimental session. Other-other: the item was other-produced (named by the participant's partner) on both iterations. Other-self: the item was other-produced on the first iteration and self-produced on the second iteration. Self-other: the item was self-produced on the first iteration and other-produced on the second iteration. Self-self: the item was self-produced on both iterations in the session. Error bars represent 95% within-subject confidence intervals, (Morey, 2008, see also Cousineau, 2005; Loftus & Masson, 1994).

Fig. 4 shows the performance on the free recall task as a function of listener display, which was kept constant across both sessions of a day, and participant activity in the two sessions preceding the memory test (self-named vs other-named), collapsed across testing days. The effect of activity on free recall of an item was assessed in a multi-level binary logistic regression model as implemented in the lme4 package (Bates et al., 2015) in R (using the BOBYQA optimizer; Powell, 2009). Data for one participant were excluded from these analyses due to an error in the administration of the free recall task. The maximal random effect structure supported by the data included random intercepts for subjects and items and random by-subject slopes for the effect of listener display. The fixed effects tested in the model were the participant activity for an item on each session on that testing day as a categorical predictor coded as a factor with four levels (other-other, other-self, self-other, self-self) and listener display as a categorical predictor with two levels (blank or picture, contrast coded) as well the interaction of listener display and speaker activity.

Across listener-display versions, participant activity had a significant effect on recall rates such that self-named items were remembered better than other-named items. Items that were self-named at least once were recalled significantly better than items that were never self-named (comparing other-other items to self-other items: $B = 0.571$, $SE = 0.061$, $z = 9.36$, and other-other items to other-self items: $B = 0.620$, $SE = 0.061$, $z = 10.18$), but there was no difference between items that were self-named during the first session (self-other) and items self-named during the second session (other-self), $B = -0.049$, $SE = 0.058$, $z = -0.85$. Participants showed the highest recall rates for items they had named twice (self-self items compared to other-other items: $B = 0.843$, $SE = 0.060$, $z = 14.04$, self-self items compared to self-other items: $B = 0.273$, $SE = 0.058$, $z = 4.71$, and self-self items compared to other-self items: $B = 0.224$, $SE = 0.058$, $z = 3.87$).

The effect of listener display on recall memory was assessed separately for each participant activity. Broadly, other-named items were

recalled better when the listener display contained a picture compared to when the listener display was blank. That is, the presence of a picture on the listener display significantly improved recall memory for other-other items, $B = 0.625$, $SE = 0.126$, $z = 4.97$, and self-other items, $B = 0.218$, $SE = 0.086$, $z = 2.53$, and this facilitatory effect of picture presence was significantly larger for the other-other items than the self-other items, $B = -0.376$, $SE = 0.121$, $z = 3.12$. However, for items that were self-named during the second session (other-self items), the presence of a picture for the listener on the first item did not significantly improve recall, $B = 0.039$, $SE = 0.081$, $z = 0.48$ (To achieve convergence, the by-subject slope for the effect of listener display was not included for this subset analysis). As expected, listener display did not affect recall of items that were self-named during both sessions (self-self items), $B = -0.025$, $SE = 0.097$, $z = -0.26$, as the speaker was necessarily presented with a picture of every self-named item regardless of the listener display condition. Supplementary analyses also showed that recall rates did not vary as a function of ordinal position within the category. These results confirmed earlier pilot data from a single-speaker version of the continuous CSI task.

3. Discussion

The current study used a joint continuous picture naming task to assess the effects of semantic context in the form of self-produced and other-produced words on subsequent language production. In the classic single-speaker version of the continuous naming paradigm naming latencies increase for successive presentations of exemplars from the same category, a phenomenon known as Cumulative Semantic Interference (CSI, Alario & del Prado Martín, 2010; Belke, 2013; Howard et al., 2006; Navarrete et al., 2010; Oppenheim et al., 2010). In the current study, the continuous naming task was performed jointly by pairs of participants, and items were assigned such that the pair alternated naming turns within a category (e.g., Speaker A named the first exemplar, Speaker B the second exemplar, etc.). As expected, performance on the joint naming task showed a within-speaker CSI effect, such that naming latencies increased following category exemplars named previously by the participant. Crucially, we also observed an across-speaker CSI effect, such that naming latencies slowed following category members named by the participant's task partner. This finding is consistent with results from Kuhlen and Abdel Rahman (subm.), who observed more within-speaker CSI when every other category exemplar was named by a partner compared to when every other category exemplar was presented as a picture but not named by either partner. Importantly, the magnitude of the across-speaker CSI effect did not depend on whether or not the listening participant could see the pictures their partner was about to name.

To estimate the relative magnitude of the within- and across-speaker CSI effect, we compared naming latencies in the current study to those in Belke (2013, Experiment 1), which included a single-speaker version of the continuous naming task using the same stimuli in the same language (German). In Belke's experiment, naming latencies increased by an estimated average of 20.7 ms for each additional category exemplar whereas in the current experiment average increase per exemplar was 10.4 ms. Taken together, these results suggest that CSI can occur following both within and across speakers, but it appears that the across-speaker CSI effect is smaller than the within-speaker CSI effect. Similarly, Belke (2013) and Riley et al. (2015) reported cumulative facilitation effects in semantic classification that were numerically smaller (in absolute magnitude) than the interference effect typically observed for object naming. Arguably, the magnitude of the interference effect is linked to the depth of conceptual processing (cf. Kroll & Stewart, 1994), which might differ between object naming on the one hand and auditory comprehension or semantic classification on the other. Future research will have to establish whether there is a systematic difference between comprehension-induced and production-induced cumulative semantic effects.

The Conceptual Accumulation account (Belke, 2013) localizes the origin of the CSI effect at the conceptual level. When speakers access the same semantic category repeatedly, incremental learning causes activation to accumulate in a small section of the conceptual system. This activation accumulation feeds forward to the lexical level, where it hampers the selection of one lexical entry from among the strongly co-activated competitors. Conceptual representations are commonly assumed to be shared between production and comprehension, so a straightforward prediction from the Conceptual Accumulation account is that the CSI effect arises following production as well as in comprehension of category coordinates. This is the result we obtained, suggesting that the CSI effect reflects incremental learning that takes place at the conceptual level. This interpretation is consistent with the finding that CSI occurs across languages when tested within bilingual speakers (Runnqvist, Strijkers, Alario, & Costa, 2012; but see Castro, Strijkers, Costa, & Alario, 2008).

As discussed, the models proposed by Howard et al. (2006) and Oppenheim et al. (2010) propose that CSI results from incremental changes in the conceptual-to-lexical mappings. This generates the prediction that the CSI effect only occurs following operations that require such mapping, such as picture naming but not printed word naming or listening to someone else naming a picture. However, these models could account for the observed across-speaker CSI effect if listeners engaged in covert naming of the other-produced items. As specified by Oppenheim et al., learning within the model is error-based and does not depend on whether or not the correct name is actually selected and produced, meaning the effect is also expected to occur following erroneous or omitted responses. Consequently, the effect may also emerge when item names are activated but not overtly pronounced. However, on this view the across-speaker CSI effect should only be seen when the picture is also presented on the listener's display. This is because covert naming or anticipation of a picture name is only possible when the listening participant can see the picture their partner is about to name. Contrary to this prediction, no systematic effect of the presence of the picture was observed (see also Kleinman, Runnqvist, and Ferreira (2015) for effects of prediction on CSI). Of course, participants could covertly repeat their partner's utterances even in the absence of the picture. However, it is unclear to what extent this process involves the conceptual-to-lexical mapping considered necessary for CSI, as without the picture listeners would not have access to conceptual information until after the lexical information had been processed.

Barring an effect of covert naming, there may be other ways to adjust the models by Howard et al. (2006) and Oppenheim et al. (2010) to account for the current findings, for example by implementing a comprehension mechanism that allows lexical representations to activate associated conceptual features. However, as their current accounts of the CSI effect is based specifically on the process of conceptual-to-lexical mapping, the model would require further specification regarding the type of incremental learning following the reverse, lexical-to-conceptual, mapping and the way it might result in cumulative interference during subsequent picture naming (see Belke, 2013, p. 253, for discussion of a proposal by Gary Oppenheim of an expansion of their model to cover the semantic classification task). Unlike the models by Howard et al. (2006) and Oppenheim et al. (2010), the Conceptual Accumulation account (Belke, 2013) has not been computationally implemented. It is based on two implemented models of semantic memory and lexical retrieval, but neither of these models incorporates incremental learning, leaving the mechanisms proposed in the Conceptual Accumulation account less specified than in the other models. At the same time, the Conceptual Accumulation account is, as yet, the only account that can accommodate the present and all previous findings on cumulative semantic context effects.

A computational model that could account for both the within- and across-speaker CSI effects, whether taking the Howard et al. (2006) and Oppenheim et al. (2010) models or the Conceptual Accumulation account as its starting point, would benefit from a more detailed

understanding of how and under which circumstances the across-speaker CSI effect arises. An important aspect of such a model may be the matter of what is considered 'relevant experience' to the production system. As described by Oppenheim et al., the semantic interference effect accumulates with relevant experience and it is insensitive to irrelevant experience. Thus far, only naming of related items had been shown to constitute relevant experience, whereas naming unrelated items, non-verbal semantic processing of objects, word naming and unrelated filler activities (see Damian & Als, 2005) have all been demonstrated to constitute irrelevant experience in the sense that they do not result in cumulative interference. On the other hand, Abdel Rahman and Melinger (2011) demonstrated that speakers can establish associative relations between seemingly unrelated entities when their relatedness is made explicit (shovel, handkerchiefs, flowers, suit, and coffee are categorically unrelated but associatively related to a funeral). Their findings suggest that relevance can be contextually induced. Importantly, the across-speaker CSI effect observed in the current study suggests that processing of other-produced items should be considered relevant experience to the production system (see also Kuhlen and Abdel Rahman, *subm.*).

The current study showed a dissociation between explicit learning, observed in the form of recall memory, and implicit learning as signified by CSI. Whereas self-produced and other-produced items both resulted in a CSI effect on naming latencies, self-produced items were remembered better than other-produced items, and other-produced items for which the listener had seen the picture were recalled better than other-produced items during which the listener had seen a blank display. These results are consistent with the production effect (Conway & Gathercole, 1987; MacLeod et al., 2010; Gathercole & Conway, 1988) and the picture-superiority effect (Paivio, Rogers, & Smythe, 1968; Paivio, 1971). The picture superiority effect refers to the observation that items presented as pictures are typically remembered better than items presented as (written) words. The production effect refers to the finding that words that are read out loud or are quietly (but overtly) mouthed are remembered better than words that are read covertly to oneself (Conway & Gathercole, 1987; Gathercole & Conway, 1988; MacLeod et al., 2010), or are read out loud by someone else (MacLeod, 2011). Both the production and picture-superiority effect can be explained according to the item's distinctiveness during the study phase. Pictures and self-produced items are hypothesized to leave a more distinct memory trace during encoding than words and not-produced or other-produced items, which makes them more likely to be recalled (Nelson, Reed, & Walling, 1976; Weldon & Coyote, 1996). According to this interpretation, the current study showed an effect of item distinctiveness on explicit memory but not implicit learning, whereas the number of times a semantic category was accessed affected implicit learning but not explicit memory. In the context of the Conceptual Accumulation account (Belke, 2013), this means that the incremental changes to conceptual representations induced by repeated access to the same semantic category affect the ease with which specific names may be retrieved during naming, but do not affect a specific item's distinctiveness in memory.

This work demonstrates that Cumulative Semantic Interference can operate across speakers, such that hearing a partner name an item can interfere with one's own subsequent production of a semantically related name. These findings are most consistent with the assumption that the CSI originates at the conceptual level, as proposed by the Conceptual Accumulation account (Belke, 2013). Nonetheless, the Howard et al. (2006) and Oppenheim et al. (2010) models might account for the across-speaker CSI if it is assumed that listeners covertly imitated their partner's utterances and this process involved semantic-to-lexical mapping, or through adjustments to the models that allowed similar incremental changes to the production system following self- and other-produced words. Implicit learning as measured by the CSI effect was largely dissociated from explicit learning as measured by item recall, suggesting that cumulative interference in the production system did

not affect individual items' distinctiveness in memory. In sum, the current study suggests that speaking and listening result in similar implicit learning at the semantic level of the language system as indicated by naming latencies, but these implicit learning effects are independent from explicit learning effects as indicated by item recall.

References

- Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 149–161.
- Alario, F.-X., & del Prado Martín, F. M. (2010). On the origin of the "cumulative semantic inhibition" effect. *Memory & Cognition*, 38, 57–66.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Belke, E. (2013). Long-lasting semantic context effects on object naming are necessarily conceptually mediated. *Journal of Memory and Language*, 69, 228–256.
- Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology*, 58(4), 667–692.
- Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer [computer program]. Version 6.0.17. retrieved 21 April 2016 from <http://www.praat.org/>
- Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 204–215.
- Castro, Y. G., Strijckers, K., Costa, A., & Alario, X. (2008). When cat competes with dog but not with perro: Evidence from the semantic competitor paradigm. *Poster presented at the Fifth International Workshop on Language Production*, Annapolis, MD.
- Collins, A. M., & Loftus, E. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428.
- Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, 26(3), 341–361.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45.
- Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1372–1384.
- Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory and Cognition*, 16(2), 110–119.
- Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, 100, 464–482.
- Kleinman, D., Runnqvist, E., & Ferreira, V. S. (2015). Single-word predictions of upcoming language during comprehension: Evidence from the cumulative semantic interference task. *Cognitive Psychology*, 79, 68–101.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- Kuhlen, A.K., & Abdel Rahman, R. (subm.) Having a partner affects lexical retrieval: Spoken word production in shared task settings.
- Levitt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1(4), 476–490.
- MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin and Review*, 18(6), 1197–1202.
- MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36(3), 671–685.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4(4), 310–322.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, 4(2), 62–64.
- Navarrete, E., Mahon, B. Z., & Caramazza (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica*, 134(3), 279–289.
- Navarrete, E., Mahon, B. Z., Lorenzoni, A., & Peressotti, F. (2016). What can written-words tell us about lexical retrieval in speech production? *Frontiers in Psychology*, 6.
- Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 523–528.
- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, 114, 227–262.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, 11(4), 137–138.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02), 169–190.
- Powell, M. J. D. (2009). The BOBYQA algorithm for bound optimization without derivatives. *Technical report DAMTP NA2009/06*. Department of Applied Mathematics and Theoretical Physics, Cambridge University: Cambridge, UK.
- R Development Core Team (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (ISBN 3-900051-07-0).

- Riley, E., McMahon, K. L., & de Zubicaray, G. (2015). Long-lasting semantic interference effects in object naming are not necessarily conceptually mediated. *Frontiers in Psychology*, 6, 578.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 904–921.
- Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex*, 59, 33–48.
- Rose, S. B., & Abdel Rahman, R. (2016). Cumulative semantic interference for associative relations in language production. *Cognition*, 152, 20–31.
- Runnqvist, E., Strijkers, K., Alario, F. X., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language*, 66(4), 850–869.
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garret, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48, 422–488.
- Vitkovitch, M., & Cooper, E. (2012). My word! Interference from reading object names implies a role for competition during picture name retrieval. *The Quarterly Journal of Experimental Psychology*, 65(6), 1229–1240.
- Vitkovitch, M., Cooper-Pye, E., & Ali, L. (2010). The long and the short of it! Naming a set of prime words before a set of related picture targets at two different intertrial intervals. *The European Journal of Cognitive Psychology*, 22, 161–171.
- Weldon, M. S., & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 670.