

Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking

Gerard Kempen^{1,2}, Nomi Olsthoorn³, and Simone Sprenger⁴

¹Max Planck Institute for Psycholinguistics, Nijmegen
The Netherlands

²Cognitive Psychology Unit, Leiden University, Leiden,
The Netherlands

³Amsterdam Centre for Language and Communication, University of
Amsterdam, Amsterdam, Netherlands

⁴Department of Psychology, University of Groningen, Groningen
The Netherlands

Grammatical encoding and grammatical decoding (in sentence production and comprehension, respectively) are often portrayed as independent modalities of grammatical performance that only share declarative resources: lexicon and grammar. The processing resources subserving these modalities are supposed to be distinct. In particular, one assumes the existence of two workspaces where grammatical structures are assembled and temporarily maintained—one for each modality. An alternative theory holds that the two modalities share many of their processing resources and postulates a single mechanism for the online assemblage and short-term storage of grammatical structures: a shared workspace. We report two experiments with a novel “grammatical multitasking” paradigm: the participants had to read (i.e., decode) and to paraphrase (encode) sentences presented in fragments, responding to each

Correspondence should be addressed to Gerard Kempen, Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands. E-mail: gerard.kempen@mpi.nl

The experiments reported herein were conducted at the Cognitive Psychology Unit of Leiden University (Experiment 1) as part of the second author's PhD project, and at the Max Planck Institute for Psycholinguistics (Experiment 2) as part of a Postdoc project awarded to the third author. We thank three anonymous reviewers for comments on earlier versions of the text and for helpful suggestions.

input fragment as fast as possible with a fragment of the paraphrase. The main finding was that grammatical constraints with respect to upcoming input that emanate from decoded sentence fragments are immediately replaced by grammatical expectations emanating from the structure of the corresponding paraphrase fragments. This evidences that the two modalities have direct access to, and operate upon, the same (i.e., token-identical) grammatical structures. This is possible only if the grammatical encoding and decoding processes command the same, shared grammatical workspace. Theoretical implications for important forms of grammatical multitasking—self-monitoring, turn-taking in dialogue, speech shadowing, and simultaneous translation—are explored.

Keywords: Language production; Language comprehension; Grammatical encoding; Grammatical decoding; Parsing; Human sentence processing; Speech shadowing; Self-monitoring; Self-repair; Simultaneous interpreting; Simultaneous translation; Multitasking; Spoken dialogue.

In our daily lives, we often find ourselves producing language and comprehending language at the same time. Involved in a conversation and listening to what our dialogue partner is saying, we sometimes produce a comment or reply already before the other person finishes speaking (overlapping dialogue turns; Sacks, Schegloff, & Jefferson, 1974; Stivers et al., 2009). Listening in to a news broadcast, we produce spoken comments on an exciting item. While talking, we listen to our own speech in order to spot and quickly rectify erroneous or infelicitous utterances (“self-monitoring”; Levelt, 1983, 1989). If we are fluent in more than one spoken language, we can translate between these languages online (“simultaneous interpreting”; Christoffels, De Groot, & Kroll, 2006). We can “shadow” relatively simple spoken sentences with an ear-voice span of about a second, in the meantime spontaneously correcting morpho-syntactically ill-formed input; and some of us accomplish this feat with much shorter ear-voice spans of less than a third of a second (Marslen-Wilson, 1973, 1985). Language production and comprehension in parallel is commonplace when dealing with written and sign language as well. Taking notes during a lecture, we produce condensed written versions of the ongoing spoken discourse. While reading out loud a newspaper headline or a telegram-style message to another person, we can insert the missing little words on the fly. During proofreading, we rephrase incorrect or infelicitous fragments of the text we are reading. Interpreters for the deaf listen and sign at the same time.

Language production and language comprehension both include cognitive processes that deal with grammatical form. In this paper, we focus on these processes, called GRAMMATICAL ENCODING (during language production) and GRAMMATICAL DECODING (during language comprehension).

Grammatical encoding and grammatical decoding both perform mappings between meaning and form—more precisely, mappings between nonlinguistic communicative intentions and lemma¹ strings that underlie spoken, written, or signed linguistic expressions of such intentions. Grammatical encoding deals with mappings from communicative intentions to linguistic expressions, grammatical decoding with mappings from linguistic expressions to communicative intentions. In both directions, the mappings are usually assumed to include submappings between communicative intentions and CONCEPTUAL structures on the one hand, and between conceptual and SYNTACTIC structures on the other. Communicative intentions include representations of states, state changes, events, etc., in the extralinguistic context, together with their relationships (e.g., locative, temporal, causal).

In the psycholinguistic community, the two modalities of grammatical performance tend to be studied separately—witness the fact that they are typically addressed in different textbook chapters, in unconnected research groups, and in separate conferences. The grammatical encoding and decoding tasks are supposed to be subserved by distinct, DEDICATED processing resources, which presumably do not have much more in common than access to the same declarative resources: the lexicon and the grammar.² Among the cognitive resources subserving grammatical processing is a workspace for the assemblage and short-term storage of syntactic and conceptual structures.³ The assumption of dedicated processing resources then suggests there are two workspaces, one subserving each modality.

An alternative theory of grammatical encoding and decoding holds that the two modalities execute the mappings between communicative intention and linguistic expressions in very similar manners, notwithstanding the

¹ Lemmas are “syntactic words”; their lexical entries specify the sentential environments in which they are allowed to occur. The morphological and phonological information associated with words is specified in another type of lexical entry called word forms or lexemes.

² It is usually acknowledged that developmental, pathological, and performance factors may lead to differential accessibility of the lexical or grammatical information needed for the two tasks.

³ The term “workspace” should not be interpreted with emphasis on “work”, *in casu* the assemblage of (key elements of) grammatical structures. The assembly process presupposes temporary storage of (partial) assembled structures (“working memory”) but we do not wish to imply that the cognitive resources underlying grammatical assembly processes are distinct from the memory resources enabling temporary storage of the products. The Unification Space model of human syntactic parsing by Vosse and Kempen (2000, 2009) exemplifies a style of neurocognitive modelling that makes do without such a division of labour. The distinction tends to be absent from brain-style modelling generally. The grammatical workspace intended here should also be distinguished from the short-term “buffer” where strings of words (more precisely, strings of phonological forms or lexemes) are temporarily stored. This buffer can be occupied by the output of the word recognition mechanism during language comprehension, and by the output of the morpho-phonological encoding process during language production.

directional difference (see Kempen, 2000, for a list of similarities). Hence, grammatical encoding and decoding could be accomplished by *SHARED* processing resources—by a single exemplar of all, or all important, parts of the cognitive resources. In the psycholinguistic literature, we find a limited number of theoretical proposals in this direction, suggesting that the mechanism responsible for executing a certain grammatical subtask within one modality could also be deployed in service of the other modality and perform there the same or a similar subtask (Garrett, 2000; Pickering & Garrod, 2007; Townsend & Bever, 2001).

The above examples of temporally overlapping varieties of language production and language comprehension do not necessarily imply that the tasks of grammatical encoding and grammatical decoding can be performed “truly simultaneously”. Simultaneity presupposes, among other things, that the tasks are executed by *DISTINCT* processing resources. Two tasks that are subserved by the *SAME* cognitive processing resources cannot be executed truly simultaneously but require task switching (alternation, time-sharing, time-division). As for terminology, we use the terms “concurrent tasks” and “multitasking” if the time interval between start and finish of the execution of one task overlaps with the corresponding interval of another task (for short: execution of different tasks in overlapping timespans). Notice that this includes the possibilities not only of task switching but also of truly simultaneous execution of several tasks.

The experiments we report in the present paper focus on the following theoretical question: Do the cognitive resources for linguistic performance include two *DEDICATED WORKSPACES* for the assemblage and temporary storage of grammatical structures, one commanded by each modality, or only *A SINGLE SHARED WORKSPACE* that is at the disposal of both modalities. We had language users carry out a task that involves “grammatical multitasking” (grammatical encoding and grammatical decoding in overlapping timespans) but did not attempt to take precise measurements of temporal relations between the two subtasks. Instead, we capitalised on a different empirical contrast predicted by the competing cognitive architectures under consideration. A dedicated-workspaces architecture predicts that encoding and decoding processes are able to assemble, during overlapping time intervals, two different grammatical forms—one being decoded, the other encoded—and that they can do this without the need to switch, within those intervals, between the tasks and between the grammatical forms involved in the tasks. The hypothesis of a single shared workspace, in contrast, predicts that concurrent assemblage of two differing grammatical forms—one by the encoder, one by the decoder—requires switching between tasks and between grammatical forms involved. The experimental paradigm we developed requires rapid alternation between decoding and encoding episodes, and the

empirical issue is whether or not the task switches go hand in hand with switches between the two differing grammatical structures being assembled in fulfillment of the tasks.

In the next section, we describe the novel grammatical multitasking paradigm. Two experiments based on this task constitute the core sections of the paper. In the discussion section we explore some important implications of our findings for well-known instances of grammatical multitasking: self-monitoring, turn-taking in dialogue, speech shadowing, and simultaneous interpreting/translation.

GRAMMATICAL MULTITASKING

In order to pit the competing architectural hypotheses against one another, we devised an experimental test requiring the participants to read a sentence and to paraphrase it on the fly. Consider a person who reads the written sentence *The government has decided to put its plan into force*, and WHILE doing that, produces a spoken SYNTACTIC PARAPHRASE that she likes better than the original input. Due to a preference for subject-verb agreement with the conceptual rather than the syntactic number value of the subject NP, she applies pluralisation. That is, she reads the sentence out loud as *The government HAVE decided to put THEIR plan into force*. For ease of exposition, assume that she processes the sentence incrementally in two parts: finite main clause, and infinitival complement clause. Any edit introduced during an encoding step potentially affects subsequent decoding steps. In the example, after letting the main verb agree with the subject's conceptual number, any pronouns taking this NP as antecedent are also expected to have plural number: *its* will be replaced by *their*.

In the experiments below, sentences were presented as sequences of fragments—one fragment per trial—and the participants responded to every fragment with a fragment of a paraphrase, executing the experimenter's instruction on how to transform the input sentence. Every fragment triggered a decoding action (reading the fragment) followed by an encoding action (encoding the required output fragment); hence, in each trial, the two modalities were active alternately: task switching. The participants reacted to each fragment with a spoken response as quickly as possible. A voice key measured the latency of every response.

In the course of processing an entire sentence, the participants switched modality repeatedly: in each trial, a decoding episode was followed by an encoding episode. On the dedicated-workspaces hypothesis, the currently active modality always works with its “own” representation of the sentence, stored in its “own” workspace (either the original input or the output paraphrase). The shared-workspace hypothesis predicts that the currently

active modality works with the sentence representation that the other modality has left behind in the shared workspace.

All experimental stimuli contained a quoted main clause (“direct speech”) with the first-person personal pronoun as subject NP; e.g., (1a). However, in some trials, the subordinating conjunction *that* was presented in a visually salient manner, indicating that, from then onward, the DIRECT-speech clause had to be realised as grammatically correct INDIRECT speech, e.g., as (1b); see Figure 1. The experimental stimuli contained either a grammatically correct reflexive pronoun (*myself*, like in (1a)) or an incorrect one (*himself*, as in (1c)). The critical measurements were the response onset latencies (RTs) to the correct output reflexive pronoun: i.e., *himself*. (NB The example

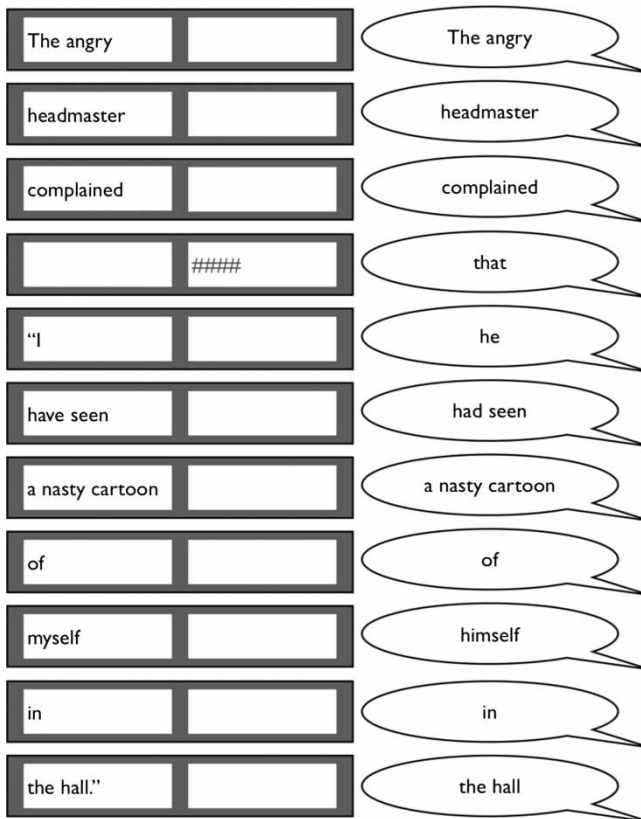


Figure 1. Sentence fragments (left) and vocal responses (right) during the Paraphrasing task (example similar to the materials of Experiment 2). The grey rectangles represent the computer screen, the enclosed white rectangles the frames where input could appear.

sentences in this Section are in English but the experiments were carried out in Dutch.)

- (1) a. The angry/headmaster/complained:/
 “I have/seen/a nasty cartoon/of/ myself/in/the hall”.
 b. The angry/headmaster/complained/
 that/he had/seen/a nasty cartoon/of/himself/in/the hall.
 c. *The angry/headmaster/complained:/
 “I have/seen/a nasty cartoon/of/himself/in/the hall”.

On the dedicated-workspaces hypothesis, once the decoder has parsed the pronoun *I*, the first-person value remains intact throughout the input sentence and is used when the reflexive pronoun is decoded later on. The fact that, in the meantime, the encoder has produced a structure with a third-person subject NP, does not affect the content of the decoding workspace. Stated differently, each modality of grammatical performance works with its own grammatical form-under-construction, and switching between modality entails switching between associated forms. On the shared-workspace hypothesis, however, the two modalities work with the same (token-identical) grammatical form-under-construction. Hence, the encoding process overwrites the value of the subject’s person feature to third-person, and the new value is used when decoding the reflexive pronoun. Because the latter has to agree with the subject in number, gender, and person, the decoding process spots a mismatch. THIS SUGGESTS THAT CORRUPTING AN INPUT SENTENCE LIKE (1A) BY CHANGING THE CORRECT *MYSELF* TO INCORRECT *HIMSELF*, AS IN (1C), WILL NOT BE NOTICED BY THE DECODING PROCESS. This prediction was tested in the experiments below.⁴

In addition to the Paraphrasing task, the experiments included a very similar control task dubbed **PROOFREADING**. Again, the participants read sentences with either a correct or an incorrect input reflexive. All they had to do was to repeat the input fragments out loud, immediately correcting any morpho-syntactically incorrect input words. For Proofreading trials with input sentences like (1c), we predict—on the dedicated-workspaces as well as the shared-workspace hypothesis—a response delay due to the combination of an expectancy violation (incorrect input reflexive) and a repair action, whereas no delay is predicted for well-formed input sentences like (1a):

⁴An informal demonstration of the sensitivity of the decoding process to parallel events in the encoding process can proceed as follows. One asks participants who have performed a Paraphrasing trial with an ungrammatical input sentence like (1c), to re-read the sentence without paraphrasing. According to the first author’s experience during lectures and in classrooms (since 2003), while re-reading, nearly all serious participants express surprise that the sentence is ill-formed and that they had failed to notice this during the Paraphrasing trial.

baseline performance. The Proofreading task provides an independent measurement, in the absence of paraphrasing operations but under otherwise identical conditions, of the time needed to deal with an expectancy violation followed by carrying out the corresponding repair.⁵

As the overt response to an input reflexive pronoun that does not fit the output sentence, is determined by sequentially executing decoding and encoding steps, any processing delays during these steps add up. The decoder may incur a delay due to VIOLATION of a grammatical expectation. The encoder is delayed when it has to execute a repair: setting the person/number/gender features of the reflexive equal to those of the subject NP, and retrieving the corresponding lexeme from the mental lexicon.⁶ Expectancy violations will often be followed by repairs, but on the dedicated-workspaces hypothesis this does not always hold for the Paraphrasing task: on this hypothesis, the encoding and the decoding processes are able to work with distinct grammatical structures concurrently. Hence, situations may arise where the decoding process runs into a violated expectation whereas the encoding process, due to having embarked upon a modified output sentence, need not launch a repair—simply because the input word that caused the expectation violation actually happens to fit syntactically into the output sentence under construction. This holds for Paraphrasing trials with sentence (1c) as input. In the opposite case, exemplified by Paraphrasing trials with sentence (1a) as input, the decoding process can proceed unhampered by any expectancy violations, but the encoding process needs to initiate a repair in order to come up with the reflexive pronoun that fits into the correct output sentence (1b).

According to the SHARED-WORKSPACE hypothesis, which entails that the decoding process always works with the same (token-identical) grammatical structure as the encoding process, an expectancy violation and a repair are either both present (causing a “double” delay: violation + repair) or both absent (no delay: baseline performance). Applied to upcoming reflexive pronouns, this means that such a pronoun either violates an expectation engendered by the decoding process, thereby indirectly triggering a repair

⁵Incidentally, we assume that, at every input fragment, Proofreading also requires a switch from grammatical decoding to grammatical encoding. During an encoding episode, the conceptual content derived from the current input fragment is grammatically encoded by using as much as possible the lexical material of the input fragment. Under this characterisation, the Proofreading task is neutral between the two architectural hypotheses under investigation.

⁶One may object that the second source of RT delays is “merely” a Stroop-effect insofar as the speaker has to withhold the primary response to the stimulus word and to produce another response instead. However, the precise response selection mechanism underlying the RT delays is irrelevant: We only assume that the RTs reflect the syntactic expectancies prevailing at the moment the speakers prepare their response to the reflexive pronoun. These serve as a window on the object of our inquiry: the grammatical structure(s) currently under assembly.

within the encoding process (e.g., in Paraphrasing trials with (1a) as input, or in Proofreading trials with (1c) as input) or does not violate any expectation, hence obviating the need for a repair (yielding baseline performance, e.g., in Paraphrasing trials with (1c), or in Proofreading trials with (1a) as input).

The DEDICATED-WORKSPACES model, in contrast, predicts a different RT pattern for Paraphrasing trials. With input sentences (1a) and (1c), the RTs should lie in between those predicted for baseline conditions in Paraphrasing and Proofreading (no delay) on the one hand, and the condition that triggers a “double” delay on the other (violation + repair: proofreading incorrect input, like sentence (1c)). During Paraphrasing, after well-formed input (1a) no violation will be spotted by the decoder, but the encoder needs to make a repair (selection and production of a new reflexive). After ill-formed input (1c), there is an expectancy violation but no repair. (In Proofreading, incorrect input (1c) gives rise to a double violation, both on the dedicated-workspaces and the shared-workspace model. See also the derivation of the predictions for the two experiments, and the Tables in the Appendix.)

In the next two sections, we describe two Paraphrasing/Proofreading experiments. In Experiment 1, the Proofreading and the Paraphrasing tasks were performed by different groups of participants (i.e., “Task” was a between-Subjects factor; we preferred this over a within-Subjects design because of fear that the participants might mix up the two tasks). Both groups of participants produced indirect-speech output sentences where third-person reflexive pronouns were correct. Experiment 2 used a within-Subject design: all participants were tested on both tasks (one block of Proofreading trials and another block of Paraphrasing trials, with order of blocks counterbalanced). Moreover, after these blocks, the participants executed a silent Self-Paced Reading (SPR) task with the same sentences (direct-speech versions). The SPR data served to verify that the ungrammatical reflexives indeed cause comprehension problems when no spoken responses are required.

EXPERIMENT 1

In Experiment 1, the two tasks (Proofreading and Paraphrasing) were administered to different groups of Dutch-speaking participants. Both groups had to read input sentences fragment by fragment, responding to each fragment as fast as possible with a spoken response—either a repetition of the input fragment or a morpho-syntactic variant thereof, depending on the task instruction. All experimental sentences contained a reflexive pronoun. Grammaticality of the input sentences was manipulated by including a reflexive pronoun that did or did not match the person feature of its antecedent (the subject NP). In the Proofreading task, with

TABLE 1

Experiment 1: Sources of response delay predicted by the dedicated-workspaces and the shared-workspace hypotheses. The numbers refer to sample sentences mentioned in the Materials subsection. For explanation see Tables A1 and A2 in the Appendix

Input reflexive pronoun	Task and correct output reflexive pronoun	
	Paraphrasing: <i>himself</i>	Proofreading: <i>himself</i>
	Dedicated workspaces	
<i>myself</i>	REPAIR (cor-inp: 2a)	VIOLATION + REPAIR (incor-inp: 3b)
<i>himself</i>	VIOLATION (incor-inp: 2b)	— (cor-inp: 3a)
	Shared workspace	
<i>myself</i>	VIOLATION + REPAIR (cor-inp: 2a)	VIOLATION + REPAIR (incor-inp: 3b)
<i>himself</i>	— (incor-inp: 2b)	— (cor-inp: 3a)

INDIRECT-SPEECH input sentences such as (1b), the input reflexive was either a first- or third-person singular reflexive (*mezelf* or *zichzelf*, the Dutch equivalents of *myself* or *him/herself*⁷). The third-person reflexive was always GRAMMATICALLY CORRECT WITH RESPECT TO THE INPUT SENTENCE, so the correct response only required overt pronunciation of this pronoun. We will refer to this condition as “Proofreading-cor-inp”. The first-person pronoun was INCORRECT WITH RESPECT TO THE INPUT SENTENCE (“Proofreading-incor-inp”) and had to be replaced by the third-person reflexive. In the Paraphrasing task, with DIRECT-SPEECH input sentences such as (1a), the input reflexive was also a first-person (“Paraphrasing-cor-inp”) or third-person (“Paraphrasing-incor-inp”) singular reflexive, and the correct response was the third-person reflexive pronoun in both input conditions. The subordinating conjunction *dat* “that” announced the onset of the to-be-paraphrased part of the input sentence. The measures of interest in both tasks were reaction times for the reflexive pronoun.

The RT predictions of Experiment 1 are summarised in Table 1. For a detailed account of their derivation, see Tables A1 and A2 in the Appendix. The dedicated-workspaces hypothesis predicts that response delays may be caused either by a decoding problem (expectancy violation with respect to the input pronoun), or an encoding problem (repair), or a combination of a decoding and an encoding problem. In conjunction with the assumption of additivity of incurred delays, this predicts the following RT pattern:

⁷Both *mezelf* and *zichzelf* are gender-neutral. So it was not necessary for the participants to retrieve the gender of the person mentioned in the antecedent NP.

Proofreading-cor-inp < [Paraphrasing-cor-inp, Paraphrasing-incor-inp]
 < Proofreading-incor-inp.

In contrast, workspace sharing predicts that the decoding problem and the encoding problem are either both present or both absent, yielding

Proofreading-cor-inp \approx Paraphrasing-incor-inp < Proofreading-incor-inp
 \approx Paraphrasing-cor-inp.

In addition to the RT tasks, the participants were asked to indicate, following each sentence in each task, whether or not the input sentence was grammatically well-formed (through a spoken *yes/no* response). Since the dedicated-workspaces hypothesis predicts that the encoded and the decoded structures are both maintained, it follows that grammaticality judgements in the Proofreading as well as the Paraphrasing tasks should be (near-)perfect. The shared-workspace hypothesis, however, predicts that the participants will maintain only the encoded structure. Hence, we expect that ungrammaticalities in the input sentences will often go unnoticed. No feedback was given about the correctness of the judgements.

Method

The Paraphrasing task

Participants. Twenty-two members of the Leiden University community were paid to participate in this experiment. All were native speakers of Dutch.

Materials. We constructed 12 Dutch experimental items, like (2), 24 fillers and five practice items.⁸ The experimental sentences consisted of a subject NP, a finite verb followed by a colon and an opening quotation mark (“), a sentence in direct speech, containing a reflexive pronoun that either did or did not agree in person with the antecedent (the subject), and a closing quotation mark (”), like (2a/b). The position of the reflexive pronoun in the sentence varied. In half of the experimental sentences, the reflexive took sentence final position; in the other half it preceded the sentence-final infinitive. The two positions are grammatically correct and semantically equivalent. Each sentence was presented in fragments, like (2c), and the desired output was as in (2d). Notice that the fragment that includes the critical reflexive pronoun, also includes the preceding preposition.

⁸The full list of sentences used in Experiments 1 and 2 is available from the first author.

- (2) a. Input in condition Paraphrasing-cor-inp
De lottowinnaar zei / "Ik heb besloten een rode auto te kopen voor mezelf".
 The lottery winner said: "I have decided a red car to buy for myself".
 "The lottery winner said: 'I have decided to buy a red car for myself'".
- b. Input in condition Paraphrasing-incor-inp
 **De lottowinnaar zei: "Ik heb besloten een rode auto te kopen voor zichzelf".*
 The lottery winner said: "I have decided a red car to buy for himself".
 "**The lottery winner said, 'I have decided to buy a red car for himself'".
- c. Input fragments
De lottowinnaar/ zei:/ [dat] /"Ik heb besloten/ een rode auto/ te kopen/ voor me|zichzelf".
- d. Desired output in both Paraphrasing conditions
De lottowinnaar zei dat hij had besloten een rode auto te kopen voor zichzelf.
 The lottery winner said that he had decided a red car to buy for himself.
 "The lottery winner said that he had decided to buy a red car for himself".

Two rectangular frames were displayed on the monitor, one to the right, the other to the left of the centre. The sentence fragments were presented one-by-one in the left-hand frame, as illustrated in Figure 1. To elicit paraphrasing, the subordinating conjunction *dat* "that" was presented in the right-hand frame after the finite verb of the main clause (cf. (2c)).

The practice sentences and fillers were in direct speech as well. Two out of five practice sentences contained reflexive pronouns, one of which was incorrect. One-third of the filler sentences contained a subject-verb agreement error or another inflectional error rendering the sentence ungrammatical. None of the fillers contained a reflexive pronoun.

The experiment started with the practice sentences, followed by the experimental sentences and the fillers in random order, with the restriction that no more than two experimental items would occur in consecutive sentences. To avoid possible learning effects, participants received each sentence in one condition only, with half of the sentences embodying condition Paraphrasing-cor-inp and the other half condition Paraphrasing-incor-inp. The assignment of these conditions to experimental sentences was rotated across participants.

Procedure. Participants were tested individually with the experimenter present, facing a computer screen positioned about 80 cm away and a microphone to register vocal response time. Reaction times were measured from the appearance of the sentence fragment on the screen until voice onset. Each sentence fragment was presented for 1,200 ms, with a 10 ms break between fragments. Immediately after each sentence the grammaticality

question was presented for 1,000 ms. After a 1,000 ms break the next sentence started. Stimulus presentation and reaction time measurements were controlled by NESU 2000 (Nijmegen Experiment Setup, developed at the MPI for Psycholinguistics in Nijmegen, documentation available from: <http://www.mpi.nl/world/tools/nesu/nesu2000.htm>).

Participants were instructed explicitly and by means of examples that their task was to paraphrase the sentences into indirect speech, and to rectify any other errors, such that the output would form a grammatical sentence. Furthermore, they were asked to judge the grammaticality of each input sentence by saying *yes* or *no* after the final input fragment. They were instructed to respond as fast as possible and to speak clearly.

Proofreading task

Participants. Fifteen members of the Leiden University community were paid to participate. All were native speakers of Dutch.

Materials. There were 12 experimental sentences, like (3): indirect-speech versions of the experimental sentences as in the Paraphrasing task. In addition, we constructed 24 fillers and five practice items. Each experimental sentence contained either a third-person singular reflexive pronoun that matched the antecedent (condition Proofreading-cor-inp; cf. (3a)), or a first-person reflexive which did not match the antecedent (condition Proofreading-incor-inp; cf. (3b)). The sentences were presented in fragments like (3c), and the desired output was as in (3d).

(3) a. Input in condition Proofreading-cor-inp

De lottowinnaar zei dat hij had besloten een rode auto te kopen voor zichzelf.

The lottery winner said that he had decided a red car to buy for himself.
 “The lottery winner said that he had decided to buy a red car for himself”.

b. Input in condition Proofreading-incor-inp

**De lottowinnaar zei dat hij had besloten een rode auto te kopen voor mezelf.*

The lottery winner said that he had decided a red car to buy for myself.
 “*The lottery winner said that he had decided to buy a red car for myself”.

c. Input fragments

De lottowinnaar/ zei dat/ hij/ had besloten/ een rode auto/ te kopen/ voor me/zichzelf.

d. Desired output in both Proofreading conditions

De lottowinnaar zei dat hij had besloten een rode auto te kopen voor zichzelf.

The lottery winner said that he had decided a red car to buy for himself.

“The lottery winner said that he had decided to buy a red car for himself”.

The practice sentences and fillers were also in indirect speech but did not contain reflexive pronouns. Three out of five practice sentences and half of the filler sentences contained a subject-verb agreement error or an idiomatic error rendering the sentence ungrammatical.

Participants saw half of the experimental items in condition Proofreading-cor-inp, the other half in Proofreading-incor-inp, according to a Latin square. The experiment started with the practice sentences, randomly followed by the experimental items and the fillers with the restriction that no more than two experimental items would occur consecutively.

Procedure. The experimental procedure was the same as in the Paraphrasing task. Participants were instructed to correct input errors as they read the sentences fragment by fragment and to respond as clearly and as quickly as possible.

Results

Participants with more than four voice key errors on the experimental items were excluded from analysis, leaving 12 participants in each task. All extreme RTs (shorter than 300 ms or longer than 1,000 ms) as well as RTs on incorrect response words, were discarded. This resulted in the removal of 13 data points (9% of the experimental items) for the Paraphrasing task and 19 data points (13%) for the Proofreading task, more or less evenly distributed across conditions and participants.

Latencies for conditions with input reflexives matching the desired output sentence (i.e., the Proofreading-cor-inp and Paraphrasing-incor-inp conditions) were on average 52 ms shorter than for conditions with reflexives that needed a change (Proofreading-incor-inp and Paraphrasing-cor-inp). Table 2 shows the mean latencies (averaged across participants and items) for each

TABLE 2
Experiment 1: Mean reaction times (milliseconds) on the reflexive pronouns in the Paraphrasing and Proofreading tasks

Input reflexive pronoun	Task and output reflexive pronoun	
	Proofreading: <i>himself</i>	Paraphrasing: <i>himself</i>
<i>myself</i> (Mismatch)	670 (incor-inp)	604 (cor-inp)
<i>himself</i> (Match)	607 (cor-inp)	645 (incor-inp)

Reflexive Pronoun Person condition. Recall that the shared-workspace hypothesis predicts a reversal of the difference between the RTs for cor-inp and incor-inp pronouns in the two tasks. This prediction is clearly confirmed: in Proofreading, the difference is 63 ms, in Paraphrasing it is -41.

We conducted two repeated-measures ANOVAs. The first one was a by-Subject analysis with Task (Paraphrasing vs. Proofreading) as between-Subjects factor and Reflexive Pronoun Match (matching in Proofreading-cor-inp and Paraphrasing-incor-inp; mismatch in Proofreading-incor-inp and Paraphrasing-cor-inp) as within-Subject factor. The second ANOVA was a by-Items analysis with Task and Reflexive Pronoun Match both treated as Within factors. Reflexive Pronoun Match yielded significant main effects both by-Subjects [$F_1(1, 22) = 8.844, p < .005$] and by-Items [$F_2(1, 11) = 8.914, p = .012$]. The main effect of Task and the interaction of Task and Reflexive Pronoun Match were not significant, neither by-Subjects nor by-Items ($p > .25$ in all analyses).

The grammaticality judgements yielded six (4%) incorrect answers in the Proofreading task and 36 (25%) in the Paraphrasing task. Corrected for chance, this means that the participants were able to correctly judge the (un)grammaticality of the last seen input in only 50% of the Paraphrasing sentences, whereas this was the case in 92% of the Proofreading sentences.

Discussion

Reaction times were considerably faster in conditions Proofreading-cor-inp and Paraphrasing-incor-inp where the input reflexive pronoun matched the desired output pronoun, than in conditions Proofreading-incor-inp and Paraphrasing-cor-inp where the input and output pronouns did not match. Stated differently, the input reflexive that was easy to process in Proofreading, was the difficult one in Paraphrasing (and vice-versa), as predicted by workspace sharing and inconsistent with the dedicated-workspaces hypothesis.

This picture is confirmed by the grammaticality judgement data. In about half of the sentences in the Paraphrasing task, participants could not determine correctly whether the input was grammatical or not, whereas they had little trouble doing this in the Proofreading task. We conclude that the results of Experiment 1 are in better agreement with the shared-workspace than with the dedicated-workspaces hypothesis.

The RT pattern predicted by the dedicated-workspaces hypothesis (see the introduction to Experiment 1 above) entails that the difference between the two Proofreading conditions (Proofreading-incor-inp vs. Proofreading-cor-inp: 670–607 ms) is larger than the corresponding difference between the Paraphrasing conditions (Paraphrasing-cor-inp vs. Paraphrasing-incor-inp: 645–604 ms). The fact that the obtained RT pattern conforms to this

prediction, might be interpreted as affording at least partial support for dedicated workspaces. However, another factor is probably responsible for the finding that correcting an unexpected pronoun is easier in the Paraphrasing than in the Proofreading task. While doing the Paraphrasing task, during every presentation of an experimental sentence, the participants read the reflexive pronoun after having modified the subject personal pronoun from first- to third-person just a few seconds ago. Performing this change may have rendered the resulting third-person agreement of the subject NP more salient/prominent/active in the Paraphrasing task than after only having read the subject pronoun in the Proofreading task. As a consequence, when the Paraphrasing participants hit upon a reflexive pronoun that failed to match the person of the subject NP, they needed less time to select the reflexive that did match (see also the Discussion of Experiment 2 below).

Another objection against our interpretation of the RT data proceeds from a trivial alternative explanation. Given the composition of the sentence materials used in the two tasks, the participants in both groups might have discovered that an ad hoc rule where grammatical expectations do not play a role, always generates the correct response: "If you read *myself*, retrieve *himself* and pronounce it; if you read *himself*, pronounce it immediately". The extra time needed to respond with *himself* could be an effect of response interference: responding to a word not with its own name but with another name. Although some of the participants may have used this nongrammatical strategy in a proportion of the sentence presentations, it cannot explain the complete pattern of results of the present experiment: It fails to provide a reason why the grammaticality judgements in the Paraphrasing task were much worse than in the Proofreading condition. Experiment 2 yields additional data that are at variance with such ad hoc strategies.

EXPERIMENT 2

The main goals of Experiment 2 were twofold. First, we wished to check whether the effects obtained in Experiment 1 generalise to an experimental setup where the to-be-DECODED input sentences rather than the to-be-ENCODED output sentences are kept constant. That is, instead of presenting the participants with direct-speech complement clauses in the Paraphrasing task and with indirect-speech complements in the Proofreading task, the present experiment uses direct-speech input clauses in both tasks. Second, when preparing Experiment 1, we had rejected the option of a design with Task as a within-Subjects variable because we feared that the participants might get confused and mix up the two tasks. In spite of this risk, the present experiment stages a design where each participant performs both the

Paraphrasing and the Proofreading task—not intermingled but in blocks, with order of blocks counterbalanced.

Additionally, Experiment 2 implements a number of smaller-scale improvements. First, we provided the participants with feedback on their speed of responding, expecting that this would lead to considerably shorter mean RTs and reduce the likelihood of ad hoc response strategies. Second, we asked content questions instead of grammaticality judgements, thereby emphasising reading for comprehension. Third, we composed a larger set of experimental (36) and filler (90) sentences. In both the Paraphrasing and the Proofreading task, the fillers had the same overall syntactic structure as the experimental items, except for the fact that they contained a nonpronominal NP instead of a reflexive pronoun. Due to this homogeneity, the participants could not predict whether or not the current sentence would contain a reflexive. Fourth, the symbol “####” instead of *that* served as trigger of the direct-to-indirect-speech conversion, thus ensuring that the subordinating conjunction would have to be retrieved from the mental lexicon as part of the encoding process and did not belong to the decoded input string. Fifth, we added a SPR task with the same input sentences that were presented during the Proofreading and Paraphrasing tasks. Because SPR does not require any verbal response to be grammatically encoded, it enabled a test of the assumption that sentences with ill-formed reflexive pronouns are more difficult to decode grammatically than their counterparts with well-formed reflexives. Finally, the presentation of a new sentence fragment was initiated immediately after the apparatus had registered offset of the vocal response to the last fragment; thus we avoided measurement errors due to very slow responses that intruded into the next trial and triggered premature RT registrations there. In all other respects, the method used in Experiment 2 was identical to that of Experiment 1.

The pattern of RTs predicted by the two processing models is the same as in Experiment 1. That is, the shared-workspace hypothesis expects

$$\text{Proofreading-cor-inp} \approx \text{Paraphrasing-incor-inp} < \text{Proofreading-incor-inp} \\ \approx \text{Paraphrasing-cor-inp},$$

whereas the dedicated-workspaces model forecasts

$$\text{Proofreading-cor-inp} < [\text{Paraphrasing-cor-inp}, \text{Paraphrasing-incor-inp}] \\ < \text{Proofreading-incor-inp}.$$

See Table 3, and Tables A3 and A4 of the Appendix, for details. Notice that, due to the use of direct-speech input sentences in the Proofreading task (invariably with first-person subject NPs), the correct output reflexives differ between tasks (first-person in Proofreading, third-person in Paraphrasing).

TABLE 3

Experiment 2: Sources of response delay predicted by the dedicated-workspaces and the shared-workspace hypotheses. The numbers refer to sample sentences mentioned in the Materials subsection. For explanation see Tables A3 and A4 in the Appendix

Input reflexive pronoun	Task and correct output reflexive pronoun	
	Paraphrasing: <i>himself</i>	Proofreading: <i>myself</i>
	Dedicated workspaces	
<i>myself</i>	REPAIR (cor-inp: 4a)	— (cor-inp: 5a)
<i>himself</i>	VIOLATION (incor-inp: 4b)	VIOLATION + REPAIR (incor-inp: 5b)
	Shared workspace	
<i>myself</i>	VIOLATION + REPAIR (cor-inp: 4a)	— (cor-inp: 5a)
<i>himself</i>	— (incor-inp: 4b)	VIOLATION + REPAIR (incor-inp: 5b)

The use of direct-speech instead of indirect-speech input sentences in both tasks also explains further differences between Table 1 and 3.

Method

Participants

Thirteen participants took part in the experiment. They all were students at the Radboud University in Nijmegen and were paid for participation. All were native speakers of Dutch. The data of one participant could not be analysed because of a temporary failure of the technical equipment. The remaining 12 participants were assigned to one of two groups of six participants according to a random scheme, one group taking the tasks in the order Proofreading–Paraphrasing–SPR, the other group in the order Paraphrasing–Proofreading–SPR.

Materials

A total of 126 Dutch sentences was created, 36 of which were experimental items like (4a) or (4b). The remaining 90 items were fillers. The experimental sentences were nearly identical with respect to grammatical structure: They embodied the same string of 15 or 16 word categories (the 15-word sentences contained one NP without an article), and their surface phrase structures were identical. They all included a direct-speech clause, in both the Paraphrasing and the Proofreading task, and were presented in 11 fragments. The reflexive pronoun was the eighth fragment in Proofreading blocks and, due to the insertion of the “####” symbol, the ninth fragment in Paraphrasing blocks.

- (4) a. Input in condition Paraphrasing-cor-inp
De ijdele artiest beaamde: “Ik ben begonnen een portret van mezelf op de wand te schilderen”.
 “The conceited artist confirmed: ‘I have started to paint a portrait of myself on the wall’”.
- b. Input in condition Paraphrasing-incor-inp
 **De ijdele artiest beaamde: “Ik ben begonnen een portret van zichzelf op de wand te schilderen”.*
 “*The conceited artist confirmed: ‘I have started to paint a portrait of himself on the wall’”.
- c. Input fragments
De ijdele artiest| beaamde:| [#####]| “Ik| ben begonnen| een portret| van me|zichzelf| op de wand| te schilderen”.
- d. Desired output in both Paraphrasing conditions
De ijdele artiest beaamde dat hij was begonnen een portret van zichzelf op de wand te schilderen
- (5) a. Input in condition Proofreading-cor-inp
De ijdele artiest beaamde: “Ik ben begonnen een portret van mezelf op de wand te schilderen”.
- b. Input in condition Proofreading-incor-inp
 **De ijdele artiest beaamde: “Ik ben begonnen een portret van zichzelf op de wand te schilderen”.*
- c. Input fragments
De ijdele artiest| beaamde:| “Ik| ben begonnen| een portret| van me|zichzelf| op de wand| te schilderen”.
- d. Desired output in both Proofreading conditions
De ijdele artiest beaamde: “Ik ben begonnen een portret van mezelf op de wand te schilderen”.

The filler sentences also included a direct-speech clause, but the number of fragments varied between nine and 11. In the experimental items, the only to-be-corrected errors were reflexive pronouns (*mezelf/zichzelf* “myself/himself”). In both tasks, input sentences with *mezelf* embodied the “cor-inp” condition, and items with *zichzelf* the “incor-inp” condition. The filler items did not contain reflexive pronouns. Instead, 36 of them had a direct-object NP introduced by a possessive pronoun: *mijn* “my” (18), *zijn* “his” (9) or *haar* “her” (9). In the Paraphrasing task, *mijn* always had to be corrected to *zijn* “his” or *haar* “her” to maintain person agreement with the subject of the indirect-speech clause. The other input possessive pronouns (*zijn* and *haar*) never needed to be replaced. Although all reflexive pronouns were preceded by a preposition, we took care that no preposition could serve as a reliable predictor of a reflexive pronoun in the next fragment (e.g., *van* “of” and *voor* “for” could precede nonpronominal as well as pronominal NPs). The experimental and filler sentences did not include any errors other than the

ones mentioned here. We composed six experimental lists in which all experimental items were rotated across condition (cor-inp/incor-inp reflexive pronoun) and across tasks (Proofreading, Paraphrasing, and SPR). Within each list, experimental items alternated with filler items according to a pseudo-random design. Participants saw each of 12 experimental items in one of the three tasks, and in one experimental condition only (six with cor-inp reflexives, and six with incor-inp reflexives; as there were three tasks, we needed 36 experimental sentences). Data from six participants were required for a complete design in which each item was shown in all three tasks with both a cor-inp and an incor-inp reflexive pronoun.

Procedure

The course of events in the Paraphrasing and the Proofreading tasks was the same as in Experiment 1, except for the procedural differences listed above in the Introduction to Experiment 2. As for the response speed feedback, if participants did not respond to a fragment within 1,000 ms, a large red exclamation mark was flashed on the screen for 150 ms. We assumed that the presence of this sign would press for faster responding in subsequent trials. The exclamation mark was immediately followed by the next sentence fragment.

An experimental session comprised three trial blocks: Proofreading–Paraphrasing–SPR or Paraphrasing–Proofreading–SPR. In the third block (SPR), participants were instructed to read the sentences silently, fragment by fragment and to press a button as quickly as possible to reveal the next fragment. Measures of interest were the reading times per fragment, i.e., the intervals between successive button presses.

To ensure that participants read for comprehension, all three tasks were interrupted from time to time by a question—one per six sentences on average—about the content of the sentence just seen (always a filler). Participants were supposed to respond with *yes* or *no* by means of a button press (in all three blocks). The SPR task was identical to the Proofreading task, except that participants were instructed to read the fragments silently and to trigger the presentation of a new fragment by means of a button press. Participants were asked to read fluently and to read for comprehension. Button press latencies were measured after each input fragment.

Results

Self-paced reading (SPR)

We first analysed the SPR data in order to check whether, given the sentence materials used in the Proofreading and Paraphrasing tasks, an incorrect input pronoun indeed caused a processing delay that cannot be

ascribed to grammatical encoding but must be located in the grammatical decoding process. Because we expected the effect of ungrammatical input fragments to be delayed and to spill over into subsequent input fragments, we averaged the RTs for the three fragments following the reflexive pronoun (no effect was visible at the pronoun itself). We did this separately for sentences with input-correct and input-incorrect reflexives. The resulting means, averaged over participants, were 582 (cor-inp) and 634 (incor-inp) ms. The 52 ms difference is significant ($t=2.56$, $df=11$, one-tailed $p=.014$). We conclude that the SPR data provide independent evidence that, in the experimental items we used, the incorrect reflexive pronouns gave rise to a time-consuming grammatical decoding problem that cannot be attributed to grammatical encoding.

Testing the central hypothesis

The proportion of erroneous responses to the reflexive pronoun fragments (i.e., incorrect choice of pronoun) amounted to 8% in the Proofreading blocks, and 15% in the Paraphrasing blocks. These data were discarded. Table 4 presents the RTs, separately for the two tasks and the two trial blocks.

We entered the remaining RTs into a by-Subject analysis of variance with Order of Tasks (Paraphrasing–Proofreading vs. Proofreading–Paraphrasing) as between-Subjects factor, and Task (Paraphrasing vs. Proofreading) and Reflexive Pronoun Person (cor-inp/*myself* vs. incor-inp/*himself*) as within-Subject factors. (A by-Items analysis could not be conducted because, after removing the errors, not all items had observations in all cells—due to the fact that each item was presented only twice per combination of Task and Reflexive Pronoun Person.) Order of Tasks turned out to be a highly significant main effect: The participants who started with the Proofreading task on average had slower RTs than the participants who started with Paraphrasing, $F(1, 10) = 7.11$, $p = .02$. The two other main effects were not

TABLE 4
Experiment 2: Mean reaction times (milliseconds) on the reflexive pronouns in the Paraphrasing and Proofreading tasks. Correct output reflexive in parentheses

Input reflexive pronoun	Task and output reflexive pronoun			
	Trial block 1		Trial block 2	
	Proofreading: (<i>myself</i>)	Paraphrasing: (<i>himself</i>)	Proofreading: (<i>myself</i>)	Paraphrasing: (<i>himself</i>)
<i>myself</i> (cor-inp)	579	504	460	612
<i>himself</i> (incor-inp)	660	453	494	629

TABLE 5
 Experiment 2: Differences (milliseconds) between RTs to incor-inp (*himself*) and cor-inp (*myself*) reflexive pronouns in first and second halves of the two trial blocks: RT(*himself*)-RT(*myself*)

Task	Block 1		Block 2	
	First half	Second half	First half	Second half
Proofreading	89	73	1	67
Paraphrasing	-42	-60	11	-43

significant ($p > .17$ in both analyses). The shared-workspace hypothesis predicts an interaction between Task and Reflexive Pronoun Person: In Proofreading, RTs to *myself* (cor-inp) should be faster than RTs to *himself* (incor-inp), but the difference should be reverse in Paraphrasing. The interaction is indeed significant, $F(1, 10) = 7.15$; $p = .02$. However, the predicted reversal only obtained in the first trial block⁹ (see Table 4). In the second block of trials, the reversal failed to materialise, presumably due to a counteracting factor that we will discuss below in the context of Table 5.

None of the other interactions was significant ($p > .15$ in all analyses). At the same time, inspection of Table 4 reveals that the RT averages are at variance with dedicated-workspaces hypothesis, which predicts that the two Paraphrasing averages should be situated within the range spanned by the Proofreading averages.

Discussion

The first-block data in Table 4 clearly show the between-task reversal of the RT pattern, in line with the prediction derived from workspace sharing. However, additional factors exerted their influence as well. Consider the large difference between the RTs to the incor-inp reflexive *himself* and the cor-inp reflexive *myself* in the participants who started with the Proofreading task. This difference amounted to 81 ms, whereas the participants who started with the Paraphrasing task needed only 51 ms extra to replace *myself* by *himself* in comparison with repeating *himself*. Although we can only speculate, we suggest that this difference is caused by two factors conspiring to make retrieval of the person agreement of the reflexive's antecedent more difficult in the Proofreading-as-first-task group than in the

⁹The crucial RT pattern—the reversal of the RT difference between cor-inp and incor-inp pronouns in the two tasks—already obtains in early trials. The mean RTs for the FIRST HALF of the first trial block executed by the participants are as follows: Proofreading-cor-inp: 578 ms; Proofreading-incor-inp: 667 ms; Paraphrasing-cor-inp: 506 ms; Paraphrasing-incor-inp: 464 ms. They are similar to the corresponding means for the first trial block as a whole (Table 4).

Paraphrasing-as-first-task group. First, as in Experiment 1, the fact that the participants in the former group only needed to read the subject NP of the complement clause, may have made the person agreement features of the antecedent less salient than it was for the participants, in the latter group who had to convert the subject NP's person feature to third-person. Second, this effect may have been aggravated by the discrepancy between first-person agreement of the reflexive pronoun's grammatical antecedent *ik* "I" of the complement clause and third-person agreement of the subject NP of the main clause—i.e., the referent of *ik*.¹⁰ The combination of these factors may have made the participants who started with Proofreading insecure about how to react to the incor-*inp* reflexive *himsel* during trial block 1.

Another puzzling aspect of the RT pattern, already noted above, is the fact that in the second block of trials the between-pronoun differences in the two tasks failed to show the reversal predicted by the shared-workspace model. In search for an explanation, we computed average RTs for the first and the second halves of each trial block separately. Table 5 shows the between-pronoun differences as they developed in the course of an experimental session. The Proofreading task gave rise to positive values, indicating that grammatically correct input reflexives (*myself*) were easier than grammatically incorrect ones (*himsel*). Negative values in the Paraphrasing task embody the reversal predicted by the shared-workspace hypothesis. Remarkable is the "neutralisation" of the between-pronoun RT difference in the first half of the second block of trials. We suggest that this is due to a perseveratory tendency to apply the old task instruction instead of the new one. The tendency apparently affected a substantial proportion of trials shortly after the task switch but subsided in the second half of the post-switch trial block, thus reinstating the predicted pattern.

In the Discussion of Experiment 1, we brought up an alternative hypothesis for the main findings: The participants might have discovered an ad hoc rule which does not presuppose any syntactic expectations about upcoming constituents. Applied to the conditions of Experiment 2, this rule could be as follows: "In Proofreading, if you see *himsel*, retrieve *myself* and pronounce it; if you see *myself*, pronounce it immediately". In the Paraphrasing task, the rule would apply with *himsel* and *myself* exchanged. However, the following empirical argument renders this interpretation in terms of strategic behaviour highly unlikely.

The reflexive pronoun was not the only one that allowed an ad hoc rule. The fragment containing the subject NP of direct-speech clause (*ik* "I") is the case in point. In the Proofreading task, the participants could pronounce this

¹⁰Note that this second factor was absent from Experiment 1, where the Proofreading task was performed with indirect-speech complement clauses whose subject NP had third-person agreement.

word unaltered whereas they had to change it to *hij/zij* “he/she” in the Paraphrasing task. Given the composition of the sentence materials, the participants might have discovered early on that *ik* always followed the main clause fragment with verb and colon. So, it was perfectly predictable on the basis of the immediately preceding input fragment, and the participants could, in principle, have responded to *ik* without decoding it in detail. Similarly, in the Paraphrasing task the participants could have noticed that the response to *ik* would be *hij* (or *zij* in a minority of sentences) and that grammatically decoding *ik* was not necessary. Hence, given this (nearly) perfect predictability and on the assumption of strategic behaviour, one would expect the mean RT to *ik* in the Paraphrasing task to be similar to the one in the Proofreading task, or somewhat longer due to the fact that in the former task the first-person feature of the Subject had to be changed to third-person. However, in contrast to this prediction, the average RT to *ik* in Paraphrasing (424 ms) was much shorter than in Proofreading (549 ms)—the difference amounting to 125 ms. (The mean RTs to *ik* for sentences with *cor-inp* and *cor-inp* reflexives further downstream were nearly identical.) Perhaps even more telling, in the Proofreading task the average latency to *ik* was almost the longest one of all input fragments belonging to the direct-speech clause, whereas it was by far the shortest in Paraphrasing. This data pattern speaks strongly against strategic behaviour based on ad hoc rules but fits well with what one expects on the basis of normal grammatical decoding processes. In the Proofreading task, the subject *ik* requires starting up a new finite clause; in Paraphrasing, this work is done during the preceding trial in the context of processing the subordinating conjunction. Moreover, to the extent that the decoding process uses transitional probabilities based on past parsing experience (rather than on contingencies holding only for the current stimulus set), the RTs mirror the fact that, after a sentence-initial nominative NP and a finite verb (e.g., *The headmaster complained.*), the likelihood of yet another nominative NP (*ik/hij*) is much lower than a nominative NP following a subordinating conjunction (*dat*).

GENERAL DISCUSSION

In the two experiments reported here, the participants performed an incremental grammatical decoding and an incremental grammatical encoding task concurrently. The decoding task consisted of reading and interpreting a sentence that was presented fragment-by-fragment; the encoding task required the participants to paraphrase the input sentence by responding to each input fragment with a fragment of the paraphrase. Some input fragments could simply be repeated in the output, but others had to be transformed into a semantically equivalent but morpho-syntactically

different output fragment that would fit into the emerging paraphrase. This paradigm enabled us to pit the shared-workspace and the dedicated-workspaces hypotheses against one another: The latter predicts that the decoder's expectations regarding upcoming input are solely determined by the structure of the input sentence and are immune to influences from the structure of the paraphrase being encoded concurrently. The former, on the other hand, entails that the structure underlying the encoded fragments overwrites the structure of the corresponding input fragments, so that henceforth decoder expectations derive from the structure of the paraphrase. This evidences that the two modalities of grammatical performance are not completely independent: When processing the same communicative intention concurrently, they work with one and the same (more precisely: token-identical) linguistic expression of that intention. This suggests that the grammatical encoding and decoding processes command the same, shared workspace for the assembly and short-term storage of grammatical forms. For, on the dedicated-workspaces hypothesis, one would expect each modality to store its "own" grammatical representation of the communicative intention in its own dedicated workspace. This, in turn would have led to different expectancies regarding the person value of upcoming reflexive pronouns than attested in the two experiments. The fact that, in a concurrent decoding and encoding task like the Paraphrasing task staged in the experiments, language users spontaneously opt for fusing the decoded with the encoded grammatical structure instead of keeping the two apart, even when executing the task for the first time (cf. footnote 4), strongly suggests that, in overlapping time, the decoding and encoding processes generate and manipulate the same, token-identical output structure.

This preference does not rule out entirely the possibility that the two modalities keep decoded and encoded grammatical structures apart. A case in point is the translation task called simultaneous interpreting, which forces the grammatical workspace to deal concurrently with two grammatically different structures expressing the same communicative intention—in *casu* the decoded source language sentence and the encoded target language translation. For these sentences, the workspace needs to create separate "threads", and to switch back and forth between them: "multithreading" (Salvucci & Taatgen, 2008). Consider an interpreter who hears Dutch sentence (6a) and translates it into (6b), starting immediately after having heard and decoded the subject NP (as indicated by the dash). Because the plural noun *hersenen* of the Dutch subject NP translates into a singular English subject NP headed by *brain*, the shared workspace now contains both the Dutch and the English subject NPs.

- (6) a. *De hersenen_{PL} van een baby / zijn_{PL} nog in ontwikkeling*
 The brain of a baby is still in development
 b. “The brain of a baby / is still developing”

The next processing steps consist of decoding the Dutch VP and encoding the English VP. In order to deal correctly with the differing values of the verb’s number feature, the two translation equivalents are to be treated as separate processing threads, and switching between decoding and encoding here requires multithreading—switching between threads, i.e., translation equivalents. Multithreading draws heavily on efficient and flexible management of the storage capacity of the grammatical workspace. In a dedicated-workspaces architecture, the two translation equivalents are kept in different workspaces, with as a consequence that the multithreading problem does not arise, and no processing costs due to multithreading are expected. So, the shared-workspaces model, in a sense, explains the high cognitive load imposed by simultaneous translation tasks.¹¹

In the remainder of this section, we turn to the question whether our conclusion in favour of the shared-workspace architecture meshes with other known facts about relationships between grammatical encoding and decoding.

Neurocognitive data

A shared workspace for grammatical encoding and decoding is compatible with neurophysiological evidence indicating that the same neural circuits, including the Left Inferior Frontal Gyrus and the Left Posterior Temporal Gyrus, subserve grammatical structure formation in both sentence comprehension and sentence production (Hagoort, 2005; Hagoort, Brown, & Osterhout, 1999; Snijders et al., 2009; Ullman et al., 2005).

Self-monitoring

Could the fact that speakers monitor and repair self-produced language utterances be construed as evidence in favour of a dedicated-workspace architecture? Levelt’s (1983, 1989) theory of self-monitoring is based on the “Perceptual Loop”: sentences delivered by the grammatical encoding mechanism can be perceived and analysed by the speaker’s own grammatical decoder—via an outer loop (spoken utterances) or an inner loop (inner speech). When the decoder spots an ill-formed word string, or when the communicative intention derived from the decoded utterance does not match the communicative intention that served as input to the encoding process, the

¹¹For a recent experimental assessment of cognitive skills required by these tasks, see Christoffels, De Groot, and Kroll (2006).

decoder issues an error/warning message. The phenomenon of self-monitoring thus is a plausible argument in favour of dedicated workspaces for grammatical encoding and decoding.

The error message leads the encoder to re-encode (part of) the current utterance. The overt realisation of the resulting repair text is then decoded, and the grammatical structure derived from it henceforth replaces, in the decoder's workspace, the structure underlying the erroneous (part of the) string. Thus aligned, the encoding and the decoding workspaces contain identical grammatical structures for the current (incomplete) sentence, and the two modalities can continue as if no error had occurred. By way of example, consider a speaker who repairs a false start as follows: *Since my son . . . uh . . . my sons have an iPad, they only read ebooks*. If the grammatical decoder, after processing *my son* and deciding that this referring expression does not square with the current communicative intention, fails to adjust the number feature of the subject NP in response to the repair *my sons*, it would still expect singular finite verbs and personal pronouns. Obviously, the encoded structure underlying a repair should be capable of influencing the decoder's expectations and decisions. In a dedicated-workspaces architecture, the inner or outer perceptual loop paves the way for this influence. We say that the perceptual loop enables ENCODING-TO-DECODING CONTENT MIRRORING. Obviously, without this mechanism, the decoder would not be able to continue monitoring correctly after a repair.¹²

However, self-monitoring and self-repair are by no means impossible in a shared-workspace architecture, provided the process of grammatical encoding meets the following criteria: (1) it is an OPTIMISATION PROCESS IN WHICH STRUCTURES UNFOLD OVER TIME, EMBODYING INCREASINGLY MANY OF THE PREVAILING CONSTRAINTS,¹³ and (2) at any point in time during optimisation, the current structure can be read out for Phonological

¹² Is there also content mirroring in the opposite direction (decoding-to-encoding)? When a speaker has "talked him/herself into a corner" and the listener has a more or less accurate idea of the communicative intention the speaker was going to express, the listener is able to join in rapidly, continuing the utterance in a grammatically well-formed fashion. This phenomenon indicates that grammatical structure emerging from decoding other-produced speech can be made available to the encoding mechanism without much delay. However, to our knowledge, no data are available concerning the time lag between speech offset by the speaker and speech onset by the interlocutor. This means that the latter, based on his/her reconstruction of the speaker's communicative intention and on short-term recollection of lexical and syntactic properties of the interrupted utterance, may have quickly re-generated the grammatical structure of the interrupted sentence and planned his/her utterance as a continuation of this structure. Consequently, this phenomenon cannot count as evidence in favour of decoding-to-encoding content mirroring in a dedicated-workspaces architecture.

¹³ We presuppose that grammaticality (or acceptability) of a string is a continuous rather than a discrete (binary) variable—in line with much recent evidence for "gradient grammaticality" (Bard, Robertson, & Sorace, 1996; Fanselow, Féry, Vogel, & Schlesewsky, 2006).

Encoding (i.e., the next stage of sentence production after Grammatical Encoding) and overt pronunciation. If, at the moment of read-out, the quality of the grammatical structure has not yet reached an optimum (“premature delivery”), syntactic or semantic errors may become manifest (as yet undetected by any language processing component). In the meantime, the optimisation process continues, which may give rise to modified or extended input for the phonological encoder, possibly to amendments in the string of word forms delivered previously. In case of the latter, the grammatical encoder may be said to have “detected” an error. Then, the amended string is processed phonologically and phonetically from a point preceding the leftmost amendment onward, thus yielding a repair (as in the example *Since my son . . . uh . . . my sons*). Crucially, in contrast to what the perceptual loop theory postulates, this dynamic self-monitoring process does not involve grammatically decoding a word form string that has just been grammatically encoded. In this scenario, the monitoring process is “production-internal” (Postma, 2000), and one grammatical workspace suffices.¹⁴ This cursory sketch of self-monitoring and self-repair as a dynamic process of grammatical structure formation indicates that the idea of a shared workspace accessible to both modalities of grammatical performance does not rule out self-monitoring.

Our view of self-monitoring and self-repair assumes that the process of structure optimisation which underlies grammatical encoding, extends over (short) time intervals, and that during such intervals the phonological encoding process can read out the encoded grammatical structures. In such a system, phonological encoding and overt production of a sentence or sentence increment may be based on early, not yet optimal structures which improve further in the course of subsequent optimisation steps. Such improvements may engender substitutions in the word form string delivered by phonological encoding. If a substitution affects word forms that have already been produced overtly, or are just being pronounced, an overt repair action ensues. That is, ongoing speech is suspended immediately, the speaker retraces to the point in the phonological output stream where the substituted and the substituting word form strings begin to diverge, and resumes the utterance from there. Importantly, this scenario entails that no error correction (replanning) phase is necessary following error detection: The very availability of the repair text triggers the repair action, and the repair text can be uttered immediately after speech suspension, i.e., with suspension-to-resumption time intervals of zero milliseconds. In contrast, the

¹⁴An important principle of structure formation meeting the two above criteria is SELF-ORGANIZATION. For computationally explicit examples, see Vosse and Kempen (2000, 2008, 2009) and Tabor & Hutchins (2004). See also Konieczny, Müller, Hachmann, Schwarzkopf, and Wolfer (2009).

perceptual loop theory assumes that the process of error correction sets in only after the decoder has detected an error and initiated suspension of ongoing speech. This “main interruption rule” was recently falsified by Seyfeddinipur, Kita, and Indefrey (2008) who observed that “speakers interrupted themselves not at the moment they detected the problem but at the moment they were ready to produce the repair” (p. 837). This is exactly what one expects if self-monitoring proceeds by way of structure optimisation, with error detection being the consequence rather than the cause of error correction. Seyfeddinipur et al. describe the behaviour of their speakers as resulting from a strategic decision to maximise output fluency (by avoiding a long pause after suspension of speech). In an optimisation system as envisaged here, it is an emergent property.

We hasten to add, though, that the scenario sketched here is not the only possible one. At least two other modes of self-monitoring are possible, both yielding positive suspension-to-resumption intervals. One is based on the perceptual loop: Speakers can decode their own utterances by “replaying” a phonological or acoustic representation stored in a short-term buffer—for instance, in response to cues from a listener, or during a speech pause caused by extraneous factors. The other mode can be initiated by speakers who midsentence “change their mind”. They suspend ongoing speech, feed a modified communicative intention into the grammatical encoder, retrace to a point preceding the leftmost revised constituent, and resume speech from that point onward.

Speech shadowing

We now turn to processing differences between close and distant shadowers. Distant shadowers are able to repeat spoken prose with ear-voice spans in the range of 500–1,500 ms. Close shadowers gain ear-voice spans of less than 300 ms, even spontaneously rectifying input errors (e.g., incorrect inflectional endings deliberately introduced by the experimenter). These latencies are not much longer than the time normal speakers need in order to repeat isolated words: typically between 150 and 250 ms (Marslen-Wilson, 1973, 1985). A shared-workspace architecture may be required to enable this feat. Suppose that structure optimisation characterises not only grammatical encoding (as discussed above) but also grammatical decoding. Furthermore, assume that in close shadowers these optimisation processes yield (near-)optimal structural solutions at such a high speed that the phonological encoding component can read them out virtually flawlessly when the rate of the perceived speech is in the normal range. Perhaps we need to assume, in addition, that the speed of phonological encoding is faster in close shadowers than in distant shadowers.

Given a shared-workspace architecture, speakers in whom these conditions are met can use the online DECODED grammatical structures directly to drive phonological encoding. This affords short ear-voice spans and spontaneous correction of input errors. A dedicated workspace architecture, on the other hand, does not allow the phonological encoder to access grammatically decoded structures and forces a time-consuming detour via grammatical encoding: feeding the communicative intention reconstructed from perceived input into the grammatical encoder, computing the grammatical structure underlying the input a second time (now as part of grammatical encoding), and feeding the resulting terminal word form string into the phonological encoder.

On the shared-workspace hypothesis, if grammatical decoding and/or phonological encoding are not fast enough to keep pace with the input, distant shadowing ensues due to a more complicated course of events. As phonological encoding and overt pronunciation begins to lag behind input, the activation of the decoded grammatical structures decays more and more, soon reaching a level too low to drive phonological encoding. From that point onward, a second decoding round is required for the tail of the input string (i.e., the substring following the last pronounced word). This presupposes the shadower can retrieve a trace of that substring from a sensory or phonological buffer and feed it into the decoder.

Turn-taking in dialogue

Finally, does the fact that conversation partners sometimes speak during overlapping time intervals (Sacks, Schegloff, & Jefferson, 1974; Stivers et al., 2009) dovetail with a shared-workspace architecture for grammatical encoding and decoding? The answer is definitely positive. First, the two modalities can use the workspace alternately, on a time-division basis. Second, and more important, recent empirical evidence shows that temporal overlap between dialogue turns is rather marginal. It is largely restricted to the final 1–2 seconds of a turn (Stivers et al., 2009). On the reasonable assumption that, towards the end of a turn, the final words are virtually fully predictable, the conversation partner need not listen to the speaker anymore and, instead of decoding the final part of the turn, can begin encoding a reply. In conclusion, the phenomenon of overlapping dialogue turns fails to provide a strong argument in favour of a dedicated-workspaces architecture.

Envoi

Our multitasking experiments suggest that grammatical encoding and grammatical decoding are subserved by a shared workspace for the assemblage and temporary storage of grammatical forms. Additional

observations reported in the psycholinguistic literature indicate that the resemblance between the two modalities of grammatical performance does not stop here; the overlap between their cognitive processing resources may be considerable, in fact. Because a shared-workspace architecture thus appears compatible with what we know about the processes of grammatical encoding and decoding, the option of a single “grammatical coder” deserves serious attention in future research. This calls for closer cooperation between the language production and language comprehension research communities.

Manuscript received 5 April 2010

Revised manuscript received 24 November 2010

First published online 26 July 2011

REFERENCES

- Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 72, 32–68.
- Christoffels, I. K., De Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, 54, 324–345.
- Fanselow, G., Féry, C., Vogel, R., & Schlesewsky, M. (Eds.). (2006). *Gradience in grammar: Generative perspectives*. Oxford, UK: Oxford University Press.
- Garrett, M. (2000). Remarks on the architecture of language processing systems. In Y. Grodzinsky, L. Shapiro, & D. Swinney (Eds.), *Language and the brain: Representation and processing*. San Diego, CA: Academic Press.
- Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, 9, 416–423.
- Hagoort, P., Brown, C., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language*. Oxford, UK: Oxford University Press.
- Kempen, G. (2000). Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral and Brain Sciences*, 23, 38–39.
- Konieczny, L., Müller, D., Hachmann, W., Schwarzkopf, S., & Wolfer, S. (2009). Local syntactic coherence interpretation. Evidence from a visual world study. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the Cognitive Science Society*, July/August, Amsterdam (pp. 1133–1138). Austin, TX: Cognitive Science Society.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41–104.
- Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
- Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244, 522–523.
- Marslen-Wilson, W. D. (1985). Speech shadowing and speech comprehension. *Speech Communication*, 4, 55–73.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11, 105–110.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77, 97–131.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn taking for conversation. *Language*, 50, 696–735.

- Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*, 101–130.
- Seyfeddinipur, M., Kita, S., & Indefrey, P. (2008). How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition*, *108*, 837–842.
- Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*, 1493–1503.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10587–10592.
- Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 431–450.
- Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
- Ullman, M. T., Pancheva, R., Love, T., Yee, E., Swinney, D., & Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language*, *93*, 185–238.
- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105–143.
- Vosse, T., & Kempen, G. (2008). Parsing verb-final clauses in German: Garden-path and ERP effects modeled by a parallel dynamic parser. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society*, July 2008, Washington, DC. Austin, TX: Cognitive Science Society.
- Vosse, T., & Kempen, G. (2009). The unification space implemented as a localist neural net: Predictions and error-tolerance in a constraint-based parser. *Cognitive Neurodynamics*, *3*, 331–346.

APPENDIX. DERIVATION OF THE PREDICTED RT PATTERNS

Tables A1–A4 spell out the essential workspace contents and processing steps hypothesised in the dedicated-workspaces and the shared-workspace models, for the tasks performed in Experiments 1 and 2. The bottom row of each Table lists the sources of response delay that are also mentioned in Table 1 and 3. An expectancy violation ensues whenever the person feature of the input reflexive does not agree with the person feature predicted from the current representation of the input sentence in the “relevant” workspace. According to the dedicated-workspaces model, such predictions are based on the grammatical structure residing in the workspace commanded by the grammatical decoding process; the shared-workspace model holds that these predictions are based on the structure computed by the encoding process, which erases and modifies (overwrites) the grammatical structure computed during an earlier decoding process for the same communicative intention. A repair is required whenever the reflexive pronoun that is required in the overt output sentence, differs from the one presented in the input sentence.

TABLE A1
Dedicated-workspaces hypothesis: derivation of the predicted RT pattern for Experiment 1 (summarised in Table 1)

	Paraphrasing		Proofreading	
	(2a) I ... myself	(2b) I ... *himself	(3b) he ... *myself	(3a) he...himself
<i>Input sentence</i>				
Person feature of input reflexive	First person	Third person	First person	Third person
Predicted reflexive, given the content of the <i>decoding</i> workspace	First person	First person	Third person	Third person
Expectancy violation during the decoding process?	No	Yes	Yes	No
<i>Output sentence</i>	(2d) <i>he ... himself</i>	(2d) <i>he ... himself</i>	(3d) <i>he ... himself</i>	(3d) <i>he...himself</i>
Predicted reflexive, given the content of the <i>encoding</i> workspace	Third person	Third person	Third person	Third person
Repair needed during the encoding process? Source(s) of response delay	Yes REPAIR	No VIOLATION	Yes VIOLATION + REPAIR	No —

TABLE A2
 Shared-workspace hypothesis: derivation of the predicted RT pattern for Experiment 1 (summarised in Table 1)

	Paraphrasing		Proofreading	
	(2a) <i>I ... myself</i>	(2b) <i>I ... *himself</i>	(3b) <i>he ... *myself</i>	(3a) <i>he ... himself</i>
Input sentence				
Person feature of input reflexive	First person	Third person	First person	Third person
Predicted reflexive, given the content of the <i>shared</i> workspace	Third person	Third person	Third person	workspace
Expectancy violation during the decoding process?	Yes	No	Yes	No
Output sentence	(2d) <i>he ... himself</i>	(2d) <i>he ... himself</i>	(3d) <i>he ... himself</i>	(3d) <i>he ... himself</i>
Predicted reflexive, given the content of the <i>shared</i> workspace	Third person	third person	Third person	Third person
Repair needed during the <i>encoding</i> process?	Yes	No	Yes	No
Source(s) of response delay	VIOLATION + REPAIR	—	VIOLATION + REPAIR	—

TABLE A3
 Dedicated-workspaces hypothesis: derivation of the predicted RT pattern for Experiment 2 (summarised in Table 3)

	Paraphrasing		Proofreading	
<i>Input sentence</i>	(4a) I ... myself	(4b) I ... *himself	(5a) I ... myself	(5b) I ... *himself
Person feature of input reflexive	First person	Third person	First person	Third person
Predicted reflexive, given the content of the <i>decoding</i> workspace	First person	First person	First person	First person
Expectancy violation during the decoding process?	No	Yes	No	Yes
<i>Output sentence</i>	(4d) <i>he ... himself</i>	(4d) <i>he ... himself</i>	(5d) <i>I ... myself</i>	(5d) <i>I ... myself</i>
Predicted reflexive, given the content of the <i>encoding</i> workspace	Third person	Third person	First person	First person
Repair needed during the encoding process?	Yes	No	No	Yes
Source(s) of response delay VIOLATION + REPAIR	REPAIR	VIOLATION	—	

TABLE A4
 Shared-workspace hypothesis: derivation of the predicted RT pattern for Experiment 2 (summarised in Table 3)

	Paraphrasing		Proofreading	
	(4a) <i>I ... myself</i>	(4b) <i>I ... *himself</i>	(5a) <i>I ... myself</i>	(5b) <i>I ... *himself</i>
Input sentence	(4a) <i>I ... myself</i>	(4b) <i>I ... *himself</i>	(5a) <i>I ... myself</i>	(5b) <i>I ... *himself</i>
Person feature of input reflexive	First person	Third person	First person	Third person
Predicted reflexive, given the content of the <i>shared</i> workspace	Third person	Third person	First person	First person
Expectancy violation during the decoding process?	Yes	No	No	Yes
Output sentence	(4d) <i>he ... himself</i>	(4d) <i>he ... himself</i>	(5d) <i>I ... myself</i>	(5d) <i>I ... myself</i>
Predicted reflexive, given the content of the <i>shared</i> workspace	Third person	Third person	First person	First person
Repair needed during the encoding process?	Yes	No	No	Yes
Source(s) of response delay	VIOLATION + REPAIR	—	—	VIOLATION + REPAIR