



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Representation of others' synchronous and asynchronous sentences interferes with sentence production

Citation for published version:

Gambi, C, Van de Cavey, J & Pickering, MJ 2022, 'Representation of others' synchronous and asynchronous sentences interferes with sentence production', *Quarterly Journal of Experimental Psychology*. <https://doi.org/10.1177/17470218221080766>

Digital Object Identifier (DOI):

[10.1177/17470218221080766](https://doi.org/10.1177/17470218221080766)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Quarterly Journal of Experimental Psychology

Publisher Rights Statement:

The final version of this paper has been published in Quarterly Journal of Experimental Psychology, 2022 by SAGE Publications Ltd, All rights reserved. © Gambi Et al, 2022. It is available at: <https://doi.org/10.1177%2F17470218221080766>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Representation of others' synchronous and asynchronous sentences interferes with sentence production

| | |
|-------------------------------|--|
| Journal: | <i>Quarterly Journal of Experimental Psychology</i> |
| Manuscript ID | QJE-STD-21-261.R1 |
| Manuscript Type: | Standard Article |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Gambi, Chiara; Cardiff University, Van de Cavey, Joris; Ghent University and Karus vzw Pickering, Martin; The University of Edinburgh School of Philosophy Psychology and Language Sciences |
| Keywords: | sentence production, picture description, co-representation, joint action, imagination |
| | |

SCHOLARONE™
Manuscripts

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Representation of others' synchronous and asynchronous sentences interferes with
sentence production.

Chiara Gambi

University of Edinburgh and Cardiff University

Joris Van de Cavey

Ghent University and Karus vzw, Belgium

Martin J. Pickering

University of Edinburgh

Address for correspondence:

Chiara Gambi

School of Psychology

70, Park Place

Cardiff University

CF10 3AT Cardiff, U.K.

Email: GambiC@cardiff.ac.uk

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Abstract

In dialogue, people represent each other's utterances in order to take turns and communicate successfully. In previous work [Gambi, C., Van de Cavey, J., & Pickering, M. J. (2015). *Interference in joint picture naming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(1)*, 1-21.], speakers who were naming single pictures or picture pairs represented whether another speaker was engaged in the same task (versus a different or no task) concurrently, but did not represent in detail the content of the other speaker's utterance. Here, we investigate co-representation of whole sentences. In three experiments, pairs of speakers imagined each other producing active or passive descriptions of transitive events. Speakers took longer to begin speaking when they believed their partner was also preparing to speak, compared to when they did not. Interference occurred when speakers believed their partners were preparing to speak at the same time as them (synchronous production and co-representation; Experiment 1), and also when speakers believed that their partner would speak only after them (asynchronous production and co-representation; Experiments 2a and 2b). However, interference was generally no greater when speakers believed their partner was preparing a different compared to a similar utterance, providing no consistent evidence that speakers represented what their partners were preparing to say. Taken together, these findings indicate that speakers can represent another's intention to speak even as they are themselves preparing to speak, but that such representation tends to lack detail.

Keywords: sentence production; picture description; imagination; joint task; co-representation

Representation of others' synchronous and asynchronous sentences interferes with

RUNNING HEAD: JOINT SENTENCE PRODUCTION

sentence production

Introduction

The norm in conversation is for speakers to take turns (Sacks, Schegloff, & Jefferson, 1974) with remarkably precise timing (Stivers et al., 2009). This turn-taking system has deep evolutionary roots (Levinson, 2016), emerges early in childhood (Hilbrink, Gattis, & Levinson, 2015), and capitalizes on neural mechanisms of prediction (e.g., Magyari, Bastiaansen, De Ruiter & Levinson, 2014) to ensure smooth coordination between speakers (Levinson, 2016). But speakers can also produce language simultaneously, for example in special ritualised forms of choric speech (e.g., in singing, chanting, or praying; Cummins, 2009; Jasmin et al., 2016).

Whether they speak simultaneously or consecutively, speakers coordinate with one another: The timing and content of their utterances is dependent on the timing and content of their partner's utterances. For example, speakers typically slow down during choric speech (Cummins, 2003, 2009), perhaps to facilitate synchronization. During conversation, they predict what their partner will say so they can begin preparing an appropriate response ahead of time (Levinson, 2016). These observations illustrate how simultaneous and consecutive speaking involve different coordination demands, which are likely to be served by different coordination mechanisms. But importantly, they also show that, both when speaking simultaneously and when speaking consecutively, speakers represent their partner's use of language. In this paper, we ask how speakers represent others' utterances: Do they represent *whether* others are preparing to speak, *what* they will say, and *when* they will say it?

Co-representation of actions and utterances

Co-representation accounts of joint action propose that co-acting individuals build representations of their partner's tasks and actions which are in the same format as representations of their own task and actions (Knoblich, Butterfill, & Sebanz, 2011; Sebanz, Knoblich, & Prinz,

RUNNING HEAD: JOINT SENTENCE PRODUCTION

2003, 2005). Such co-representation can support the coordination of actions in time and space between individuals because it means others' actions can directly affect their own: The shared representational format allows "cross-talk" between representations of their own and others' actions. To illustrate, participants in one study (Schmitz, Vesper, Sebanz, & Knoblich, 2017) had to synchronize the landing times of their reaching movements; when one of the participants encountered an obstacle that delayed her landing time, the other participant also tended to move as if there were an obstacle in his path, suggesting he represented his partner's action. Similarly, when participants were tasked with completing a sequence of two actions, they were slower if they knew their partner was simultaneously performing the same two actions in a different order; remarkably, this was the case even if partners were not asked to synchronize their landing times, suggesting that co-actors may co-represent each others' actions even without an explicit task goal to do so (Schmitz, Vesper, Sebanz, & Knoblich, 2018).

Importantly, there is evidence of cross-talk between self- and other-representations for utterances as well as for actions. Baus et al. (2014) had participants take turns naming pictures with a confederate, while EEG was recorded. There was a larger P200 component when the participant prepared to produce a low-frequency compared to a high-frequency word, which suggests this component indexes difficulty of lexical retrieval. Crucially, there was a larger P300 component when the participant stayed silent and the confederate prepared to produce a low-frequency compared to a high-frequency word, but not when the confederate stayed silent as well. Despite the delayed latency, the authors interpreted this other-related frequency effect on the P300 component as suggesting that participants were sensitive to the frequency manipulation even when their partner was preparing to name the picture (but they themselves were not), and that this was because they partly retrieved words their partner was about to produce.

Similarly, in a joint Stroop task in which one participant responded to color words by saying "red" when they were printed in red, and the other responded to them by saying "green" when they

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 were printed in green, there was a larger P300 component when participants stayed silent and their
4 partner prepared to respond (joint experiment), compared to when they both stayed silent (solo
5 experiment; Demiral, Gambi, Nieuwland, & Pickering, 2016). While it is possible the enhanced
6 P300 is due to greater attention to stimuli that the partner responded to, it may also be that speakers
7 mapped their partner's stimuli to their partner's response even though they did not have to respond
8 themselves. Accordingly, the study also found indication of reduced perceptual conflict (i.e.,
9 conflict between the ink color and the meaning of the color words, when they mismatched; e.g., the
10 word *green* printed in red ink) in the joint compared to solo experiment, which suggests that
11 speakers were less likely to treat the other color as conflicting when it was their partner's response
12 than when it was not.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 Finally, Kuhlen and Abdel Rahman (2017) and Hoedemaker, Ernst, Meyer, and Belke
28 (2017) showed that listening to another speaker name pictures delays naming times for pictures
29 from the same semantic category, similarly to naming those pictures oneself. In other words, there
30 appears to be a between-speaker cumulative semantic inhibition effect (Brown, 1981). The within-
31 speaker version of this effect is thought to result from changes in the strength of connections
32 between conceptual and lexical representations, or between features and concepts, which are the
33 result of previous retrieval episodes (Belke, 2013; Howard, Nickels, Coltheart, & Cole-Virtue,
34 2006; Oppenheim, Dell, & Schwartz, 2010). Thus, the between-speaker cumulative semantic
35 inhibition effect may suggest that other-produced utterances have a similar effect on production
36 representations as self-produced utterances. Evidence that a between-speaker effect occurs even
37 when speakers cannot hear their partner, but are merely told that their partner is naming the pictures
38 (Experiments 2 and 3; Kuhlen & Abdel Rahman, 2017), supports this interpretation. But it should
39 be noted that in a more recent attempt to replicate this finding (Kuhlen & Abdel Rahman, 2021), the
40 authors found no evidence for between-speaker cumulative semantic inhibition when the partner
41 was not present in the same room as the participant.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 In sum, a number of findings suggest that speakers represent another speaker's utterances
4 before comprehending them and, sometimes at least, even when they believe they will not be able
5 to hear them (Gambi & Pickering, 2016). These findings suggest that the representations speakers
6 form of others' utterances are similar to the representations they form of their own utterances when
7 they are preparing to speak. This evidence is therefore consistent with co-representation of
8 utterances (Gambi & Pickering, 2011, 2013).
9
10
11
12
13
14
15
16

Representing another's utterances while simultaneously preparing to speak

17
18
19
20
21 In these studies, speakers took turns with their (present or imagined) co-speakers, which
22 means they never had to represent their partner *while* preparing to speak themselves. In contrast, in
23 previous work we manipulated speakers' beliefs about the task their partner was performing in a
24 different room (so they could never hear their partner) while they prepared to name pictures
25 themselves (Gambi, Van de Cavey, & Pickering, 2015). Across four experiments, speakers were
26 told their partner would produce an utterance which was the same or different from their own
27 utterance, stay silent, or respond *yes* or *no* to a semantic categorization question. Naming latencies
28 were faster when participants believed their partner was not speaking, or was speaking but not
29 engaging in lexical retrieval (categorization condition), than when they believed their partner was
30 retrieving lexical items. This finding that merely imagining another speaker producing an utterance
31 affects concurrent production is particularly striking, because these speakers represented that their
32 partner's production system was simultaneously activated even though the task did not require them
33 to pay attention to their partner's task.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 Importantly, Brehm, Taschenberger, and Meyer (2019) found a similar result to Gambi et al.
52 (2015) using a simpler design. In two experiments, they had participants name single pictures that
53 were presented either on the right- or the left-hand side of the screen, while their partner (who was
54 sitting next to them) either named the picture presented on the other side of the screen or
55 categorised it (participants wore noise-cancelling headphones so they could not hear their partner);
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 the two pictures could either be the same or different, so that participants were either
4
5 producing/categorizing the same utterance as their partner or a different one. Experiment 1 also
6
7 included an individual condition in which the participant performed the same task alone (they were
8
9 told their partner had failed to show up for the experiment). Crucially, participants took longer to
10
11 name the pictures when they knew their partner was present and was performing a different task
12
13 (categorization) than the same task (naming). Note that this study reverses the pattern found by
14
15 Gambi et al. (2015), who instead showed *shorter* latencies when participants believed their partner
16
17 was categorizing pictures rather than naming them (see Brehm et al., 2019 for a discussion of
18
19 methodological differences that may have led to this reversal of the effect). Crucially, however, the
20
21 study indicates that speakers represented their partner's task and that these representations affected
22
23 their own production.
24
25
26
27
28

29
30 However, in both Gambi et al. (2015) and Brehm et al. (2019) there was no consistent
31
32 evidence that speakers represented the content of their partner's utterances. In Gambi et al. (2015),
33
34 naming latencies were no slower when speakers believed their partner was preparing to produce a
35
36 different utterance compared to the same utterance. While there was a tendency in some of the
37
38 experiments for naming errors to be more common when participants believed their partner was
39
40 producing a different utterance, this tendency was not consistent across experiments. Similarly, in
41
42 Brehm et al.'s (2019) Experiment 1, latencies were longer when the participant knew the partner
43
44 was naming (but not categorising) a different picture, but this was the case even when no actual
45
46 partner was present, and moreover the effect was not replicated in Experiment 2. Interestingly,
47
48 participants tended to look more often at their partner's picture when they knew they were naming
49
50 rather than categorizing, and also when the partner's picture was different to their own, but overall
51
52 there were very few looks to partner's pictures, suggesting that speakers did not regularly engage in
53
54 detailed representation of their partner's utterances.
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 In sum, the findings of Gambi et al. (2015) and Brehm et al. (2019) are consistent with other
4 work (Baus et al., 2014; Demiral et al., 2016; Kuhlen & Abdel Rahman, 2017; Hoedemaker et al.,
5 2017) in showing that speakers represent their partners' act of production in a similar way to how
6 they represent their own act of production. However, they also suggest that the content of others'
7 utterances (i.e., what they are saying) may not be represented: there was only inconsistent evidence
8 of increased interference when producing a different utterance from one's partner, compared to
9 producing the same utterance. Specifically, this finding contrasts with Baus et al. (2014) and
10 Kuhlen and Adbel Rahman (2017), who showed that people are affected by linguistic properties of
11 words that their partners have produced or are about to produce, suggesting that people can
12 represent detailed aspects of the content of others' utterances (frequency, semantic category). It also
13 contrasts with evidence for co-representation of other's actions, such as the finding that people
14 execute movement sequences less efficiently when they know their partner is executing the same
15 sequence in reverse order versus the same order (Schmitz et al., 2018).

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34 Does simultaneous production affect the degree of representation of another's utterances?
35
36

37 One explanation of these divergent findings is that language production is a demanding
38 cognitive process (Roelofs & Piai, 2011). Thus, speakers may have had less resources to allocate to
39 representing their partner's utterance when they were simultaneously preparing to speak
40 themselves, resulting in less detailed co-representation – that is, limited to co-representation of
41 *whether another is speaking but not of what they are saying* – in Gambi et al. (2015) and Brehm et
42 al. (2019) compared to the other studies. Similarly, Hoedemaker and Meyer (2018) also found only
43 limited evidence that speakers represented their partner's utterance when they had to speak
44 themselves. They had speakers name one picture, then listen to their partner name a second picture,
45 and finally name a third picture themselves. Speakers looked at the second picture more often and
46 earlier when their partner named it than when nobody named it but, importantly, looks to the second
47 picture were much longer and earlier when participants named it themselves, indicating that co-
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 representation stopped well short of planning the partner's utterances. Although in Hoedemaker and
4
5 Meyer (2018) participants took turns speaking – similarly to studies that found evidence for detailed
6
7 co-representation of utterance content (Baus et al., 2014; Kuhlen & Abdel Rahman, 2017) – picture
8
9 triplets were presented on the screen simultaneously, which likely encouraged speakers to plan the
10
11 third picture name while their partner was naming the second picture. Simultaneous planning may
12
13 have taken resources away from speakers' representation of their partner's utterance, **making them**
14
15 **less likely to represent its content in detail**. In fact, in other conditions speakers may have not
16
17 represented their partner's act of production at all: When speakers were asked to name the first of
18
19 two pictures, they were as fast to begin naming the first picture whether they knew their partner
20
21 later named the second picture or they knew their partner would remain silent (whereas they were
22
23 much slower when they themselves had to name the second picture as well).

24
25
26
27
28
29 In sum, speakers sometimes form detailed representations of **the content of** their partner's
30
31 utterances, but whether they do so may depend on available resources. For example, if either
32
33 representing another's utterance or producing their own utterance is particularly demanding,
34
35 speakers may not represent their partner's utterances in any detail (or indeed, they may not
36
37 represent their partner at all). Conversely, if either co-representation or production (or both) are
38
39 made easier, speakers may be able to form more detailed representation of their partner's utterances
40
41 even if they are speaking simultaneously. For example, Brehm et al. (2019) showed that co-
42
43 representation can have a facilitatory effect on production when the task of representing another
44
45 speaker is made easier: In their study, beliefs about the partner's task were manipulated across
46
47 blocks and they found shorter latencies when participants believed their partner was naming
48
49 pictures (like them) than when they believed their partner was categorizing pictures. In contrast,
50
51 Gambi et al. (2015) manipulated beliefs on a trial-by-trial basis (and used a more difficult picture
52
53 naming task) and found longer latencies when speakers believed their partner was naming (like
54
55 them) compared to categorizing. This may suggest that, as task difficulty increases, co-
56
57 representation turns from a potentially facilitatory factor to a source of interference with concurrent
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 language production, making speakers less likely to engage with it. However, even with a much
4
5 simpler task, Brehm et al. still found little evidence of detailed representation of the content of the
6
7 partner's utterances, so it remains unclear whether limited resources are the only reason why some
8
9 studies show little or no evidence for detailed co-representation.
10
11

12 13 The current study 14 15

16 In the current study, instead of simplifying the task, we opted to make it harder. Naming
17
18 pictures is cognitively demanding, but planning whole sentences is even more so (Ferreira, 1991).
19
20 Thus, speakers may not be able to represent others' utterances at all when they are asked to
21
22 concurrently plan a whole sentence. Previous work has used comparatively simple production tasks,
23
24 where speakers produce either single words, or pairs of words (Baus et al., 2014; Brehm et al.,
25
26 2019; Demiral et al., 2016; Gambi et al., 2015; Hoedemaker et al., 2017; Hoedemaker & Meyer,
27
28 2018; Kuhlen & Abdel Rahman, 2017). In contrast, we asked speakers to describe pictured events
29
30 using full transitive sentences (e.g., *The nun follows the doctor*). Producing sentences involves
31
32 additional planning stages (Bock & Levelt, 1994) that are not involved in producing single words
33
34 (or even ordered pairs of words). Specifically, thematic roles (agent and patient) must be assigned
35
36 to syntactic positions (subject and object) so that an appropriate syntactic structure can be chosen
37
38 (Ferreira, 1994; Segaert, Wheeldon, & Hagoort, 2016). Importantly, transitive events can be
39
40 described using an active structure (as in the example above), or a passive structure (as in *The*
41
42 *doctor is followed by the nun*). To ensure that speakers focused on such additional planning stages,
43
44 they produced active and passive sentences, and sentence voice was varied on on a trial-by-trial
45
46 basis.
47
48
49
50
51

52
53 It is possible that the complex task of preparing to produce a whole utterance may entirely
54
55 prevent speakers from representing their partner's act of production. If so, production latencies
56
57 should be unaffected by whether speakers believe their partner is also preparing to produce a
58
59 sentence, or simply remaining silent. Such a finding would in turn suggest a limited role for other-
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 representations in utterance coordination, because such representations would not be available to
4
5 interlocutors who are either speaking (e.g., in choric speech) or preparing to speak (e.g., when they
6
7 are about to take a turn in dialogue).
8
9

10
11 If, however, speakers represent their partner's act of production even while performing a
12
13 demanding production task, it would suggest that they can routinely represent another speaker's act
14
15 of production while concurrently preparing to speak. If so, we expected slower production latencies
16
17 when speakers believed their partner was engaging in language production, compared to when they
18
19 believed their partner was silent, similarly to our previous research using words and word pairs
20
21 (Gambi et al., 2015). We also compared production latencies when speakers believed their partner
22
23 was producing a sentence in a different voice, to when speakers believed their partner was
24
25 producing an utterance in the same voice. Even if increasing task difficulty does not make
26
27 representation of others' utterances less likely, previous work (Gambi et al., 2015; Brehem et al.,
28
29 2019) suggests such representations are unlikely to be very detailed. If so, production latencies
30
31 should not be affected by whether speakers believe their partner is producing an utterance in the
32
33 same or a different voice, even if they are affected by whether their partner is preparing to produce
34
35 an utterance at all. This is because the voice of the partner's utterance would matter only if speakers
36
37 formed a representation of the partner's utterance that was detailed enough to distinguish between a
38
39 passive and an active utterance.
40
41
42
43
44
45

46
47 In addition, planning a passive sentence may be more difficult than planning an active
48
49 sentence – for example, Ferreira (1994) reported that passive sentences can take up to 1000-1400ms
50
51 longer to formulate than corresponding active sentences in certain conditions. If this is the case, it is
52
53 also possible that the degree to which participants represent their partner's utterances will vary
54
55 depending on whether they themselves are planning a passive (i.e., more difficult) or active (less
56
57 difficult) utterance. This would be reflected in an interaction between our manipulation of partner's
58
59 task and sentence voice. However, other studies have found that passive sentences are planned as
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 quickly as active sentences, for example in the no-syntactic repetition condition of Experiment 2 of
4
5 Segaert et al. (2016) and the no-word repetition condition of Segaert, Menenti, Weber, and Hagoort
6
7 (2011). So it is also possible that there will be no voice-related differences in planning latencies in
8
9
10 our study.

11
12
13 Finally, we have assumed that when speakers believe their partner will speak at the same
14
15 time as them, they experience interference because production and representation of others'
16
17 utterances are cognitively demanding processes that compete for resources *and* take place
18
19 synchronously. But what if these processes are asynchronous, such as when speakers believe their
20
21 partner will speak after them? It is possible that no interference would be observed in this situation,
22
23 as speakers may delay representing their partner's utterance until after they have prepared their own
24
25 utterance. In accord with this possibility, Hoedemaker and Meyer (2018) showed that participants
26
27 did not begin speaking later when they knew their partner would speak after them than when they
28
29 knew nobody would do so.
30
31
32

33
34
35 But given the evidence that speakers' representations of others' utterances often lack detail,
36
37 it is also possible that speakers represent only that their partner is intending to perform an act of
38
39 production, but not the timing of this act. For example, speakers may represent their partner's
40
41 intention to speak without representing *when* their partner will actually plan an utterance. This
42
43 would be compatible with evidence that generating an intention to speak and planning an utterance
44
45 are separate components of the production process (Levelt, 1989). For example, there is a
46
47 dissociation between processes determining what to say and processes determining when to speak
48
49 during turn-taking (Corps, Crossley, Gambi, & Pickering, 2018). Further, lexical access is faster
50
51 when there is a conscious intention to speak compared to when this is absent (Strijkers, Holcomb, &
52
53 Costa, 2011). In sum, if this account is correct, it will make no difference whether speakers believe
54
55 their partner intends to speak at the same time as them or after them – in both cases, they should
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 experience interference compared to conditions in which they believe the partner does not intend to
4
5 speak.
6
7

8 In conclusion, previous work has shown that speakers can represent others' utterances and
9 that doing so affects their own production, but there is mixed evidence regarding the ease with
10 which these representations are formed and the degree of detail involved. Relevant factors may be
11 the ease of speakers' own production and the extent to which preparation of speakers' own
12 utterances and representation of others' utterances overlap in time. In this study we explored how
13 these factors may affect the degree of interference from representations of others' utterances on
14 production.
15
16
17
18
19
20
21
22
23
24

25 In all experiments, participants produced full sentences to increase the difficulty of the
26 production task compared to previous work. In Experiment 1 we had speakers believe they were
27 speaking simultaneously (as in Gambi et al., 2015), whereas in Experiment 2 (2a and 2b) we had
28 speakers believe their partner would speak after them. Importantly, participants sat in separate
29 soundproof booths and were not able to hear one another at any point during either experiment.
30 Thus, in Experiment 2a we led participants to believe their partner would speak after them, even
31 though they in fact spoke simultaneously, which allowed us to collect comparable data from both
32 speakers in a pair. In Experiment 2b, instead, we had one randomly-selected participant actually
33 speak after their partner; although this meant we could use data from only half of the participants
34 (those who spoke first), it provided a replication of Experiment 2a that did not rely on creating a
35 false belief in the participants.
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 51 **Experiment 1: Synchronous production and representation of another's utterance**

52 In Experiment 1, either pairs of participants described pictures simultaneously or one
53 speaker described a picture while the other remained silent (as in Gambi et al., 2015, Experiment 1).
54 Thus, production and representation (of other utterances) were synchronous. Speakers produced
55
56
57
58
59
60

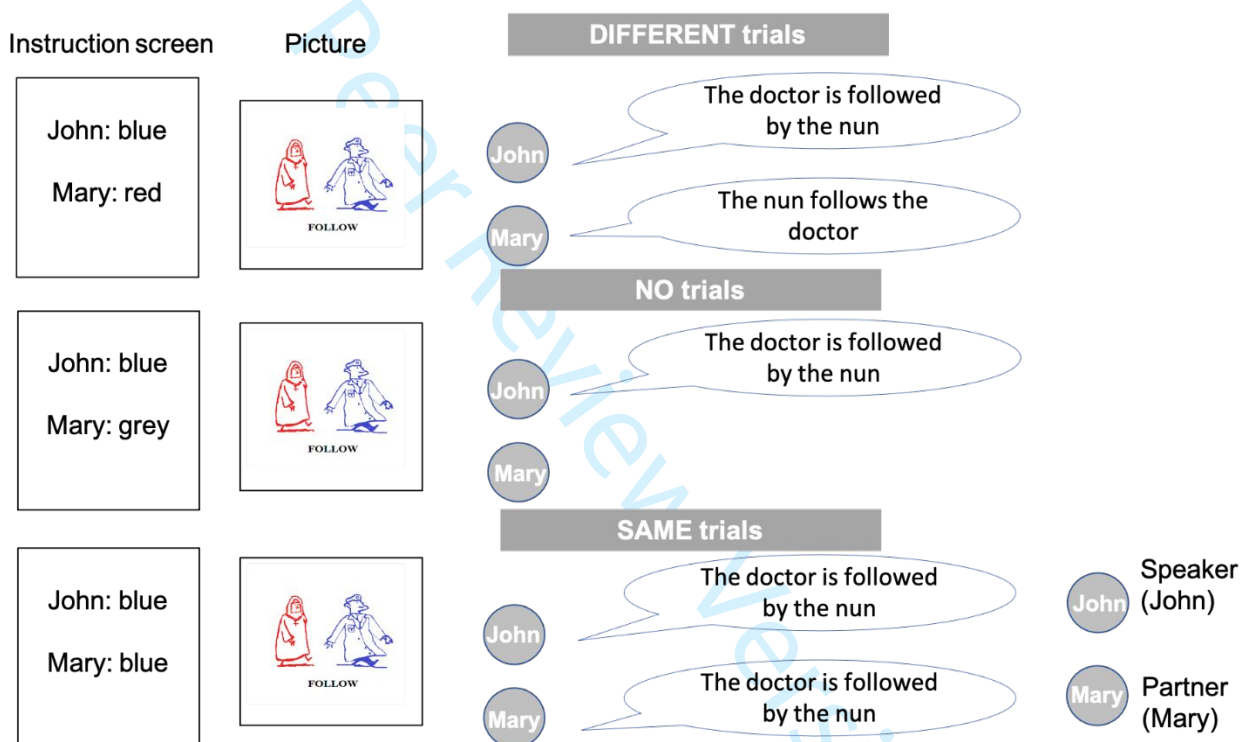
RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 short descriptions of transitive events (e.g., of a nun following a doctor, in Figure 1) in either the
4
5 active (e.g., *The nun follows the doctor*) or the passive voice (e.g., *The doctor is followed by the*
6
7 *nun*). We controlled these descriptions using a variant of the so-called “traffic-light” paradigm
8
9 (e.g., Menenti, Gierhan, Segaert, & Hagoort, 2011), in which participants were assigned a color on
10
11 each trial, and were instructed to first describe the character who appeared in that color. For
12
13 example, in Figure 1, top row, John has been assigned the color blue; since the picture displays the
14
15 patient of the action in blue, John will begin his description with the patient, and will thus produce a
16
17 passive sentence. In contrast, Mary has been assigned the color red, so she will start her description
18
19 with the agent and produce an active sentence.
20
21
22
23

24
25 Crucially, on every trial, the instruction screen also alerted participants to what their partner
26
27 was to do once the picture appeared, so we could manipulate what speakers believed their partner
28
29 was doing (Partner’s Task). There were three types of trials (Figure 1): On *Different* trials (top
30
31 row), participants were assigned different colors (John: blue, Mary: red or John: red, Mary: blue),
32
33 so that they described the picture starting with different characters, and produced sentences of
34
35 different voice; on *No* trials (middle row), the speaker described the picture starting with either
36
37 character while their partner remained silent (as indicated by the color grey; John: blue – Mary:
38
39 grey, John: red – Mary: grey, or John: grey, Mary: blue, John: grey, Mary: red); finally, on *Same*
40
41 trials (bottom row), participants were assigned the same color word (e.g., John: red – Mary: red or
42
43 John: blue – Mary: blue) and thus described the picture starting with the same character, producing
44
45 a sentence in the same voice as each other (both actives or both passives).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Figure 1. Schematic representation of the design used in Experiment 1, 2a, and 2b. Participants saw their actual names on the instruction screen (*John* and *Mary* are used here for illustrative purposes). In this example, John produces a passive description because he has been assigned the color blue and the blue character is the patient in the picture. John is the “speaker” and Mary is the “partner” because John is the one speaking on No trials in this example, but on other trials their roles were reversed (i.e., we collected data from both participants in all conditions).



If speakers who are preparing to produce a whole sentence do not concurrently represent their partner's act of production, speech onset latencies should be unaffected by our Partner's Task manipulation: Latencies should be similar across all three types of trials (*Different*, *No*, and *Same*). But if speakers represent their partner's act of production, we expect to find a joint interference effect (similar to Gambi et al., 2015): Speakers should take longer to begin their utterances when they believe their partner is also speaking (*Different* and *Same*), compared to when they believe their partner is silent (*No*).

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Further, if speakers merely represent that their partner is producing, then latencies should be unaffected by whether speakers believe their partner is producing the same utterance as themselves, or a different utterance. But if speakers represent the content of their partner's act of production, then we would also expect a joint interference effect that depended on whether the speakers' utterance is the same as their partner's (presumed) utterance or not. Presumably, there would be greater interference when self- and other-representations conflicted, as speakers would find the incompatible representation of their partner's utterance interfering with their own act of production. Specifically, speakers should have more difficulty producing an active if they believed their partner was producing a passive than an active, and more difficulty producing a passive if they believed their partner was producing an active than a passive (i.e., latencies should be longer on Different than Same trials, irrespective of Voice).

Finally, we know that sentential planning is incremental; specifically, when speakers are preparing to produce simple sentences containing two nouns (e.g., *the arrow is next to the bag*), they appear to retrieve both lemmas before beginning to speak, but also that they retrieve the phonological representation for the second noun only *after* speech onset (Meyer, 1996; Smith & Wheeldon, 2004). If so, it may be that evidence for interference due to representing the other act of production can be measured *after* as well as (or instead of) before speech onset. Therefore, we examined not only speech onset latencies, but also utterance durations.

Method

Participants. Gambi et al. (2015) found reliable effects with sample sizes varying between 24 and 32 participants, but since the current experiment included fewer trials per participant (192 vs. 300), we increased sample size and recruited 40 native English-speaking participants from the student population of the University of Edinburgh, who were assigned to 20 pairs. Participants were paired with each other on the basis of availability and no other criteria were followed when pairing participants, except that members of a pair did not know each other before taking part in the

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 experiment. All procedures were approved by the Ethics Committee of the Department of
4
5 Psychology, University of Edinburgh and participants gave informed written consent to
6
7 participating. They were paid £6 or given course credit for taking part. Data from one participant
8
9 were discarded prior to analysis because their average speech onset latency (1,681 ms) fell more
10
11 than 2 standard deviations above the average onset latency across all participants (1,164 ms); data
12
13 for their partner were retained since they could not have been aware of the other participant's
14
15 unusually delayed responding.
16
17

18
19
20 **Materials and Design.** We created 32 pictures depicting simple scenes comprising two human
21
22 characters and a printed verb. The pictures were created by combining 8 different transitive verbs
23
24 (*follow, chase, hit, punch, shoot, kill, tickle, touch*), each repeated 4 times, and 14 characters (*nun,*
25
26 *doctor, boxer, cowboy, robber, pirate, swimmer, waitress, ballerina, painter, monk, sailor, soldier,*
27
28 *clown*), each repeated a variable number of times. Across repetitions of the same verb, we
29
30 counterbalanced both the position of the agent (left or right) and the color of the agent (blue or red);
31
32 the patient was always depicted in the other color.
33
34
35
36

37 The design included 6 conditions obtained by crossing Partner's Task (Same, Different, No)
38
39 with Voice (Active vs. Passive); we manipulated both Partner's Task and Voice by assigning
40
41 participants to color words on each trial using an instruction screen (see above and Figure 1). Note
42
43 that in order to collect data from each participant in each condition, there were two additional
44
45 "dummy" sets of trials ("dummy" because they provided no data for the participant in question), in
46
47 which the speaker (John in Figure 1) remained silent but their partner (Mary in Figure 1) described
48
49 the picture instead, in either the Passive or Active voice. Participants encountered each of the 32
50
51 pictures once in each of these 8 sets of trials (i.e., 6 conditions + 2 "dummy" sets of trials).
52
53
54
55

56 **Procedure.** Participants were tested in adjacent soundproof rooms. The experimenter controlled
57
58 stimulus presentation using E-Prime (Version 2.0) from a separate room. Stimulus presentation was
59
60 simultaneous on each participant's screen. Participants could not hear each other during the

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 experiment. There was a window between the participants' rooms (so that they were aware of their
4
5 partner) but this window was perpendicular to the participants' line of vision when they faced their
6
7 monitors.
8
9

10
11 Participants were first introduced to one another and the experimenter entered their first
12
13 names in the program so they could be displayed on the instruction screens. They were then taken
14
15 into their booths and practiced naming the characters they would later encounter during the
16
17 experiment. They did this at the same time and without hearing one another. When a participant
18
19 made a mistake, the experimenter provided corrective feedback to that participant (but both
20
21 participants could hear the feedback). Participants were instructed how to describe pictures of
22
23 transitive events (through examples) based on the color cue that was assigned to them. These
24
25 instructions were delivered to both participants at the same time in the control room. The
26
27 experimenter emphasized the "joint" nature of the task, noting that participants would now work
28
29 together, and drawing attention to how they would be doing the same or different tasks as each
30
31 other, depending on the instruction screen. The participants then went back to their respective
32
33 booths to practice, before beginning the first of four experimental blocks (practice and experimental
34
35 trials were similar, but practice trials used a different set of pictures). In total, participants
36
37 completed 256 experimental trials (i.e., 32 pictures, each repeated 8 times); within each block there
38
39 were two uses of each picture.
40
41
42
43
44
45

46
47 On each trial, a fixation cross was displayed for 1,000 ms, followed by the instruction screen
48
49 for 2,000 ms. After a 500-ms blank screen, the picture appeared for 1,500 ms. Participants were
50
51 instructed to produce their description as quickly and accurately as possible. Each trial ended with a
52
53 1,000-ms blank screen before the next trial began. Between blocks, participants were left free to rest
54
55 until both were ready to continue, and the experiment resumed when both indicated they were ready
56
57 to continue. The experiment lasted about 1 hour.
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Recording and data analysis. Both participants' responses were recorded and coded for accuracy offline. A participant's description was marked as incorrect if the participant was disfluent, used a word for the character names other than the trained one, used a verb other than the one printed on the picture, or produced an active instead of a passive, or *vice versa*. Accuracy was not affected by our manipulations in Experiment 1, so we do not discuss it further (see Online Supplementary material, Tables S1.1-4, for statistical analyses). For the purpose of the analyses reported below, all incorrect responses were discarded (see Results for percentages). Recordings were then pre-processed to reduce background noise, with speech onsets being automatically tagged (using the Silence finder algorithm in Audacity, Version 1.2.5) and then manually checked for non-speech noises (e.g., lip smacks). In addition, speech offsets were manually coded, and utterance durations were defined as speech offset minus speech onset (i.e., including mid-utterance silences).

In our main analysis, we fitted linear mixed-effects models (Gaussian link function) separately to description latencies and durations. The fixed effect model structure was always $\sim 1 + \text{Voice} + \text{Partner Task} + \text{Voice: Partner Task}$. Voice compared passive to active utterances (to assess whether speakers took longer to produce passive utterances). For Partner Task, we defined two contrasts: The *Speaking* contrast compared the average of the Same and Different conditions to the No condition, to assess whether speakers slowed down when they believed their partner was speaking at the same time as them; the *Form* contrast compared the Same to the Different condition, to assess whether speakers were faster when they believed the form of their partner's utterance to be the same as the form of their own utterance (i.e., in the same voice).

We initially fit the full random structure justified by our design, for both participants and items (with correlations between random factors set to zero to aid convergence; Bates, Kliegl, Vasishth, & Baayen, 2015). However, models with full random structure were overparametrised, as indicated by many of the random factors being estimated to be zero or near-zero. We thus re-fit models with simplified random structures, after removing all slopes that were estimated to be less

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 than 10^{-3} . In text we report coefficients, standard deviations, t values, and 95% confidence intervals
4
5 (from the *confint* function, method = “Wald”) from these simplified models, but the models with
6
7 full random structure always yielded comparable findings (see Online Supplementary Material,
8
9 Tables S1.5-6, S2.5-6, S3.5-6 for latencies and Tables S1.14-15, S2.14-15, S3.14-15 for durations).
10
11 We consider $|t|$ values ≥ 2 to indicate significant findings.
12
13
14

15
16 As in Gambi et al. (2015), the models were fitted to the data after excluding extreme outliers
17
18 (< 300 ms or > 3000 ms for description latencies; > 3500 ms for utterance durations). To assess the
19
20 robustness of the findings, we performed additional analyses after data that were more than 3SD
21
22 over or below each participant’s mean were either discarded (trimmed dataset; see Online
23
24 Supplementary Material, Tables S1.12-13, S2.12-13, S3.12-13 for latencies and S1.21-22, S2.21-
25
26 22, S3.21-22 for durations) or replaced with the cut-off value (winsorised dataset; see Online
27
28 Supplementary Material, Tables S1.10-11, S2.10-11, S3.10-11 for latencies and S1.19-20, S2.19-
29
30 20, S3.19-20 for durations). Such additional analyses used simplified random structures to avoid
31
32 overparametrization as described above, and typically yielded the same pattern as the main analyses
33
34 reported below, which excluded only extreme outliers, but we note when they diverged. Full models
35
36 outputs for all analyses are reported in the Online Supplementary Material. All data and analyses
37
38 scripts are available at <https://osf.io/v7t5z/>
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Results

After removing incorrect responses (Different: 13.26%, Same: 12.06%, No: 11.70%), we discarded 7 outliers for the description latency analyses, and 1 outlier for the utterance duration analyses.

Table 1. Description latencies (means and SDs, in ms) by Voice and Partner's Task.

| Partner Task | Experiment 1 | | | Experiment 2a | | | Experiment 2b | | |
|--------------|--------------|--------------|--------------|---------------|---------------|---------------|---------------|--------------|--------------|
| | Diff. | No | Same | Diff. | No | Same | Diff. | No | Same |
| Voice | | | | | | | | | |
| Active | 958 (245) | 940 (241) | 953 (240) | 1025 (316) | 1027 (313) | 1045 (329) | 955 (386) | 953 (385) | 970 (420) |
| Passive | 948 (247) | 936 (235) | 960 (244) | 1027 (324) | 1021 (316) | 1046 (327) | 972 (381) | 907 (345) | 969 (375) |
| Total | 953 (246) | 938 (238) | 956 (242) | 1026 (320) | 1024 (315) | 1046 (328) | 963 (384) | 930 (366) | 970 (398) |

Description Latency. Speakers took longer to begin their description when they believed their partner was speaking at the same time as them ($M = 955$ ms, $SD = 244$ ms) than when they believed their partner was silent ($M = 938$ ms, $SD = 238$ ms; see Table 1 and Figure 2, left panel).

Accordingly, the Speaking contrast was significant ($B = 19.13$, $SE = 5.43$, $t = 3.52$, $CI = [8.48, 29.77]$). In contrast, description latencies did not differ when speakers believed their partners were preparing to say the same utterance as them and when they were preparing to say a different-voice utterance (Form: $B = 3.61$, $SE = 6.89$, $t = 0.52$, $CI = [-9.89, 17.12]$; see Table 1).

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Speakers did not take longer to initiate passive ($M = 948$ ms, $SD = 242$ ms) than active descriptions ($M = 950$ ms, $SD = 242$ ms) (Voice: $B = -0.79$, $SE = 11.93$, $t = -0.07$, $CI = [-24.18, 22.59]$). There was also no evidence of an interaction between Voice and Partner (Speaking X Voice: $B = 2.95$, $SE = 10.86$, $t = 0.27$, $CI = [-18.34, 24.23]$; Form X Voice: $B = 14.33$, $SE = 13.70$, $t = 1.05$, $CI = [-12.53, 41.18]$; see Table 1).

Table 2. Utterance durations (means and SDs, in ms) by Voice and Partner's Task.

| Partner Task | Experiment 1 | | | Experiment 2a | | | Experiment 2b | | |
|--------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | Diff. | No | Same | Diff. | No | Same | Diff. | No | Same |
| Voice | | | | | | | | | |
| Active | 1379 (308) | 1355 (293) | 1359 (280) | 1436 (327) | 1422 (312) | 1423 (317) | 1630 (467) | 1607 (440) | 1582 (406) |
| Passive | 1545 (291) | 1529 (277) | 1523 (269) | 1646 (341) | 1666 (354) | 1648 (355) | 1736 (421) | 1754 (420) | 1742 (437) |
| Total | 1461 (311) | 1443 (298) | 1441 (287) | 1541 (350) | 1544 (355) | 1535 (355) | 1684 (447) | 1679 (437) | 1662 (429) |

Utterance Duration. The analysis of utterance duration revealed a different and unexpected pattern compared to the analysis of description latencies. First, durations did not differ whether speakers believed their partner was speaking ($M = 1,451$ ms, $SD = 299$ ms) or was silent ($M = 1,443$ ms, $SD = 298$ ms; see Figure 3, left panel, and Table 2) (Speaking: $B = 8.94$, $SE = 6.10$, $t = 1.46$, $CI = [-3.03, 20.90]$). In contrast, speakers produced longer utterances when they believed their partner was

RUNNING HEAD: JOINT SENTENCE PRODUCTION

producing a different utterance than when they were producing the same utterance (Form: $B = -20.61$, $SE = 7.92$, $t = -2.60$, $CI = [-36.13, -5.08]$; see Table 2).¹

Unsurprisingly, passive utterances had longer durations than active ones ($M = 1,532$ ms, $SD = 279$ ms vs. $M = 1,364$ ms, $SD = 294$ ms; Voice: $B = 167.73$, $SE = 12.21$, $t = 13.73$, $CI = [143.80, 191.67]$), but there were no interactions between Partner and Voice (Speaking X Voice: $B = -7.40$, $SE = 12.91$, $t = -0.57$, $CI = [-32.70, 17.91]$; Form X Voice: $B = 0.47$, $SE = 15.42$, $t = 0.03$, $CI = [-29.75, 30.69]$; see Table 2).

Discussion

Experiment 1 demonstrated a joint interference effect in the production of sentences. Thus, speakers are able to allocate resources to imagining their partner's act of production even when they are concurrently engaged in planning a whole sentence. As in previous research using simpler production tasks (Gambi et al., 2015), speakers took longer to begin their utterances when they believed their partner was preparing to produce an utterance at the same time as them, compared to when they believed their partner was remaining silent.

¹ Form was significant ($t = -2.46$) also when we replaced data points above and below 3SD from each by-participant average with the cut-off value. However, the effect was not significant ($t = -1.12$) when we instead discarded these data points (1.14%), suggesting that the effect may have been driven by a subset of very long durations. In support of this, when we ran a Bayesian exponential-Gaussian distributional analysis (Bürkner, 2016), we found a small effect of Form on the mean of the Gaussian component (95% Credible Interval = $[0.01, 0.03]$), but a much larger effect on the rate of the exponential component (95% Credible Interval = $[-0.31, -0.11]$), which captures the thickness of the right tail of the distribution; only the latter effect remained significant after trimming.

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 The nature of the representations speakers formed of their partner's imagined utterances was
4 also similar in this task to simpler production tasks (Gambi et al., 2015; Brehm et al., 2019): There
5 was no evidence that speakers took longer to begin their utterances when they believed their partner
6 was preparing to produce an utterance in a different, compared to the same, voice. However, when
7 we analysed utterance durations, we unexpectedly found that speakers spent more time producing
8 their utterances when the partner produced a different utterance than when they produced the same
9 utterance. Since this finding appeared to be driven by a subset of very long durations (see Footnote
10 2), we address it further in Experiment 2.
11
12
13
14
15
16
17
18
19
20
21

22 Finally, we did not find any evidence that participants took longer to plan a passive than an
23 active utterance. As mentioned in the introduction, differences in latencies between actives and
24 passives are not consistently found in the literature (see e.g., Segaert et al., 2011, 2016). Thus, we
25 could not test whether participants would be less likely to represent their partner's utterance when
26 they themselves were planning a more difficult utterance. But importantly, utterance planning was
27 in general more difficult in the current study than in previous work: Latencies in Experiment 1 were
28 about 70 ms longer than in Gambi et al.'s (2015) Experiment 1. Also importantly, we could still test
29 whether speakers formed detailed representations of the voice of their partners' utterances – such
30 detailed representation should have caused interference in the Different compared to Same
31 condition regardless of whether sentence structures differ in planning difficulty, because the
32 interference is due to the mismatch between representations of one's own and one's partner's
33 utterance rather than to the degree of difficulty associated with planning one's partner's utterance.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Figure 2. Speech onset latencies for picture descriptions across Different, No, and Same trials in Experiments 1, 2a, and 2b. Error bars represented 95% bootstrapped confidence intervals.

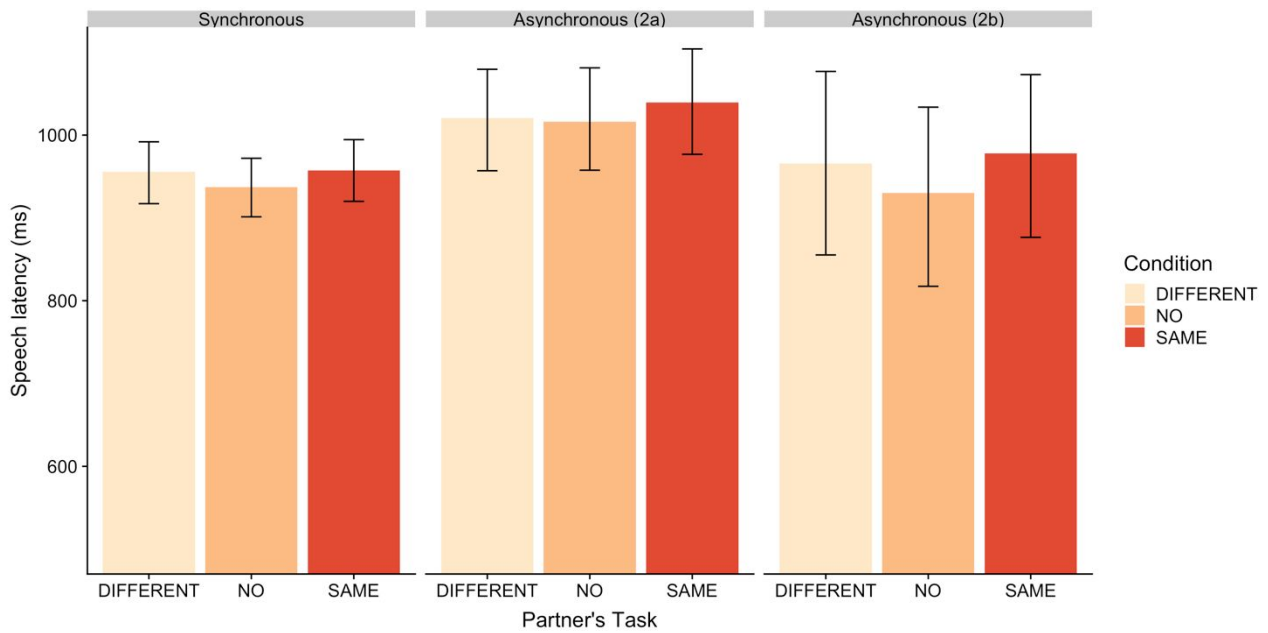
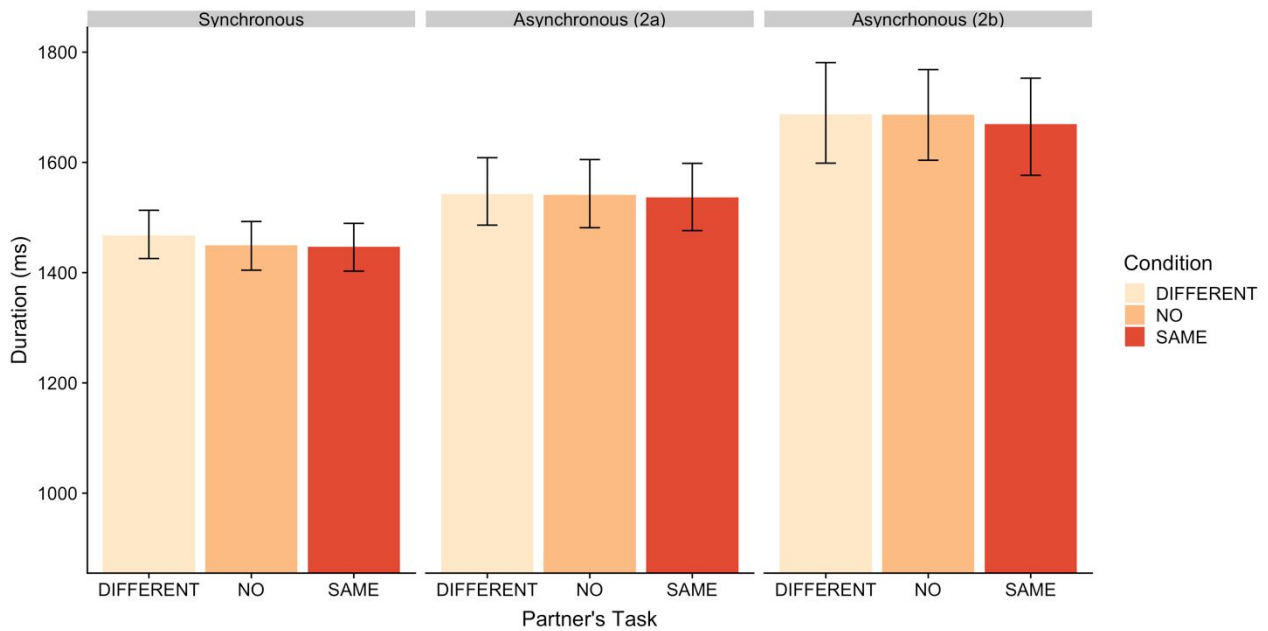


Figure 3. Utterance durations for picture descriptions across Different, No, and Same trials in Experiments 1, 2a, and 2b. Error bars represented 95% bootstrapped confidence intervals.



RUNNING HEAD: JOINT SENTENCE PRODUCTION

Experiment 2: Asynchronous production and representation of another's utterance

In Experiment 2, we asked speakers to plan a sentence and imagine their partner preparing to speak after them, so that planning and representation of their partner's act of production could be asynchronous. If speakers represent their partner's act of production at the time when they believe their partner will speak, we should find no evidence for joint interference effects. In contrast, if speakers represent their partner's intention to speak in the future in the same way as they represent their intention to speak now (or if they do not represent when their partner intends to speak), we should find longer latencies when speakers believe their partner is preparing to speak after them compared to when they believe their partner will remain silent. Since Experiment 1 showed some indication that Different trials may lead to more interference than Same trials, at least after speech onset (utterance duration analyses), we also retained the Voice manipulation, and analysed both latencies and durations.

In Experiment 2a, we had speakers believe their partner would speak after them on Same and Different trials (while remaining silent on No trials), but in fact partners spoke at the same time as the speakers. This method allowed us to collect (equivalent) data from both participants. The purpose of Experiment 2b was to replicate Experiment 2a's findings without deception (i.e., when speakers held a true, rather than a false belief that their partner would speak after them), to rule out the possibility that participants may have been aware that in fact their partner always spoke at the same time as them; e.g., by picking up movement cues in peripheral vision (see Method below).

Method

Participants. An additional 42 participants from the same population as in Experiment 1 took part in Experiment 2a, and were assigned to pairs; one pair was excluded due to experimenter error, leaving 20 pairs for analysis. A further 40 participants, divided into 20 pairs, took part in Experiment 2b. In Experiment 2b, half of the pairs were made up of two friends, and half were

RUNNING HEAD: JOINT SENTENCE PRODUCTION

made up of strangers. This factor did not interact with any within-participant manipulation, and will not be discussed further. Data from two pairs were excluded due to experimenter error, leaving 18 pairs for analysis.

Materials and Design. These were the same as in Experiment 1.

Procedure. The procedure was the same as in Experiment 1, except as follows. In Experiment 2a, both participants in a pair described the picture as soon as it appeared on the screen, as in Experiment 1. However, participants were told that their partner would speak after them, and would do so when a green square (i.e., a go signal) appeared on the screen. The go signal was displayed (for 400 ms) 2,500 ms after the picture appeared on the screen (i.e., the picture was displayed for 1,500 ms, as in Experiment 1, and then a blank screen was displayed for 1,000 ms). The timing of this go signal was set based on the sum of average latencies and durations in Experiment 1, which were just less than 2,500 ms. Finally, the go signal was followed by a blank screen for 2,500 ms. At the start of the session, participants were told that the computer would randomly assign one of them to the role of first speaker, and the other one to the role of second speaker; in fact, they were both informed that they had been assigned the role of first speaker, *so they both believed their partner was speaking after them (while in reality they spoke at the same time)*. In Experiment 2b, participants saw the same sequence of events on each trial, but now one participant in each pair was randomly assigned the role of second speaker at the start of the experiment, and they actually produced their descriptions only after the go signal (the green square) appeared on the screen.

Recording and data analysis. These were also the same as in Experiment 1. In Experiment 2a, we analysed data from both participants (as in Experiment 1). However, in Experiment 2b we analysed data only from first speakers since the aim was to investigate effects of representing another's act of production on speech production processes, and we could not be sure at what point in time second speakers planned their utterance (i.e., immediately as the picture appeared, or closer to the go signal, as they got ready to speak).

RUNNING HEAD: JOINT SENTENCE PRODUCTION

Results

We first discarded incorrect responses. In Experiment 2a, they accounted for 18.01% of the data in the Different condition, 17.85% in the Same condition, and 17.19% in the No condition; in Experiment 2b they accounted for 28.39% of the data in the Different condition, 29.95% in the Same condition, and 29.51% in the No condition. As in Experiment 1, the error rate was not affected by any of our manipulations in either Experiment 2a (see Tables S2.1-4, Online supplementary material) nor in Experiment 2b (see Tables S3.1-4, Online supplementary materials), so we do not discuss it further.² In Experiment 2a, we then discarded 14 outliers from the description latency analysis, and 4 outliers from the utterance duration analysis; the corresponding numbers for Experiment 2b were 9 and 7, respectively.

Description Latency. In Experiment 2a (Figure 2, middle panel and Table 1), speakers took longer to begin speaking ($M = 1,036$ ms, $SD = 324$ ms) when they believed their partner would produce either the same or a different sentence after them, compared to when they believed their partner would be silent ($M = 1,024$ ms, $SD = 315$ ms; Speaking: $B = 14.76$, $SE = 7.10$, $t = 2.08$, CI

² The error rates were higher in Experiment 2b than in Experiment 2a. It is unclear what accounted for this difference; participants performed very similar tasks and underwent the same amount of practice at the start of the experiment. Perhaps more significantly, error rates were fairly high (>10%) across all experiments (cf. Segaert et al., 2011, 2016 – who used a similar paradigm to elicit production of actives and passives – and reported error rates around 6-7%). This high error rate was probably down to some of our pictures being highly confusable (e.g., *nun* and *monk*, *sailor* and *soldier*). Please visit the stimuli&prime_files folder in the OSF repository <https://osf.io/v7t5z/> to view all pictures.

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 = $[0.83, 28.68]$ ³, thus indicating that they experienced interference from their partner's imagined act
4
5 of production even though production and imagined act appeared asynchronous. Unexpectedly,
6
7 speakers also took longer to begin speaking when they believed their partner would be producing
8
9 the same utterance as themselves, compared to when they believed their partner would be producing
10
11 a different utterance (see Table 1; Form: $B = 18.71$, $SE = 7.88$, $t = 2.37$, $CI = [3.26, 34.17]$). As can
12
13 be seen in Figure 1, middle panel, interference appeared to occur primarily in the Same condition
14
15 (and numerically the average latency in the Different condition is very similar to the average for the
16
17 No condition).
18
19
20
21

22
23 Importantly, Experiment 2b suggests that these differences between Experiment 1 and 2a are
24
25 unlikely to be related to the fact that participants believed their partner would be speaking after
26
27 them. This is because Experiment 2b revealed a pattern similar to the one in Experiment 1 (see
28
29 Table 1 and Figure 2, right panel): longer latencies when the partner produced a description after
30
31 the participant ($M = 966$ ms, $SD = 391$ ms) compared to when the partner did not produce a
32
33 description at all ($M = 930$ ms, $SD = 366$ ms; Speaking: $B = 39.99$, $SE = 13.92$, $t = 2.87$, $CI =$
34
35 $[12.72, 67.26]$), but similar latencies regardless of whether the partner produced the same or a
36
37 different utterance (Form: $B = 11.68$, $SE = 14.79$, $t = 0.79$, $CI = [-17.31, 40.67]$; see Table 1 and
38
39 Figure 2, right panel).
40
41
42
43

44 As in Experiment 1, description latencies were not affected by whether speakers were about
45
46 to produce an active or a passive (Exp. 2a: active $M = 1,033$ ms, $SD = 320$ ms vs. passive $M =$
47
48

49
50 ³ Speaking was significant also when we replaced data points above and below 3SD from each by-
51
52 participant average with the cut-off value ($t = 2.06$), but not ($t = 1.95$) when these data points were
53
54 discarded (1.35%). It is possible that the effect of Speaking was driven by a few extreme data
55
56 points, but Bayesian ex-Gaussian analyses did not reveal any effect of Speaking on the rate of the
57
58 exponential component, nor on the mean of the Gaussian component.
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1,031 ms, SD = 322 ms; Exp. 2b: active M = 959 ms, SD = 397 ms vs. passive M = 950 ms, SD = 369 ms), as confirmed by the lack of any significant effects involving Voice. Experiment 2a: Voice, $B = 3.81$, $SE = 17.38$, $t = 0.22$, $CI = [-30.26, 37.88]$; Voice X Speaking, $B = 7.42$, $SE = 13.03$, $t = 0.57$, $CI = [-18.12, 32.97]$; Voice X Form, $B = 1.79$, $SE = 15.09$, $t = 0.12$, $CI = [-27.79, 31.37]$. Experiment 2b: Voice, $B = -15.44$, $SE = 22.24$, $t = -0.69$, $CI = [-59.03, 28.15]$; Voice X Speaking, $B = 55.93$, $SE = 39.25$, $t = 1.42$, $CI = [-21.00, 132.86]$; Voice X Form, $B = -12.87$, $SE = 37.93$, $t = -0.34$, $CI = [-87.20, 61.47]$; see Table 1.

Utterance Duration. Unlike Experiment 1, there was no indication that representing the partner's task affected how long participants spent describing the pictures, in either Experiment 2a (see Figure 2, middle panel) or 2b (see Figure 2, right panel). Utterance durations were highly similar in all three conditions (Table 2). Accordingly, neither Speaking (Exp 2a: $B = -2.35$, $SE = 7.28$, $t = -0.32$, $CI = [-16.62, 11.91]$; Exp. 2b: $B = -14.41$, $SE = 16.91$, $t = -0.85$, $CI = [-47.56, 18.73]$) nor Form (Exp 2a: $B = -3.65$, $SE = 8.15$, $t = -0.45$, $CI = [-19.62, 12.33]$; Exp 2b: $B = -16.27$, $SE = 19.31$, $t = -0.84$, $CI = [-54.12, 21.58]$) was significant.

Unsurprisingly, passive utterances had longer durations than active utterances (Exp 2a: Passive, $M = 1,653$ ms, $SD = 350$ ms; Active, $M = 1,427$ ms, $SD = 319$ ms; Exp 2b: Passive, $M = 1,744$ ms, $SD = 426$ ms; Active, $M = 1,606$ ms, $SD = 439$ ms), as confirmed by a significant main effect of Voice in both experiments (Exp 2a: $B = 225.72$, $SE = 12.73$, $t = 17.73$, $CI = [200.78, 250.67]$; Exp 2b: $B = 142.89$, $SE = 19.48$, $t = 7.33$, $CI = [104.71, 181.08]$), but there were no further interactions with Voice. Exp 2a: Speaking X Voice, $B = -26.83$, $SE = 13.75$, $t = -1.95$, $CI = [-53.77, 0.11]$; Form X Voice: $B = 10.18$, $SE = 15.42$, $t = 0.66$, $CI = [-20.04, 40.41]$). Exp 2b: Speaking X Voice, $B = -30.11$, $SE = 32.14$, $t = -0.94$, $CI = [-93.10, 32.88]$; Form X Voice: $B = 51.95$, $SE = 37.00$, $t = 1.40$, $CI = [-20.58, 124.48]$; see Table 2.

Discussion

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 In Experiment 2, we tested whether speakers experience interference from representing their
4 partner's utterance when they believe that their partner will speak after them, rather than at the same
5 time as them. We were motivated to do so to explore the nature of other-representation. If speakers
6 represent others' utterances at the time when they believe such utterances are being produced, then
7 we should have found no evidence for interference when speakers believe their partner will speak
8 after them. However, in Experiment 2a we found that speakers took longer to begin speaking when
9 they believed their partner would speak after them, compared to when they believed their partner
10 would remain silent, a finding which was further replicated in Experiment 2b. Thus, our findings
11 suggest that speakers represent whether their partner intends to speak, and they represent their
12 partner's intention to speak in the future much in the same way as they represent their intention to
13 speak at the same time as them. In other words, speakers' representations of others' utterances do
14 not distinguish between synchronous or asynchronous speaking – that is, speakers represent that
15 their partners are or are about to speak but not precisely when they will speak. Accordingly,
16 evidence for joint interference was present in the asynchronous task just as in the synchronous task
17 used in Experiment 1.

18
19 In Experiment 2a, latencies were also longer when speakers believed their partner was
20 producing an utterance in the same, compared to in a different voice, but this effect was not
21 replicated in Experiment 2b, where latencies did not differ as a function of what the partner was
22 instructed to say. Taken together, the findings from Experiment 2 thus confirm that, when
23 participants are preparing to speak while they represent their partner's act of production (which was
24 the case in both our synchronous and our asynchronous task), concurrent production limits the
25 resources available for other-representation. As a result, we do not have clear evidence that
26 speakers typically represent any detailed aspect of their partner's utterance planning (e.g., retrieval
27 of specific syntactic representations). This conclusion is clearly supported by Experiment 2b, where
28 neither latencies nor utterance durations were affected by whether the partner would produce a
29 different or the same utterance as the participant. Utterance durations were also unaffected by this

RUNNING HEAD: JOINT SENTENCE PRODUCTION

manipulation in Experiment 2a, and although description latencies were affected in this experiment, speakers were actually *slower* when they believed their partner would produce the same utterance, rather than a different utterance. In sum, there was no indication that speakers experienced more conflict when representing an utterance that was different from the one they were preparing to produce.

Note that there were other differences between Experiment 2a and 2b. For example, speech latencies were somewhat shorter in Experiment 2b, but utterance duration was longer, possibly suggesting that speakers in Experiment 2b planned less before starting to speak and more while speaking compared to speakers in Experiment 2a. But overall, the pattern of findings was remarkably similar in Experiment 2a and 2b. In both experiments, the partner's task did not affect utterance durations and speech latencies were no longer when partners prepared to produce a different compared to the same utterance; however, speech latencies were shorter when partners remained silent.

General Discussion

In three experiments, we showed that speakers take longer to begin producing a sentence when they believe their partner intends to speak, compared to when they believe their partner does not intend to speak. Importantly, in all experiments we observed this interference effect between production and representation of another's act of production, even though speakers were asked to plan a whole sentence in either the active or the passive form – a task which is more demanding than tasks used in previous research. Moreover, we observed interference whether speakers believed their partner would speak at the same time as them (Experiment 1) or after them (Experiments 2a and 2b), suggesting that co-representations of other speakers' utterances do not include timing (see Pickering and Garrod, 2021, chapter 9).

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 In contrast, across the three experiments there was little indication that speakers represented
4 the content of others' utterances. In Experiment 1, utterance durations were longer when
5 participants believed their partner was producing a different (rather than the same) utterance at the
6 same time. This finding may reflect interference due to conflict between active and passive
7 syntactic representations, thus suggesting that, given more time, speakers can represent *what* their
8 partner is saying. Importantly, however, this finding was not replicated in either Experiment 2a or
9 2b. While it is possible that this discrepancy between Experiment 1 and Experiment 2 is related to
10 the change from a synchronous to an asynchronous task, we propose this is highly unlikely, because
11 our findings are otherwise consistent across the synchronous and the asynchronous tasks, with both
12 showing evidence of co-representation of the partner's intention to speak, but not of the content of
13 their utterance. Overall, we have now found this same pattern of results consistently across six
14 different experiments (three reported here, three in Gambi et al., 2015). *Similarly, we think it is*
15 *unlikely that utterance durations are affected differently from speech onset latencies by co-*
16 *representation. Although the clearest evidence for detailed co-representation came from utterance*
17 *durations in Experiment 1, the other two experiments did not replicate this finding and the most*
18 *likely conclusion based on these data and the rest of the literature (Gambi et al., 2015; Brehm et al.,*
19 *2019) is that speakers do not engage in detailed co-representation of others' utterances while they*
20 *prepare to speak themselves.*

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45 In this study, our first aim was to clarify to what extent speakers are able to imagine others'
46 utterances while they are concurrently speaking. In Gambi et al. (2015), we showed that speakers
47 represent that another person is engaging in lexical retrieval while they concurrently prepare to
48 produce a simple one- or two-word utterance. This is one of several studies that support utterance
49 co-representation – that is, that people represent others' utterances in a format that is similar to that
50 in which they represent their own utterances. In the Introduction, we noted that some studies
51 support a stronger version of co-representation, in which speakers represent detailed aspects of
52 others' utterances (e.g., frequency, semantic category), and appear to be engaging in at least partial

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 planning of others' utterances (Baus et al., 2014; Kuhlen & Abdel Rahman, 2017). However, other
4
5 studies support a weaker version of co-representation, where speakers' production processes are
6
7 affected by the fact that another person is engaging in language production but not by the content of
8
9 what they are producing: Rather than planning others' utterances, speakers appear to merely
10
11 represent the fact that others are speaking or about to speak (Gambi et al., 2015); in some cases,
12
13 speakers may not represent their partners' utterance at all (Hoedemaker & Meyer, 2018).
14
15

16
17 We suggest that these apparent discrepancies can be reconciled if task differences are
18
19 carefully considered. Specifically, we argue that content co-representation (i.e., detailed
20
21 representation of what the partner is preparing to say) occurs only when speakers are not
22
23 concurrently tasked with preparing to speak. In contrast, a weaker form of co-representation, co-
24
25 representation of the partner's intention to speak, takes place when concurrent speaking is required,
26
27 because production is cognitively demanding and limits the resources available for co-
28
29 representation.
30
31
32

33
34 In our experiments, we aimed to probe the limits of co-representation by asking speakers to
35
36 produce more complex utterances than used in previous research. *While a direct comparison to
37
38 previous work is not possible, the average latencies across all three experiments in this study were
39
40 longer than in any of the picture naming studies reported by Gambi et al. (2015).* For example,
41
42 latencies in Experiment 1 of this study were about 70 ms longer than in Gambi et al.'s Experiment
43
44 1. Despite this increase in complexity, speakers in all three experiments still clearly engaged in the
45
46 weaker form of co-representation, indicating that they are able to at least represent whether another
47
48 is speaking in a wider range of situations than had previously been considered. Since in most
49
50 instances of joint language production, some degree of overlap between planning of one's own
51
52 utterances and representation of another's utterances is necessary *(not only in choric speech, but
53
54 also when taking turns; Levinson, 2016), this finding is important because it shows that a weak
55
56 form of co-representation is possible even while concurrently planning a fairly complex utterance.*
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 In addition, we showed that, while speakers represent whether their partner is planning to
4 speak or not, this representation mechanism does not include information about timing. This
5 suggests that this form of weak co-representation could affect coordination both in situations when
6 people speak simultaneously (e.g., choric speech) and in those where people take turns (e.g., in
7 conversation). There is already evidence that people slow down during choric speech (Cummins,
8 2009), and we suggest that interference between representation of the other's utterance and
9 production of one's own utterance may contribute to this slowing down.
10
11
12
13
14
15
16
17
18
19

20 In conversation, it may be that this weak co-representation mechanism helps listeners who
21 are preparing to initiate their turn to delay their entry when they detect the current speaker's
22 intention to carry on speaking, or indeed another interlocutor's intention to take the floor. Thus, it
23 may help prevent simultaneous starts during conversation. Future research could test these
24 proposals using versions of our task that more closely resemble natural conversational exchanges,
25 for example by incorporating a turn-taking component.
26
27
28
29
30
31
32
33
34

35 Finally, the finding that people represent future utterances in much the same way as current
36 utterances is compatible with the idea that co-representation uses similar mechanisms as prediction,
37 and that both rely on taking the production system "offline", and using it to simulate another
38 speaker (Gambi & Pickering, 2016). Our findings are consistent with the proposal that
39 representation of others' utterances uses the production system, because we have shown
40 interference between other-representation and language production. However we have not
41 conclusively demonstrated that the production system is directly involved in other-representation.
42 This is because it is possible that other-representation instead relied on a different, non-language
43 specific mechanism (e.g., rehearsal of a working memory representation of the other's task
44 instruction), which also happened to (indirectly) affect the speaker's act of production (i.e., because
45 participants were also trying to rehearse their own task instruction). But importantly, this alternative
46 explanation is rendered much less likely by the fact that in previous work we found that interference
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

1
2
3 was greater when speakers imagined the partner was engaged in lexical retrieval, compared to a
4 semantic categorization task (Gambi et al., 2015; see also Brehm et al., 2019). It is therefore more
5 likely that language production mechanisms are involved in other-representation (Gambi &
6 Pickering, 2016).
7
8
9
10
11
12

13 Note that one finding is potentially inconsistent with the account we have presented.
14
15 Speakers in Hoedemaker and Meyer (2018) were no slower at naming a picture when they believed
16 their partner would name a second picture after them, as opposed to believing nobody would name
17 that picture (they were instead much slower when they had to name the second picture themselves).
18 This finding is in contrast to the delayed latencies in our Experiment 2. While there are several
19 potential reasons for this discrepancy (as the studies used quite different tasks), one possibility is
20 that speakers in Hoedemaker and Meyer's study were under the influence of two competing effects.
21 On the one hand, they experienced interference due to representing that their partner was about to
22 start speaking (as in our Experiment 2). On the other, they were under pressure to speed up
23 production to avoid overlapping with their partner, which would have led them to invest extra
24 resources to begin production more quickly. If this is correct, Hoedemaker and Meyer may have
25 found no evidence for interference due to co-representation in consecutive production because this
26 effect was masked by a competing pressure to speak quickly. In contrast, participants in our
27 Experiment 2 were not required to coordinate their utterances (and in fact, could not hear each other
28 at all), so that no pressure to speak quickly should have been present.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 In sum, we have shown that speakers experience interference when they represent that
49 another speaker also intends to speak while, or soon after, they speak. We propose that this joint
50 interference effect occurs because speakers represent their partner's intention to speak and that the
51 locus of the effect occurs when a decision is made to activate the language production system.
52
53
54
55
56
57
58
59
60

Acknowledgments

We thank Anna Catherine Mackenzie and Kristen Nelissen for assistance with data collection. CG was supported by a scholarship from the School of Philosophy, Psychology and Language Sciences, University of Edinburgh. JvDC was supported by an EU Erasmus scholarship to visit the University of the Edinburgh.

References

- Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Retrieved from <http://arxiv.org/pdf/1506.04967.pdf>
- Baus, C., Sebanz, N., de la Fuente, V., Branzi, F. M., Martin, C., & Costa, A. (2014). On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition*, *133*(2), 395-407.
- Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of Memory and Language*, *69*(3), 228-256.
- Bock, K., & Levelt, W. J. M. (1994). Language Production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego: Academic Press.
- Brehm, L., Taschenberger, L., & Meyer, A. (2019). Mental representations of partner task cause interference in picture naming. *Acta Psychologica*, *199*, 102888.
- Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 204-215.
- Bürkner, P.-C. (2016). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1-28.

RUNNING HEAD: JOINT SENTENCE PRODUCTION

- 1
2
3 Corps, R., Crossley, A., Gambi, C., & Pickering, M. J. (2018). Early preparation during turn-taking:
4
5 listeners use content predictions to determine what to say but not when to say it. *Cognition*,
6
7 *175*, 77-95.
8
9
- 10 Cummins, F. (2003). Practice and performance in speech produced synchronously. *Journal of*
11
12 *Phonetics*, *31*(2), 139-148.
13
14
- 15 Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of*
16
17 *Phonetics*, *37*(1), 16-28.
18
- 19 Demiral, Ş. B., Gambi, C., Nieuwland, M. S., & Pickering, M. J. (2016). Neural correlates of verbal
20
21 joint action: ERPs reveal common perception and action systems in a shared-Stroop task.
22
23 *Brain Research*, *1649*, 79-89.
24
25
- 26 Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared
27
28 utterances. *Journal of Memory and Language*, *30*(2), 210-233.
29
- 30 Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of*
31
32 *Memory and Language*, *33*, 715-715.
33
34
- 35 Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances.
36
37 *Frontiers in Psychology*, *2*. doi:10.3389/fpsyg.2011.00275
38
- 39 Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*,
40
41 *4*, doi:10.3389/fpsyg.2013.00340.
42
43
- 44 Gambi, C., & Pickering, M. J. (2016). Predicting and imagining language. *Language, Cognition*
45
46 *and Neuroscience*, *31*(1), 60-72.
47
48
- 49 Gambi, C., Van de Cavey, J., & Pickering, M. J. (2015). Interference in joint picture naming.
50
51 *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 1-21.
52
- 53 Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of
54
55 turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in psychology*, *6*,
56
57 1492. <https://doi.org/10.3389/fpsyg.2015.01492>
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

- 1
2
3 Hoedemaker, R. S., Ernst, J., Meyer, A. S., & Belke, E. (2017). Language production in a shared
4 task: Cumulative semantic interference from self-and other-produced context words. *Acta*
5
6 *Psychologica*, 172, 55-63.
7
8
9
10 Hoedemaker, R. S., & Meyer, A. S. (2018). Planning and coordination of utterances in a joint
11 naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
12
13 *Advance online publication*. doi:10.1037/xlm0000603.
14
15
16
17 Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in
18 picture naming: Experimental and computational studies. *Cognition*, 100(3), 464-482.
19
20
21 Jasmin, K. M., McGettigan, C., Agnew, Z. K., Lavan, N., Josephs, O., Cummins, F., & Scott, S. K.
22
23 (2016). Cohesion and joint speech: right hemisphere contributions to synchronized vocal
24 production. *Journal of Neuroscience*, 36(17), 4669-4680.
25
26
27
28 Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory
29 and data. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 54, pp. 59-101).
30
31 Burlington: Academic Press.
32
33
34
35 Kuhlen, A. K., & Abdel Rahman, R. (2017). Having a task partner affects lexical retrieval: Spoken
36 word production in shared task settings. *Cognition*, 166, 94-106.
37
38
39
40 Kuhlen, A. K., & Abdel Rahman, R. (2021). Joint language production: An electrophysiological
41 investigation of simulated lexical access on behalf of a task partner.. *Journal of*
42
43 *Experimental Psychology: Language, Memory, and Cognition*.
44
45 <https://doi.org/10.1037/xlm0001025>
46
47
48
49 Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
50
51
52 Levinson, S. C. (2016). Turn-taking in human communication—origins and implications for
53 language processing. *Trends in Cognitive Sciences*, 20(1), 6-14.
54
55
56 Magyari, L., Bastiaansen, M. C., De Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies
57 behind the speed of response in conversation. *Journal of Cognitive Neuroscience*, 26(11),
58
59 2530-2539.
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

- 1
2
3 Menenti, L., Gierhan, S., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and
4
5 segregation of the neuronal infrastructure for speaking and listening revealed by fMRI.
6
7 *Psychological Science*, 22(9), 1173-1182.
8
9
- 10 Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture-word
11
12 interference experiments. *Journal of Memory and Language*, 35(4), 477-496.
13
14
- 15 Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A
16
17 model of cumulative semantic interference during lexical access in speech production.
18
19 *Cognition*, 114(2), 227-252.
20
- 21 Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social*
22
23 *interaction*. Cambridge, UK: Cambridge University Press.
24
25
- 26 Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in*
27
28 *Psychology*, 2, 10.3389/fpsyg.2011.00307.
29
- 30 Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of
31
32 turn-taking for conversation. *Language*, 50(4), 696-735.
33
34
- 35 Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task
36
37 constraints in joint action. *Journal of Experimental Psychology: Human Perception and*
38
39 *Performance*, 43(8), 1480-1493.
40
41
- 42 Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2018). Co-actors represent the order of each
43
44 other's actions. *Cognition*, 181, 65-79.
45
46
- 47 Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own?
48
49 *Cognition*, 88(3), B11-B21.
50
- 51 Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-
52
53 response mappings. *Journal of Experimental Psychology: Human Perception and*
54
55 *Performance*, 31(6), 1234-1246.
56
57
58
59
60

RUNNING HEAD: JOINT SENTENCE PRODUCTION

- 1
2
3 Segaert K., Menenti L., Weber K., & Hagoort P. (2011) A paradox of syntactic priming: Why
4
5 response tendencies show priming for passives, and response latencies show priming for
6
7 actives. *PLoS ONE* 6(10): e24209. <https://doi.org/10.1371/journal.pone.0024209>
8
9
10 Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic
11
12 choices and timing of sentence generation. *Journal of Memory and Language*, 91, 59-80.
13
14 Smith, M., & Wheeldon, L. (2004). Horizontal information flow in spoken language production.
15
16 *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3), 675-686.
17
18
19 Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G.,
20
21 Rossano, F., de Ruiter, J.P., Yoon, K-E, & Levinson, S. C. (2009). Universals and cultural
22
23 variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*,
24
25 106(26), 10587–10592.
26
27
28 Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively
29
30 facilitates lexical access during overt object naming. *Journal of Memory and Language*,
31
32 65(4), 345-362.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60