The relationship between monitoring, control, conscious awareness and attention in language production

Nazbanou Nozari[1,2]

[1]Department of Psychological and Brain Sciences, Indiana University

[2]Cognitive Science Program, Indiana University

Corresponding author: Nazbanou Nozari

Address: 1101 E 10th St., Bloomington, IN 47405

Phone number: 812-855-5818

Email: bnozari@iu.edu

Lab website: https://www.nozarilab.com/

Twitter handle: @NozariL

Abstract

This paper discusses the relationship between monitoring, control, conscious awareness, and attention in language production. Instead of focusing on a specific theory, I will examine these relationships within a framework that accommodates multiple (complementary) monitoring views, and discuss key differences between situations where competition is resolved internally vs. those that recruit external control. The takeaway message is that production performance is optimized by self-regulating monitoring-control loops, which operate largely subconsciously, but conscious awareness can be —and often is— triggered by the monitor. When triggered, in conjunction with the control system, such awareness can lead to attentional control of both the primary production process, as well as the monitoring process. I will also touch upon the repair process and its relation to these issues, and end by discussing some of the open questions as possible avenues for future research.

Keywords

Monitoring, control, consciousness, language production, conflict monitoring, forward models, repairs

The relationship between monitoring, control, conscious awareness and attention in language production

For many years, the main focus of language production research has been on understanding the nature of representations and operations that allow speakers to produce speech (Levelt et al., 1999; see Dell et al., 2014, for a review). A positive development in recent years has been the growing interest in understanding how the production process is monitored and regulated (see Hartsuiker, 2014; Nozari, 2018; Nozari & Novick, 2017, for reviews). But this interest has also engendered some confusion regarding the definition and scope of monitoring, control, and their relationship to each other and to conscious awareness and attention. The goal of the current paper is to address these issues in the form of two broad questions: *(1) What is the relationship between monitoring and control?* And *(2) What is the relationship between these two and conscious awareness and attention?*

The paper is intentionally written in a relatively simple language and for a broad audience, ranging from students with little familiarity with theories of language production, monitoring or control, to experts. It does not aim to present a detailed overview of the main monitoring and control theories in language production, which exists elsewhere (Gauvin & Hartsuiker, 2020; Nozari, 2020, 2024). Whenever the two questions posed above can be answered in a model-independent manner, I will do so. However, in some cases, the answer depends critically on the specific theory and its underlying mechanisms. In such cases, I will discuss the aspects of the theory that are relevant to the issues at hand and refer the reader to the original or review papers for more detail. Similarly, the paper is not meant to present a neural model of monitoring and control, as details of neural implementation depend on specific monitoring theories (see Guenther & Vladusich, 2012; Nozari, 2022, for reviews). However, I will occasionally use neural data to highlight key distinctions.

1. What is the relationship between monitoring and control?

In the language production literature, the terms monitoring and control have sometimes been used interchangeably. Partly to blame for this is that the most prominent outcome of the monitoring and control processes is a single behavior, i.e., speech error repairs, which makes it difficult to distinguish between the two processes preceding it. In addition, the term "cognitive control", which refers to an entire field dealing with the regulation of goal-oriented behavior, makes it even more confusing to disentangle the processes underlying monitoring and those involved in the implementation of control. Although this confusion is understandable, I believe that distinguishing monitoring from control implementation and understanding how the two relate to one another is essential for grasping how the language production system is regulated. In this paper, I use the term "control" to refer to "control implementation" rather than its broader sense of the ensemble of mechanisms that collectively regulate behavior. I will unpack this by first working through an example of a simple and familiar monitoring-control system. I will then extend the principles extracted from this example to the more complex language production system. As many of the terms used throughout this paper do not have a universally agreed-upon definition, I have included a glossary in Box 1 to define how I use these terms.

Box 1. Glossary

The key terms, as I use them in this paper, are defined in alphabetical order.

Attention. The ability to consciously and deliberately enhance a mental process.

Conflict-based monitoring. Monitoring the level of conflict, i.e., how close the activation levels of competing representations are, within the language production system to estimate the probability of an error.

Conscious awareness. The ability to report an action or a mental process.

Control. A process that implements a regulatory action, the need for which has been signaled by the monitor.

Criterion. A cutoff point in the information collected by the monitor past which the need for a control is signaled.

External control. Resolving the competition between target and competitors through boosting the goal-relevant response by neural regions outside of those storing task-specific representations. Critical when the stimulus-driven information disfavors the goal-relevant response.

Forward-model monitoring. Monitoring one's speech through comparing the anticipated and actual perceptual outcomes of a produced utterance.

Internal control*. Resolving the competition between target and competitors through lateral inhibition. Sufficient when the stimulus-driven information favors the goal-relevant response.

Listener-driven monitoring. Monitoring processes that are critically dependent on the listener's feedback or the speaker's beliefs about the listener's mental state.

Monitoring. Collecting information about the current (and potentially future) state of a system.

Perceptual loop monitoring. Monitoring one's speech through the comprehension system, i.e., the same system involved in monitoring other people's speech.

Repair. A new production to substitute an already-selected response.

Speaker-internal monitoring. Monitoring processes that are based entirely on the speaker's own action and cognition, without relying on the listener's feedback.

* While referred to as "control" to acknowledge its role in suppressing the non-target representations, internal control is a natural and continuous part of the production process and is not engaged/disengaged based on the signal by the monitor. In contrast, external control fits the canonical definition of control.

1.1 An illustrative example

Generally speaking, monitoring refers to collecting information regarding the current state of the system. Since many current states are closely linked to future states, the information about the current state can also be predictive of future states. An easy-to-understand example of a monitor is a thermostat. A thermostat registers the current temperature. Unless changed by some means, the current temperature is also predictive of the temperature within the next few minutes. This is useful information that can form the basis of a decision to change (or not to change) the current state. If the temperature is just right, nothing needs to be done. If it is too high, the air conditioning (AC) system is triggered. Either way, the thermostat itself does not actively change the temperature. It merely registers and passes the information on to a system that takes action to change the temperature, i.e., the control system. In our example, the control system is the AC. In a nutshell, the monitoring system *collects* the information, whereas the control system *acts* upon that information.

The distinction between monitoring and control is important for three reasons: (1) they have different and non-interchangeable mechanisms. (2) These unique mechanisms are implemented by unique and non-interchangeable hardware. Finally, (3) monitoring and control can be damaged independently. If the thermostat is broken, the AC will not turn on and off automatically but if switched on manually could still cool the room. Similarly, if the AC is broken, the thermostat could still accurately show the current temperature. This is called a "double dissociation" and is strong evidence for the separability (albeit not necessarily total modularity) of components/pathways within a larger system (McCloskey, 2015; Plaut, 1995; Van Orden et al., 2001).

While understanding the distinction between monitoring and control processes is important for these reasons, it is equally important to note that separation does not mean complete independence. The thermostat and the AC jointly regulate the temperature by forming a loop, without the need for interference from an outside agent. If the temperature exceeds a set threshold, a signal is sent to the AC, turns it on and the air is cooled. When the temperature drops below the threshold, another signal sent to the AC turns it off. If either the thermostat or the AC breaks, the loop is disrupted, and the temperature will no longer be regulated in this manner. In short, this simple example describes specialized, non-interchangeable, processes for monitoring and control, which nevertheless form a loop within which the outcome of each process directly affects the outcome of the other process.

1.2 Applying the above framework to language production

The example above helped provide a general definition for monitoring and control processes, and show that the two are distinct, but interact to form a self-regulating loop. All this is directly applicable to language production. However, the complexity of the production system raises new questions. For instance, given its simple goal, it is easy to see that the information to be monitored by the thermostat is temperature. But what information does the monitor collect in the language production system? How does the monitor know if the current state is desirable or not? The answer to these questions depends on the specific monitoring theory. For example, forward models assume that the monitor uses the sensory outcomes of production, i.e., acoustic and proprioceptive information. If these do not match the anticipated sensory information associated with a given production, e.g., if /p/ is produced, but the representation for /b/ is activated in the acoustic space, the current state is marked as a potential error (Bohland et al., 2010; Guenther, 1994, 2016). Conflict-monitoring models assume that the monitor uses information from within the production system, namely, the difference between the levels of coactivation (i.e., conflict) of items competing for selection. Since higher levels of conflict are associated with greater probability of errors compared to lower levels of conflict, the monitor can label high-conflict states as potential errors (M. M. Botvinick et al., 2001; Hanley et al., 2016; Nozari et al., 2011; Pinet & Nozari, 2021). Although using different measures, both kinds of monitoring systems ultimately label states that are closer to an error as problematic states.

If the current state is deemed problematic, control processes are triggered. As described above, unlike the monitor, control processes take action to regulate production. The action taken by the control system depends on the problem at hand. For example, if a word has already been spoken, a monitoring red flag signals a potential need for a repair. If, on the other hand, the red flag is raised before the word has been spoken, it signals the need for halting articulation and reexamining the prepared output (Nozari & Hepner, 2019b). Regardless of the specific action, the relationship between the monitoring and control systems is the same as the thermostat-AC: once an action is taken, the monitoring system continues to survey the state of production to assess whether continued control is or is not necessary. For example, if a repair is successful, the information picked up by the monitor will no longer signal a problem. Therefore, control will no longer be necessary and the production routine will be resumed (Logan, 2018).

The same three arguments for the separation of monitoring and control processes discussed in the context of the thermostat-AC example also hold for language production. In terms of distinct mechanisms, monitoring processes collect information (e.g., sensory output, conflict, etc.) for evaluative purposes, whereas control processes implement actions (e.g., repairs, halts, etc.). Correspondingly, monitoring and control have different neural correlates. For example, monitoring models that use acoustic information, require the involvement of the auditory cortex, i.e., superior temporal gyrus (STG; Bohland et al., 2010; Guenther & Vladusich, 2012), but the implementation of control takes place within the production system and not in the auditory cortex (see Guenther, 2016, for details). Finally, in terms of neuropsychological double dissociations, some individuals with aphasia are incapable of detecting errors in their speech (failure of monitoring) but if prompted to repair a response may be capable of it, whereas others show the ability to detect an error while not being able to fix the problem (failure of control) (Nozari, 2019; Nozari et al., 2011).

At the same time, the idea of a self-regulating monitoring-control loop holds here as well. For example, in conflict monitoring, conflict information for lexical items competing for selection is gathered from the middle and posterior temporal gyrus, where such representations are stored (Krieger-Redwood & Jefferies, 2014; Lewis & Poeppel, 2014; see Nozari, 2022, for a review), and picked up by the anterior cingulate cortex (ACC), which is often considered to have an evaluative role over performance (Akam et al., 2021; M. M. Botvinick et al., 2004), most likely in a hierarchical way, with rostral parts monitoring for higher-level task goals and mid-cingulate parts monitoring subgoals and lower-level actions (Alejandro & Holroyd, 2024). The richly connected ACC then communicates the need for action regulation to regions such as the lateral prefrontal cortex (LPFC), which are involved in resolving conflict in a goal-directed manner (Nozari et al., 2014; Nozari & Thompson-Schill, 2016). During control implementation, conflict is continuously monitored by the ACC, and when the current state is deemed satisfactory, the LPFC involvement is reduced or discontinued. The close relationship between the ACC (monitoring) and the lateral PFC (control) (e.g., De Pisapia & Braver, 2006) is clear in this example, and well-supported by empirical evidence. However the two regions have distinct functions, and create distinct patterns of abnormal behavior when lesioned selectively (Haddon & Killcross, 2006).

We are now prepared to answer the first question, *What is the relationship between monitoring and control?* The two are separate processes with distinct neural correlates, which, nevertheless form a self-regulating loop.

2. The relationship between monitoring/control and consciousness/attention

Does the monitoring-control loop require conscious awareness and attention to operate? This question goes to the heart of the mechanisms underlying monitoring and control. Given the relative separation of the two, discussed in the previous section, I will tackle this question separately for monitoring and for control in the following sections, but let us first define the slippery terms "consciousness" and "attention".

Consciousness is notoriously difficult to define (Chalmers, 1995). Here, I define conscious awareness as the ability to report an action or a mental process. I do not mean the ability to report the mechanisms underlying the mental process; rather, the ability to simply report whether an action or a mental process has taken place or not. For example, a speaker would be considered "aware" of an error if they can correctly report that they produced an error. Conversely, if after committing an error, the speaker genuinely insists that no error has occurred, they would be considered "unaware" of that error. Basing conscious awareness on explicit reports has the caveat of being confounded with truthfulness (perhaps the speaker is aware of the error but is consciously denying it). However, in most cognitive studies, there is little incentive for being dishonest. A more likely confound is memory failure (perhaps the speaker was aware of the error when it occurred but later forgot it). This confound can be minimized by seeking the report immediately after the action or mental process has occurred. I do not claim that this definition of conscious awareness is the best one, but it is an experimentally feasible one and a useful one for the purpose at hand.

Attention also does not have a universal definition. Here, I define it as the ability to consciously and deliberately enhance a mental process. From this definition, it is clear that attention presupposes conscious awareness. A possible exception is attentional capture by salient environmental stimuli (e.g., an unexpectedly loud noise), which simultaneously triggers both attention and conscious awareness of the stimulus, rather than making attention contingent on conscious awareness of the stimulus (Treisman & Gelade, 1980). But when speaking of attentional control in language production, we often mean focusing on a goal, e.g., higher accuracy, and voluntarily directing mental processes to meet that goal (Nozari & Dell, 2012; Nozari & Thompson-Schill, 2013). This is the sense in which I will use the word "attention".

2.1 Monitoring and conscious awareness

2.1.1 Speaker-internal vs. listener-driven monitoring

Does monitoring require conscious awareness? The answer depends on the specific monitoring theory. But which monitoring theory is the right one, so we can focus on that? I have previously argued that there is no one "right" monitoring theory; rather, monitoring in language production must be viewed as a multi-process operation (Nozari, 2020, 2024). While there are several technical arguments to support this position (see Nozari, 2020), I will focus on an intuitive one here: which aspects of language production need to be monitored? In the simple thermostat-AC example, there is only one element that requires monitoring: temperature. In contrast, many aspects of language production need to be monitored. Speakers must ensure that they choose the right words and sentence structures, pronounce the words correctly, and speak with the right volume and rate. When in the presence of an interlocutor, they must also ensure that their speech communicates their message effectively. It is evident from this short description that some of these aspects, e.g., choosing and pronouncing the words correctly can be judged by the speaker even in the absence of an audience. In contrast, other aspects, e.g., whether the speech was clear enough to convey the message, require feedback from an audience.

Corresponding to the distinction above, I will define two types of monitoring mechanisms, *speaker-internal* and *listener-driven* mechanisms (see Lane & Ferreira, 2008, for a parallel categorization of primary linguistic processes). Speaker-internal mechanisms refer to processes that are based entirely on the speaker's own action and cognition, without relying on the interlocutor's feedback. Most formal theories of monitoring such as forward models (e.g., Guenther, 2016) and conflict monitoring (Nozari et al., 2011) are examples of speaker-internal mechanisms. In contrast, listener-driven monitoring refers to processes that critically hinge on listeners' feedback or the speaker's beliefs about the interlocutor's mental state. The latter can use a wide range of cues. In fact, any sign that the listener has failed to grasp the speaker's communicative intentions can act as a critical input to listener-driven monitoring mechanisms. Such failures can arise due to low linguistic proficiency of the interlocutors, noise in the communicative channel, or lack of common ground between the speaker and the listener.

It is worth noting that speakers often set the parameters of conversation in a way to accommodate their interlocutor. For example, when speaking to a child, or a new learner of English, one usually uses shorter sentences and avoids low-frequency words. When speaking in a noisy restaurant, one speaks louder (although this adjustment also has a speaker-internal mechanism, i.e., The Lombard effect; Lombard, 1911). Finally, one usually considers the knowledge and perspective of one's audience during a conversation (Brown-Schmidt et al., 2008; Yoon et al., 2012). However, choosing an appropriate set of parameters does not guarantee that miscommunication will not arise. A given 3-year-old may or may not know a certain word, an increase in the voice volume may not be sufficient to overcome environmental noise, and a speaker may miscalculate what is or is not in the common ground with the listener (Horton & Keysar, 1996; Lane & Ferreira, 2008;

McKinley et al., 2017; Trude & Nozari, 2017). These temporary mishaps can derail comprehension. In such cases, listeners often signal the problem verbally (e.g., "What do you mean?", "Pardon me?") or non-verbally (e.g., looking puzzled). If perceived by the speaker, these signs provide a reliable signal that a change is needed.

2.1.2 Conscious awareness and speaker-internal vs. listener-driven monitoring

The division of monitoring mechanisms into speaker-internal and listener-driven has critical implications for the involvement of conscious awareness in monitoring. Let us start with the more intuitive listener-driven monitoring. Monitoring a listener's reactions to one's speech requires attending to them (see also Cirillo et al., 2024, for the effects of attention on joint production tasks). If you are not looking at your interlocutor or are too absorbed in your own thoughts to process their responses to you, it is easy to miss signs of incomprehension. In terms of formal monitoring theories, the most appropriate monitoring mechanism for listener-based monitoring is the outer loop of Levelt's classic perceptual loop theory (Levelt, 1983). According to this theory, speakers monitor their own speech through the comprehension system, the same system used to monitor other people's speech. The account's emphasis on the similarity between processing self-produced and other-produced signals for monitoring makes perceptual loop a good candidate for monitoring listener-generated feedback. Although the account's original scope is auditory comprehension, it can be easily extended to include other sensory systems such as the visual processing system. Any information picked up by the speaker's sensory systems that signals a miscommunication raises a red flag for the monitor. This could be classic speech errors or ambiguous utterances (e.g., "the glass" instead of "the tall glass", when there is more than one glass; Levelt, 1983, 1989). Since such information comes from the outside, they must first find a meaningful internal representation (i.e., recognition of another person's comprehension difficulties) to be used by the monitor. A rich body of literature on inattentional blindness tells us that mere exposure to sensory information is not sufficient for processing that information to a meaningful level (Mack, 2003; Simons, 2000). Consequently, a monitoring mechanism that relies on processing sensory information from the environment, by definition, requires conscious awareness and attending to the source of external information. This matches Levelt's own definition of the perceptual loop as an attentional monitoring mechanism (Levelt, 1983, p. 50).

Let us now turn to speaker-internal monitoring. Individuals can still detect and repair their errors when there is no audience to provide external feedback. This can happen through several mechanisms. At first glance, perceptual loop seems to be sufficient here as well, as the comprehension system can also operate on one's own speech. However, empirical data show that the ability to detect one's own speech errors is dissociable from that of detecting errors in other people's speech (Marshall et al., 1998). This dissociation points to a speaker-internal component to the monitoring process, which distinguishes it from processing other people's speech. Two monitoring mechanisms are a good fit here, forward models and conflict monitoring. I musr note that perceptual loop also contains an inner loop that is meant to address internal monitoring. The nature of the inner loop has been debated over the years (see Nozari et al., 2011; Vigliocco & Hartsuiker, 2002, for summaries of arguments) but the most recent definition (Roelofs, 2020) discusses inner loop as an interaction between production representations and their corresponding perceptual representations. Since this is essentially the same idea as forward models, I will not discuss it separately.

As alluded to briefly in the earlier sections, in forward models, an error is detected when the perceptual consequences of production, i.e., acoustic and proprioceptive feedback, do not match the anticipated perceptual consequences (Guenther, 1994, 2016). Forward models are supported by much data, and there is little doubt that they are critical to the detection of sublexical speech errors (Civier et al., 2010; Max et al., 2004; Miller & Guenther, 2021; Runnqvist et al., 2016; Tourville & Guenther, 2011; see Kearney & Guenther, 2019, for a review). However, it is unclear whether they are plausible for monitoring parts of the language system where no clear separation between sensory and motor representations exists (see Nozari, 2020 and Nozari, 2024, for a full discussion of this issue). Moreover, reducing sensory feedback (e.g., by preventing typists from seeing what they are typing) changes the electrophysiological profile of monitoring (Pinet & Nozari, 2020). Specifically, such reduction leads to a more pronounced appearance of an electrophysiological signature of errors, called the error related negativity (ERN; Falkenstein et al., 2000; Riès et al., 2011). This finding is compatible with a mechanism that weights production-internal representations more heavily than distal perceptual representations. An example of such a mechanism is conflict monitoring. In conflict monitoring, when the difference in the activation of two representations that compete for selection is smaller than a criterion, a signal is generated marking the high probability of an error (M. M. Botvinick et al., 2001; Nozari et al., 2011).

It is worth noting that forward models and conflict monitoring are not conflicting theories (Nozari, forthcoming). In fact, conflict monitoring must have its origin in a type of forward model. To understand this, think about how the conflict monitoring system is trained. For conflict monitoring to be useful, conflict must be reliably linked to performance outcome. This "linking" initially relies strongly on evaluating the production outcome, i.e., the spoken utterance, carried out by the perceptual systems. As speakers gain more experience, they learn to link earlier states of the system to the final performance outcome. Once these links are trained, the system can rely on earlier states even in the absence of the final outcome. Different kinds of "early" information can potentially be used, including perceptual and production-internal (e.g., conflict) representations. This process works well even in complex hierarchical systems like language production and is generally described under the framework of model-based hierarchical reinforcement learning (Botvinick & Weinstein, 2014), which is itself an example of the more general reinforcement learning theory (Sutton & Barto, 2018). Reinforcement learning taps into a network of reward processing that includes regions like the medial prefrontal cortex (Zarr & Brown, 2016) and cerebellum (Kostadinov & Häusser, 2022). Naturally, as instances of reinforcement learning models, both forward models of motor speech control and conflict monitoring are also expected to involve these neural regions.

Both forward models and conflict monitoring operate largely subconsciously, which allows them to continuously monitor production without exhausting the speaker or disrupting the primary production process. Moreover, the automaticity that underlies these two mechanisms further allows for extremely fast detection of many errors (Blackmer & Mitton, 1991; R. J. Hartsuiker & Kolk, 2001; Nooteboom & Quené, 2017) as well subtle and subconscious adjustments to speech based on auditory feedback (e.g., Niziolek & Parrell, 2021). Finally, their being subconscious is compatible with the ERN not being critically dependent on conscious awareness (Endrass et al., 2007; Pinet & Nozari, 2020).

2.1.3 Supplemental subconscious monitoring with attention: criterion setting

So far, I have argued that of the three monitoring mechanisms, the outer loop of perceptual loop requires conscious awareness and attentional processing of the incoming information, whereas the basic processes underlying forward models and conflict monitoring do not depend critically on conscious awareness. However, a mechanism that is not critically dependent on attention could, nevertheless, be modulated by attention. This is true of both forward models and conflict monitoring. How? The answer lies in the criterion that speakers set for detecting an undesirable state. I will unpack this with an example below.

As described earlier, forward models detect the discrepancy between the anticipated and actual perceptual consequences of a produced utterance. But how much discrepancy is worth raising a red flag? Articulation is quite variable. Figure 1 shows the distributions of initial fundamental frequency (F0) in repeated productions of words “beer” and “pier” by a single speaker (62 “beer” and 58 “pier” productions; data from Murphy et al., 2024). The figure shows that while the average values of F0 (dashed lines) discriminate between “beer” and “pier”, there is considerable variability in individual productions, creating overlapping F0 distributions for the two words. Since F0 corresponds closely to pitch in perception, we can assume similar overlapping distributions in perception. Now imagine that, intending to produce “pier”, the speaker produces token x (see Figure 1), which falls in the zone of overlap between “beer” and “pier” distributions. Correspondingly, the acoustic feedback would be somewhere in the gray zone.

How does the monitoring system determine whether this is a good enough “pier” or whether production needs to be amended? The decision depends on the monitoring criterion. On figure 2, c1 (solid black line) and c2 (solid red line) show two placements of the criterion. Everything above the criterion is labeled as “pier” and everything below it, as “beer”. With c1 as the criterion, production x is a good enough “pier” and no red flag is raised. But imagine that the speaker is an opera singer or an actor who wants to perfect the diction of their /p/. In other words, they do not want any of their “pier” productions to possibly be heard as “beer”. This can be achieved by shifting the criterion from c1 to c2. Now, the same production x falls below the criterion and is detected as an error that needs an adjustment to production. In practice, such a shift represents focusing attention on producing a highly accurate “pier”. Over time, this can reduce the variability in the two distributions, further shrinking the overlap and making production cleaner.

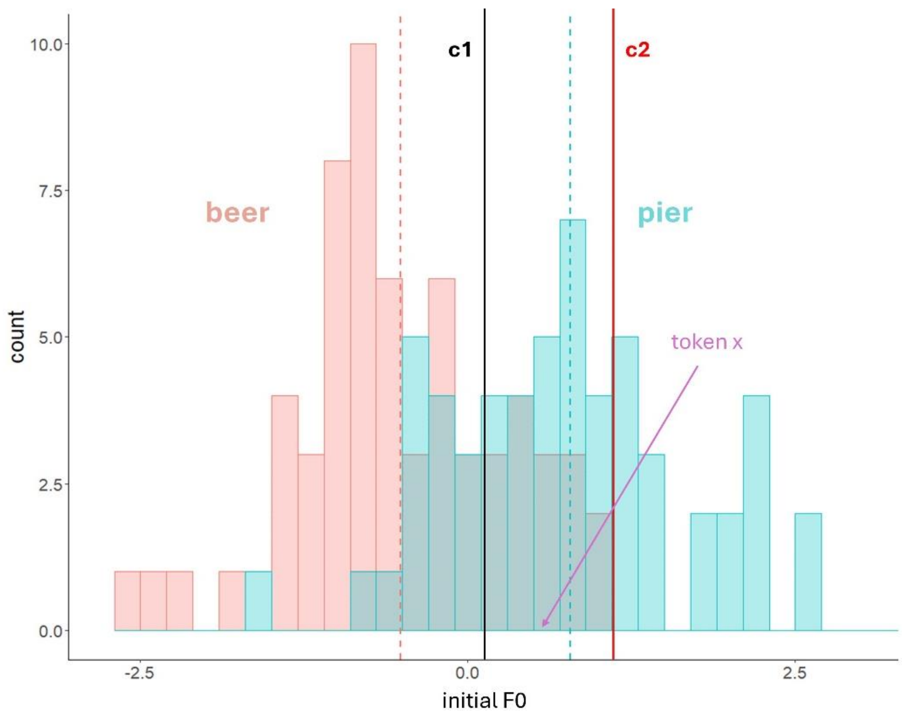

Figure 1. Initial (normalized) F0 distributions of "beer" and "pier" productions by a single speaker. Two placements of the criterion (c1 and c2) are shown.

Criterion shifts are not unique to forward models. The same idea applies to other monitoring theories as well. For example, if we swap the F0 axis in Figure 1 for "conflict" (i.e., the difference in the activation of competing representations), the distributions on the left and right could be thought of as distribution of correct vs. error responses at any stage of production, and the placement of the criterion determines how stringent the monitor is (see Nozari & Hepner, 2019b, 2019a, for an extended discussion on criterion setting in language production, and Pinet & Nozari, 2021, for an application). To summarize, regardless of whether the basic monitoring mechanism does or does not critically depend on conscious awareness, it can be modulated by explicit attentional mechanisms, which tune it to the speaker's goals. In line with the claim that subconscious monitoring mechanisms could be modulated by conscious and deliberate processes, ERN shows sensitivity to manipulations of awareness and reporting accuracy (Steinhauser & Yeung, 2010; see Wessel, 2012, for a review).

2.1.4 Monitoring vs. search: a critical distinction

Before I end this section, I must draw the reader's attention to a critical distinction between "monitoring" and "search". A search task is a task that requires participants to look for a specific target among distractors. Unless the target possesses certain salient features among largely uniform distractors to create a pop-out effect (Treisman & Gelade, 1980), such a search requires top-down attention (Wolfe & Horowitz, 2017). An example of such a search task in language production is the "phoneme monitoring" task, in which speakers must look for a specific phoneme within a word. While originally recognized as a task that tapped into the nature of representation and attentional mechanisms (e.g., Titone, 1996), it later became a tool for studying monitoring processes that underlie speech error detection (e.g., Özdemir et al., 2007).

However, there is little reason to believe that attentional search is the underlying mechanism for speech monitoring. As reviewed above, attention is not a core feature of language monitoring. Moreover, even the perceptual loop, which depends more on attention than the other monitoring mechanisms, does not set an explicit a priori search target (e.g., /b/) to detect as an error. Some have argued for higher-level categories, e.g., lexical vs. non-lexical, as search targets (Baars et al., 1975; Hartsuiker et al., 2005; Nooteboom, 2005). However, conscious categorization of self-produced utterances based on lexical status shows a timing profile that is incompatible with the timeline of error detection in language production (Nozari & Dell, 2009). Therefore, even if such categories matter for monitoring, the underlying mechanism for their detection is not a search. In short, tasks that require a deliberate search for certain elements or categories in linguistic stimuli have their own utility, but they are not well-suited for indexing monitoring processes in language production.

2.2 Control and conscious awareness

The independence of some of the monitoring mechanisms from conscious awareness allows speakers to monitor language production continuously and effortlessly. As long as all is well, production can carry on smoothly. But what if a red flag is raised and something needs to change? Can control be implemented without conscious awareness? As in monitoring, answering this question requires a discussion of the mechanisms that underlie control in language production. Control is difficult to measure directly. Therefore, researchers often measure indirect proxies of control, such as behavioral interference, which often indexes conflict. In a task like Stroop, when the word RED is printed in blue (incongruent), participants take longer and make more errors when naming the print color compared to when it is printed in red (congruent). These indices of behavioral interference are taken as indirect evidence that control was required (and presumably recruited) in the incongruent condition. Similarly, neuroimaging studies of control often take the greater activation of certain regions, such as the dorsolateral and ventrolateral prefrontal cortex in the incongruent conditions as an index of control recruitment (e.g., Thompson-Schill et al., 1997; Wagner et al., 2001). However, not all tasks that induce behavioral interference show a similar neural profile (e.g., Pinet & Nozari, 2023), casting doubt on the proposition that control is always implemented in the same way in language production. I will unpack this in the next section.

2.2.1 Internal vs. external control

Let us start the discussion of the control mechanisms in language production with an example of two conflict-inducing tasks with the same stimuli but different task instructions. In both tasks, participants complete multiple blocks, each with only two pictures, presented randomly multiple times, and must name the pictures as quickly and accurately as possible. One task manipulates the similarity between the two pictures, e.g., they could be taxonomically related (cat/dog) or unrelated (cat/pen). The other task asks participants to switch the name of the two pictures, i.e., name the picture of a dog as a "cat" and vice versa, thus creating a Stroop-like effect. Both contextual similarity and Stroop-like manipulation create behavioral interference (e.g., Breining et al., 2016; Harrison et al., 2020; Nozari et al., 2016; Oppenheim & Nozari, 2024; Schnur et al., 2006; see Nozari & Pinet, 2020, for a review). But do these two tasks recruit the same control mechanism?

Figure 2 shows a schematic of the production process for the first task, i.e., the contextual similarity manipulation. The picture of a dog has been named on the previous trial and now the speaker views the picture of a cat. The picture activates its corresponding semantic features, some of which are unique to cat (green circles), some are shared between cat and dog (purple circles), and some may even be unique features of dog lingering from the last presentation (blue circle). Activation then spreads from semantic features to the lexical items. Since it is compatible with both “cat” and “dog”, both lexical items initially gain some activation (timepoint 1). But selection is not yet possible, because the activation levels are close and conflict is high. However, as time passes, “cat” receives greater activation, because more semantic features support it compared to “dog”, until its activation far surpasses that of “dog” (timepoint 2). At this stage, “cat” can be selected for production.

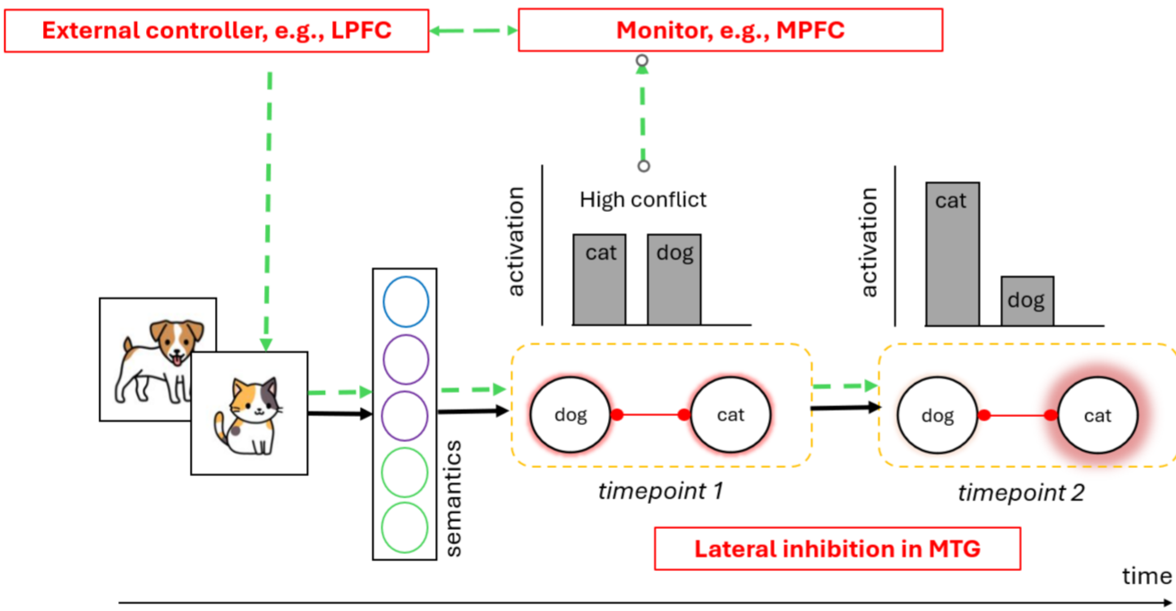

Figure 2. Schematic of the contextual similarity task. Black arrows show the basic production process. Lateral inhibition (red connections in yellow boxes) is sufficient to resolve competition, but external control can be optionally recruited to facilitate the process (dashed green arrows). LPFC = lateral prefrontal cortex; MPFC = medial prefrontal cortex; MTG = middle temporal gyrus.

As the reader might have noticed, the process was completed simply by following the black arrows on the picture and without any mention of a “control” operation. Strictly speaking, control is not necessary when the stimulus-driven information, in this case, the cat picture, provides stronger cues for the target than the competitor. However, some consider the mutual inhibition between "cat" and "dog" (i.e., lateral inhibition[1]) to be a form of *internal control*. This is shown in the dashed yellow boxes in Figure 2. Because "cat" and "dog" have mutually inhibitory connections, as "cat" gains more activation, it naturally enforces stronger inhibition over "dog" and ultimately suppresses it sufficiently to be selected (see Oppenheim & Nozari, 2024 and Nozari, forthcoming, for a more extensive discussion of selection mechanisms in language production). When competition is stronger (e.g., related vs. unrelated competitors), the target must gain more activation to sufficiently suppress the competitor, which manifests as behavioral interference (e.g., Roelofs, 2018). As noted earlier, since interference is often taken as an index of control implementation, it has been hypothesized that contextual similarity might recruit control (e.g., Crowther & Martin, 2014). However, note that this is different from the kind of control we defined within the monitoring-control loop, i.e., "A process which implements a regulatory action, the need for which has been signaled by the monitor." (Box 1). Lateral inhibition is simply a natural consequence of spreading activation, and not a process that is uniquely recruited upon the detection of an undesirable state.

Now let us turn to the second task, i.e., the Stroop-like task. The speaker is looking at the picture of a cat. Therefore, the process of spreading activation and lateral inhibition shown in Figure 2 will naturally lead production towards the word "cat". But this is the incorrect response according to the task goal. The discrepancy between the stimulus-driven and goal-driven responses creates a high-conflict state that cannot be satisfactorily resolved without a shift of processing away from the prepotent response. This situation calls for *external control*, i.e., a mechanism that can resolve the competition between target and competitors by boosting the goal-relevant response. While internal control is implemented where task representations are, e.g., middle and posterior temporal gyrus for lexical representations (Krieger-Redwood & Jefferies, 2014; Lewis & Poeppel, 2014;

---

[1] It is worth noting that the presence of lateral inhibition among lexical items in the production system is not undisputed (e.g., Mahon et al., 2007), and even when not theoretically rejected, such mutual inhibitory connections are not always part of the computational models of language production (e.g., Dell, 1986; Oppenheim et al., 2010). However, given the presence of lateral inhibition in almost all other cognitive systems, it is reasonable to assume that they also exist in the language production system (Nozari, forthcoming).

see Nozari, 2022, for a review), external control involves areas that do not store domain-specific representations. These include frontoparietal networks, as well as domain-general subcortical networks (Wessel & Anderson, 2024).

The distinction between the internal and external control mechanisms is supported by the empirical data. Using the two tasks described above, we found that the magnitude of interference in the Stroop-like task was highly correlated across pairs, regardless of their contextual similarity, but there was no correlation between the magnitude of this interference and the interference induced by contextual similarity (Nozari, Freund, et al., 2016). Neurally, we found that electrophysiological signatures of interference were also distinct in these two tasks (Pinet & Nozari, 2023). Neuroimaging evidence from other labs is consistent with this picture: Stroop-like tasks engage prefrontal regions more consistently than contextual similarity tasks, both for monitoring (medial regions) and control (lateral regions) (de Zubicaray et al., 2014; de Zubicaray et al., 2013; Peterson et al., 1999; Vanderhasselt et al., 2009). Finally, the two types of control are distinguishable by congruency sequence effects (CSE; Gratton et al., 1992). CSE refers to the modulation of interference on a current trial as a function of the level of interference on a previous trial. Specifically, performance shows less interference on a current incongruent trial if it is preceded by another incongruent, compared to a congruent, trial. The mechanism underlying CSE is thought to be as follows: When the system encounters an incongruent or high-conflict trial, the difficulty associated with resolving conflict between the prepotent and the goal-appropriate response causes the monitoring system to signal the need for control. Once recruited, control mechanisms resolve the competition on the current trial, but are also on hand for the next incongruent trial, reducing the behavioral interference on those high-conflict trials compared to the ones that were not receded by another high-conflict trial.

CSE is a robust finding in both linguistic and non-linguistic domains, and showcases the self-regulating monitoring-control loop that was introduced at the beginning of this paper (Duthoo et al., 2014; Freund & Nozari, 2018; Zhu et al., 2024; see Braem et al., 2019; Egner, 2007, for reviews). In contrast to Stroop-like tasks, contextual similarity effects do not trigger CSE. In fact, repeated encounters with similar items cumulatively *increases*, rather than decreases, interference (Costa et al., 2009; Schnur, 2014), showing that internal control does not form the same kind of self-regulating monitoring-control loop as the external control.

2.2.2 Conscious awareness, attention and internal and external control

In the last section, I argued that although lateral inhibition can be thought of as a kind of internal control, it is part of the basic and continuous production process. Consequently, in so far as those basic processes are subconscious, so is internal control. While speakers are aware of the general message and certain linguistic choices, they are largely unaware of the dynamics underlying production (Nozari, 2018). This extends to internal control. However, as with monitoring, a control process that does not critically depend on attention can nevertheless be enhanced by conscious attention. For example, while lateral inhibition naturally resolves competition in favor of the target response, interference from the competitor can slow down this process. Deliberately deciding to pay more attention to be faster without losing accuracy, can indeed engage external control, although this engagement is optional. This optionality explains the contradicting reports on the necessity and engagement of LPFC in production tasks that tap contextual similarity (de Zubicaray et al., 2013; Piai et al., 2013, 2016). When engaged, LPFC can help processing by focusing attention on the relevant stimulus and thus facilitate stimulus-response mapping (e.g., Nozari, Mirman, et al., 2016). This is shown in the green connections in Figure 2.

In contrast to the largely subconscious internal control mechanisms that only electively tap attentional control, actively focusing on the task goal is necessary in tasks where stimulus-driven information biases the speaker away from the correct response. This is usually referred to as "attention" to the goal-relevant information. To do well in the Stoop task, one must "attend" to the ink color and not the written word. Attention implies a certain degree of intentionality and conscious deliberation. Although maintaining task goals, which activates the monitoring-control loop requires conscious awareness and attention, it does not imply that speakers are consciously aware of the inner workings of the loop, such as moment-by-moment regulations observed in CSEs in language production tasks.

2.2.3 Consciousness and repairs

One of the most basic, and yet most important, functions of a monitor is to flag errors. Once flagged, such errors should ideally be repaired. Since repairs are actions taken following a monitoring red flag, they fit under the definition of control processes. However, they are special in the sense that they must dial back processing that has gone far enough in the wrong direction to have resulted in a mis-selection. This property brings up interesting questions. For example, is the repair process a continuation of the process that led to the incorrect selection, or does the system have to be reset and start processing from scratch (Hartsuiker et al., 2008; Nooteboom & Quené, 2020; Nozari et al., 2019; Seyfeddinipur et al., 2008)? For this reason, I will discuss repairs separately.

Compared to other processes, the study of repairs has received little attention (cf., Gauvin & Hartsuiker, 2020; Nooteboom & Quené, 2013, 2017, 2020; Nozari et al., 2019). We have recently laid out the structure of a computational model that explains the basic properties of repairs (Burgess & Nozari, 2022; Nozari, forthcoming). In a nutshell, the model implements a respond-&-check mechanism, meaning that selection at the time of responding is compared to the selection at a timepoint shortly after response selection. If there is a discrepancy, the second selection will be used as a repair. This basic mechanism, which combines primary selection processes in language production with the idea of post-selection processing from the decision-making literature (Pleskac & Busemeyer, 2010), can generate repairs that are fast and largely automatic, in the sense of using the output of the same primary production processes. But what is the role of conscious awareness in this process?

Empirically, this question is quite difficult to answer. Ideally, speakers would be asked to report repairs right after they have been committed, to avoid memory errors. However, forcing speakers to continually interrupt running speech to report repairs would make the primary task difficult and unnatural. A potential solution is to turn to single-word production, but neurotypical native adult speakers rarely make errors on spoken single-word production tasks. To avoid these challenges, we conducted several single-word typing-to-dictation experiments (Pinet & Nozari, 2020, 2022). Participants heard a word, typed it under a temporal deadline, and were allowed to use the backspace and repair any errors as in normal typing. After each trial, they were asked whether they had committed an error, and if yes, whether they had repaired it. Across multiple experiments, we replicated the finding that between 10-20% of repairs were not consciously reported, even immediately after the repair had been applied. We ruled out factors such as misunderstanding the question, pressing the wrong button to report an error/repair, and memory failure, as alternative explanations (Pinet & Nozari, 2022). Therefore, we can say with certainty that, at least in some modalities of language production, repairs do not critically depend on conscious awareness. However, given the differences between typing and speaking, an immediate extension to spoken production is premature. Moreover, the majority of repairs were associated with conscious awareness, showing that even if not critical to the repair process, conscious awareness is often triggered when a repair is committed. However, rather than being a precursor of repair, it is likely to be triggered by the monitoring process, in parallel with the repair.

Interestingly, repairs show an adaptive property: when the proportion of errors increases, so does the proportion of repairs (Levelt, 1989; Nozari et al., 2019). Although with a different time scale than CSE, the spirit of the effect is similar: detecting greater difficulty in the production system upregulates the control processes that mediate repairs. Therefore, repairs fit well within the closed monitoring-control loops discussed earlier in this paper. We have proposed that this loop operates largely subconsciously (Burgess & Nozari, 2022). However, as in monitoring and control processes other than repairs, this loop can be upregulated by attention, i.e., by focusing specifically on the goal of repairing as many errors as possible.

We can now answer the second question, *What is the relationship between monitoring and control and conscious awareness and attention?* Although the answer depends, to some extent, on the specific mechanism, monitoring and control processes are largely subconscious. However, they often trigger conscious awareness, which can, in turn, trigger explicit attentional processes that affect both the primary production process, as well as monitoring and control processes.

3. Conclusions

The conclusions of this paper are summarized in the five points below.

(a) Monitoring and control are separate but interacting processes. In most cases, they jointly form a self-regulating loop to optimize performance.

(b) Monitoring is not a singular process. Multiple monitoring mechanisms contribute to language production. Broadly speaking, these can be divided into speaker-internal and listener-driven processes.

(c) While listener-driven monitoring mechanisms require conscious awareness of, and attending to, the listener, speaker-driven monitoring mechanisms are largely subconscious. However, these subconscious processes can also be subject to attentional regulation based on task goals.

(d) Behavioral interference in language production does not always index the recruitment of the same control mechanism. When stimulus-driven information naturally drives processing toward the correct response, lateral inhibition is often sufficient. While lateral inhibition can be thought of as an "internal" control mechanism, it is part of the basic continuous production process and is not recruited uniquely upon detecting an undesirable state. In contrast, situations in which the stimulus-driven information is insufficient to arrive at the correct response recruit external control. External control forms a self-regulating loop with monitoring processes, which mediates the moment-by-moment adjustments to the language production system in the form of CSEs.

(e) Repairs are a special case of control processes. Although understudied, current data suggest that the primary repair process can be carried out subconsciously, but the same process that triggers repairs often triggers conscious awareness in parallel. Moreover, repairs show the same kind of adaptation observed in CSE, albeit on a longer time scale, suggesting that they are part of a monitoring-control loop.

4. A visual summary

Figure 3 shows a schematic of the relationship between language production, monitoring, control, conscious awareness, and attention. The black arrows show the basic monitoring-control loop. There is a close interaction between language production and perception systems; we constantly hear ourselves speak and what we hear affects what we say and how we say it. The information generated within these two systems is continuously picked up by the monitor. The monitor determines, based on the criterion, whether the received information signals an optimal or a suboptimal state of production. If the state is deemed suboptimal, a signal is sent to the (external) control system, which, in turn, implements control over the language production system, until the information conveyed to the monitor signals that the desired state has been reached. At least two monitoring mechanisms, conflict-based detection and forward modeling, can be carried out within these black loops without depending critically on conscious awareness.

The blue arrows show the involvement of consciousness. In most cases, speaking requires conscious planning. Similarly, extracting meaning from the speech input requires some degree of conscious awareness. These are shown in the top-down connections between conscious awareness and the language production and perception systems. However, the monitor can also make speakers aware of a problem they may not have been aware of before. Finally, conscious awareness can trigger control processes in a voluntary manner (red arrows). We often refer to this as "paying attention". The downstream consequences of attention affect the connections between the control system and the language production system (and if need be, the perception system; not shown to avoid cluttering the figure) but can also directly affect monitoring by allowing the speaker to shift the criterion in a goal-directed manner to focus on perfecting a certain aspect of production.

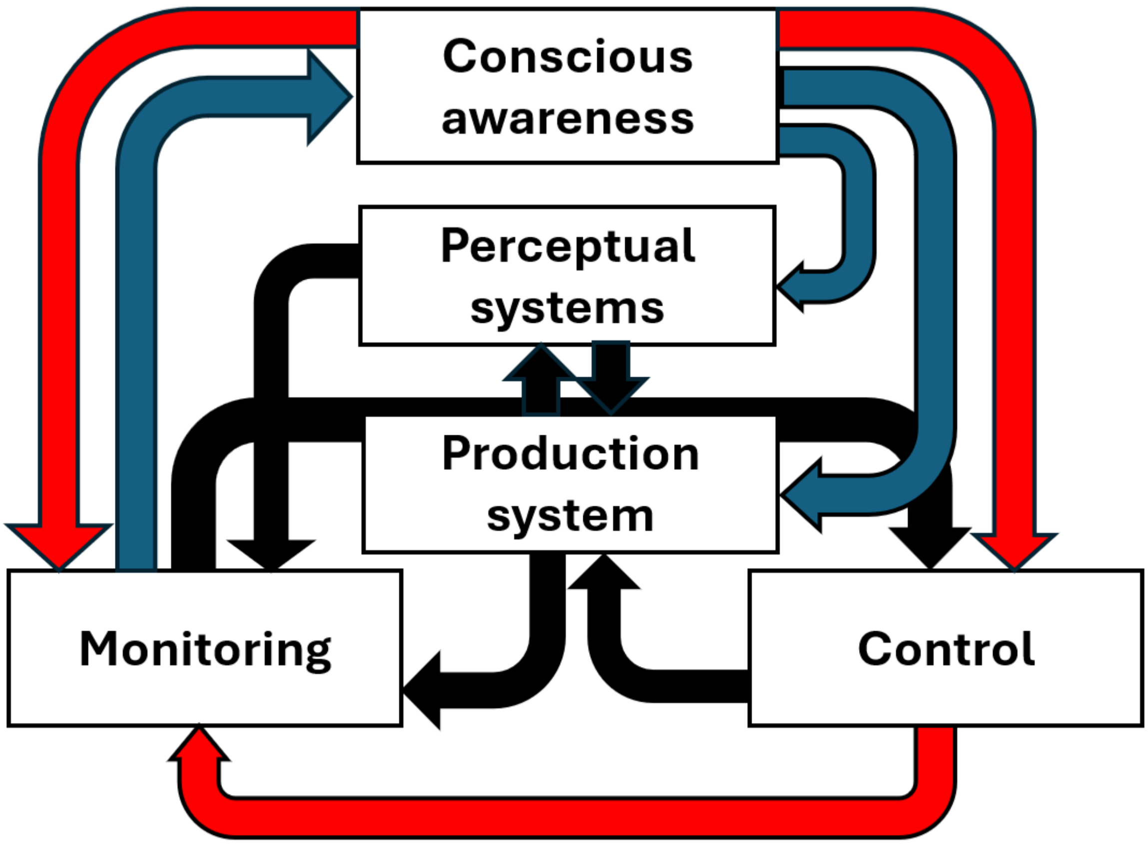

Figure 3. A schematic of the relationship between language production, perception, monitoring, control, conscious awareness, and attention. Black arrows show the basic monitoring-control loops that do not critically depend on conscious awareness or attention. Blue arrows show the involvement of conscious awareness. Red arrows show how explicit attention is recruited and applied. Note that once attention is recruited, it can affect the implementation of control in the production (and perception; not shown) system.

Box 2. Open questions and future directions

This box lists some open questions related to the issues discussed in this paper.

1. There is some evidence that repairs can be carried out without consciousness in typing. Is the same true for spoken production?

2. Evidence for subconscious error corrections described in the text concerns segmental errors (e.g., "cap" for "cat"). Can lexical repairs be carried out without conscious awareness?

3. Repairs show general adaptation effects. Do they also show trial-by-trial adaptation as observed in CSEs?

4. Is the monitoring-control loop regulating repairs the same as the monitoring loop regulating production to avoid repairs?

5. Acknowledgement

This work was supported by the NSF grant BCS-2317121 to N.N.

6. References

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R. F., Dayan, P., & Costa, R. M. (2021). The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron*, *109*(1), 149-163. https://doi.org/10.1016/j.neuron.2020.10.013

Alejandro, R. J., & Holroyd, C. B. (2024). Hierarchical control over foraging behavior by anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, *160*, 105623. https://doi.org/10.1016/j.neubiorev.2024.105623

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*(4), 382–391. https://doi.org/10.1016/S0022-5371(75)80017-X

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, *39*(3), 173–194. https://doi.org/10.1016/0010-0277(91)90052-6

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural Representations and Mechanisms for the Performance of Simple Speech Sequences. *Journal of Cognitive Neuroscience*, *22*(7), 1504–1529. https://doi.org/10.1162/jocn.2009.21306

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M., & Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130480. https://doi.org/10.1098/rstb.2013.0480

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*.

Breining, B., Nozari, N., & Rapp, B. (2016). Does segmental overlap help or hurt? Evidence from blocked cyclic naming in spoken and written production. *Psychonomic Bulletin & Review*, *23*(2), 500–506. https://doi.org/10.3758/s13423-015-0900-x

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, *107*(3), 1122–1134. https://doi.org/10.1016/j.cognition.2007.11.005

Burgess, J. L., & Nozari, N. (2022). A conflict-based model of speech error repairs in humans. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44). https://escholarship.org/uc/item/2qg342vp

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.

Cirillo, G., Strijkers, K., Runnqvist, E., & Baus, C. (2024). Effects of Shared Attention on joint language production across processing stages. *Language, Cognition and Neuroscience*, *39*(2), 203–214. https://doi.org/10.1080/23273798.2023.2260021

Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders*, *35*(3), 246–279. https://doi.org/10.1016/j.jfludis.2010.05.002

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences*, *106*(50), 21442–21446. https://doi.org/10.1073/pnas.0908921106

Crowther, J. E., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: The role of cognitive control and learning. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00009

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*(10), 1322–1326. https://doi.org/10.1016/j.neucom.2005.12.100

de Zubicaray, G. I., Hansen, S., & McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *Journal of Experimental Psychology: General*, *142*(1), 131–142. https://doi.org/10.1037/a0028717

de Zubicaray, G., Johnson, K., Howard, D., & McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex*, *54*, 135–149. https://doi.org/10.1016/j.cortex.2014.01.018

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Nozari, N., & Oppenheim, G. M. (2014). Lexical access: Behavioral and computational considerations. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production* (pp. 88-104). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199735471.013.014

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., & Notebaert, W. (2014). The congruency sequence effect 3.0: A critical test of conflict adaptation. *PloS One*, *9*(10), e110462.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 380–390. https://doi.org/10.3758/CABN.7.4.380

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*(6), 1714–1720. https://doi.org/10.1111/j.1460-9568.2007.05785.x

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, *51*(2), 87–107. https://doi.org/10.1016/S0301-0511(99)00031-9

Freund, M., & Nozari, N. (2018). Is adaptive control in language production mediated by learning? *Cognition*, *176*, 107–130. https://doi.org/10.1016/j.cognition.2018.03.009

Gauvin, H. S., & Hartsuiker, R. J. (n.d.). Towards a New Model of Verbal Monitoring. *Journal of Cognition*, *3*(1), 17. https://doi.org/10.5334/joc.81

Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards a New Model of Verbal Monitoring. *Journal of Cognition*, *3*(1), 17. https://doi.org/10.5334/joc.81

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*, *72*(1), 43–53. https://doi.org/10.1007/BF00206237

Guenther, F. H. (2016). *Neural Control of Speech*. MIT Press.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, *25*(5), 408–422. https://doi.org/10.1016/j.jneuroling.2009.08.006

Haddon, J. E., & Killcross, S. (2006). Prefrontal Cortex Lesions Disrupt the Contextual Control of Response Conflict. *Journal of Neuroscience*, *26*(11), 2933–2940. https://doi.org/10.1523/JNEUROSCI.3243-05.2006

Hanley, J. R., Cortis, C., Budd, M.-J., & Nozari, N. (2016). Did I say dog or cat? A study of semantic error detection and correction in children. *Journal of Experimental Child Psychology*, *142*, 36–47. https://doi.org/10.1016/j.jecp.2015.09.008

Harrison, W., Hepner, C. R., & Nozari, N. (2020). *Is Segmental Interference Position-dependent?*

Hartsuiker, R. (2014). Monitoring and control of the production system. *The Oxford Handbook of Language Production*, 417–436.

Hartsuiker, R. J., Catchpole, C. M., de Jong, N. H., & Pickering, M. J. (2008). Concurrent processing of words and their replacements during speech. *Cognition*, *108*(3), 601–607. https://doi.org/10.1016/j.cognition.2008.04.005

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related beply to Baars et al. (1975). *Journal of Memory and Language*, *52*(1), 58–70. https://doi.org/10.1016/j.jml.2004.07.006

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error Monitoring in Speech Production: A Computational Test of the Perceptual Loop Theory. *Cognitive Psychology*, *42*(2), 113–157. https://doi.org/10.1006/cogp.2000.0744

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

Kearney, E., & Guenther, F. H. (2019). Articulating: The neural mechanisms of speech production. *Language, Cognition and Neuroscience*, *34*(9), 1214–1229. https://doi.org/10.1080/23273798.2019.1589541

Kostadinov, D., & Häusser, M. (2022). Reward signals in the cerebellum: Origins, targets, and functional implications. *Neuron*, *110*(8), 1290–1303. https://doi.org/10.1016/j.neuron.2022.02.015

Krieger-Redwood, K., & Jefferies, E. (2014). TMS interferes with lexical-semantic retrieval in left inferior frontal gyrus and posterior middle temporal gyrus: Evidence from cyclical picture naming. *Neuropsychologia*, *64*, 24–32. https://doi.org/10.1016/j.neuropsychologia.2014.09.014

Lane, L. W., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1466–1481. https://doi.org/10.1037/a0013353

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Lewis, G., & Poeppel, D. (2014). The role of visual representations during the lexical access of spoken words. *Brain and Language*, *134*, 1–10. https://doi.org/10.1016/j.bandl.2014.03.008

Logan, G. D. (2018). Automatic control: How experts act without thinking. *Psychological Review*, *125*(4), 453.

Mack, A. (2003). Inattentional Blindness: Looking Without Seeing. *Current Directions in Psychological Science*, *12*(5), 180–184. https://doi.org/10.1111/1467-8721.01256

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Marshall, J., Robson, J., Pring, T., & Chiat, S. (1998). Why Does Monitoring Fail in Jargon Aphasia? Comprehension, Judgment, and Therapy Evidence. *Brain and Language*, *63*(1), 79–107. https://doi.org/10.1006/brln.1997.1936

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or Insufficiently Activated Internal Models and Feedback-Biased Motor Control as Sources of Dysfluency: A Theoretical Model of Stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*(Spring), 105–122. https://doi.org/10.1044/cicsd_31_S_105

McCloskey, M. (2015). The future of cognitive neuropsychology. In *Handbook of cognitive neuropsychology* (pp. 593–608). Psychology Press. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315783017-33/future-cognitive-neuropsychology-michael-mccloskey

McKinley, G. L., Brown-Schmidt, S., & Benjamin, A. S. (2017). Memory for conversation and the development of common ground. *Memory & Cognition*, *45*(8), 1281–1294. https://doi.org/10.3758/s13421-017-0730-3

Miller, H. E., & Guenther, F. H. (2021). Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology*, *35*(4), 424–441. https://doi.org/10.1080/02687038.2020.1765307

Murphy, T. K., Nozari, N., & Holt, L. L. (2024). Transfer of statistical learning from passive speech perception to speech production. *Psychonomic Bulletin & Review*, *31*(3), 1193–1205. https://doi.org/10.3758/s13423-023-02399-8

Niziolek, C. A., & Parrell, B. (2021). Responses to Auditory Feedback Manipulations in Speech May Be Affected by Previous Exposure to Auditory Errors. *Journal of Speech, Language, and Hearing Research*, 1–13.

Nooteboom, S. G. (2005). Listening to oneself: Monitoring speech production. In *Phonological encoding and monitoring in normal and pathological speech* (pp. 179–198). Psychology Press. https://www.taylorfrancis.com/chapters/edit/10.4324/9780203506196-16/listening-oneself-monitoring-speech-production-sieb-nooteboom

Nooteboom, S. G., & Quené, H. (2013). Parallels between self-monitoring for speech errors and identification of the misspoken segments. *Journal of Memory and Language*, *69*(3), 417–428.

Nooteboom, S. G., & Quené, H. (2017). Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback. *Journal of Memory and Language*, *95*, 19–35. https://doi.org/10.1016/j.jml.2017.01.007

Nooteboom, S. G., & Quené, H. (2020). Repairing speech errors: Competition as a source of repairs. *Journal of Memory and Language*, *111*, 104069. https://doi.org/10.1016/j.jml.2019.104069

Nozari, N. (2018). How special is language production? Perspectives from monitoring and control. In K. Federmeier and D. Watson (Eds.), *Psychology of Learning and Motivation: Current Topics in Language* (Vol. 68, pp. 179-213). Cambridge, MA: Academic Press. https://doi.org/10.1016/bs.plm.2018.08.006

Nozari, N. (2019). The dual origin of semantic errors in access deficit: activation vs. inhibition deficit. *Cognitive Neuropsychology, 36,* 31-53. https://doi.org/10.1080/02643294.2019.1587397

Nozari, N. (2020). A comprehension- or a production-based monitor? Response to Roelofs (2020). *Journal of Cognition, 3*(1): 19, 1–21. https://doi.org/10.5334/joc.102

Nozari, N. (2021). Neural basis of word production. In L. R. Gleitman, A. Papafragou, & J. C. Trueswell (Eds.), *The Oxford Handbook of The Mental Lexicon* (pp. 552-574). Oxford: Oxford University Press.

https://doi.org/10.1093/oxfordhb/9780198845003.013.12

Nozari, N. (forthcoming). Monitoring, control, and repair in word production. *Nature Reviews Psychology.*

Nozari, N., Arnold, J. E., & Thompson-Schill, S. L. (2014). The Effects of Anodal Stimulation of the Left Prefrontal Cortex on Sentence Production. *Brain Stimulation*, *7*(6), 784–792. https://doi.org/10.1016/j.brs.2014.07.035

Nozari, N., & Dell, G. S. (2009). More on lexical bias: How efficient can a "lexical editor" be? *Journal of Memory and Language*, *60*(2), 291–307. https://doi.org/10.1016/j.jml.2008.09.006

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1084–1090. https://doi.org/10.1037/a0026933

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, *63*(1), 1–33. https://doi.org/10.1016/j.cogpsych.2011.05.001

Nozari, N., Freund, M., Breining, B., Rapp, B., & Gordon, B. (2016). Cognitive control during selection and repair in word production. *Language, Cognition and Neuroscience*, *31*(7), 886–903. https://doi.org/10.1080/23273798.2016.1157194

Nozari, N., & Hepner, C. R. (2019a). To select or to wait? Response to the commentaries. *Cognitive Neuropsychology*, *36*(5–6), 226–233. https://doi.org/10.1080/02643294.2019.1632280

Nozari, N., & Hepner, C. R. (2019b). To select or to wait? The importance of criterion setting in debates of competitive lexical selection. *Cognitive Neuropsychology*, *36*(5–6), 193–207. https://doi.org/10.1080/02643294.2018.1476335

Nozari, N., Martin, C. D., & McCloskey, N. (2019). Is repairing speech errors an automatic or a controlled process? Insights from the relationship between error and repair probabilities in English and Spanish *Language, Cognition, & Neuroscience. 34*(9), 1230-1245. http://dx.doi.org/10.1080/23273798.2019.1637007

Nozari, N., Mirman, D., & Thompson-Schill, S. L. (2016). The ventrolateral prefrontal cortex facilitates processing of sentential context to locate referents. *Brain and Language*, *157–158*, 1–13. https://doi.org/10.1016/j.bandl.2016.04.006

Nozari, N., & Novick, J. (2017). Monitoring and Control in Language Production. *Current Directions in Psychological Science*, *26*(5), 403–410. https://doi.org/10.1177/0963721417702419

Nozari, N., Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics, 53,* 1-19. https://doi.org/10.1016/j.jneuroling.2019.100875

Nozari, N., & Thompson-Schill, S. L. (2013). More attention when speaking: Does it help or does it hurt? *Neuropsychologia*, *51*(13), 2770–2780. https://doi.org/10.1016/j.neuropsychologia.2013.08.019

Nozari, N., & Thompson-Schill, S. L. (2016). Left ventrolateral prefrontal cortex in processing of words and sentences. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 569–584). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00046-8

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Oppenheim, G. M., & Nozari, N. (2024). Similarity-induced interference or facilitation in language production reflects representation, not selection. *Cognition*, *245*, 105720. https://doi.org/10.1016/j.cognition.2024.105720

Özdemir, R., Roelofs, A., & Levelt, W. J. M. (2007). Perceptual uniqueness point effects in monitoring internal speech. *Cognition*, *105*(2), 457–465. https://doi.org/10.1016/j.cognition.2006.10.006

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of stroop word-color interference: Evidence for cingulate subregions subserving multiple distributed attentional systems. *Biological Psychiatry*, *45*(10), 1237–1258. https://doi.org/10.1016/S0006-3223(99)00056-6

Piai, V., Riès, S. K., & Swick, D. (2016). Lesions to Lateral Prefrontal Cortex Impair Lexical Interference Control in Word Production. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00721

Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: Domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00832

Pinet, S. & Nozari, N. (2020). Electrophysiological Correlates of Monitoring in Typing with and without Visual Feedback. *Journal of Cognitive Neuroscience, 32*(4), 603-620. https://doi.org/10.1162/jocn_a_01500

Pinet, S., & Nozari, N. (2021). The role of visual feedback in detecting and correcting typing errors: A signal detection approach. *Journal of Memory and Language*, *117*, 104193. https://doi.org/10.1016/j.jml.2020.104193

Pinet, S., & Nozari, N. (2022). Correction without consciousness in complex tasks: Evidence from typing. *Journal of Cognition*, *5*(1). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8740635/

Pinet, S., & Nozari, N. (2023). Different Electrophysiological Signatures of Similarity-induced and Stroop-like Interference in Language Production. *Journal of Cognitive Neuroscience*, *35*(8), 1329–1349. https://doi.org/10.1162/jocn_a_02014

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*(2), 291–321. https://doi.org/10.1080/01688639508405124

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. https://doi.org/10.1037/a0019737

Riès, S., Janssen, N., Dufau, S., Alario, F.-X., & Burle, B. (2011). General-Purpose Monitoring during Speech Production. *Journal of Cognitive Neuroscience*, *23*(6), 1419–1436. https://doi.org/10.1162/jocn.2010.21467

Roelofs, A. (n.d.). Self-Monitoring in Speaking: In Defense of a Comprehension-Based Account. *Journal of Cognition*, *3*(1), 18. https://doi.org/10.5334/joc.61

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, *172*, 59–72. https://doi.org/10.1016/j.cognition.2017.12.007

Runnqvist, E., Bonnard, M., Gauvin, H. S., Attarian, S., Trébuchon, A., Hartsuiker, R. J., & Alario, F.-X. (2016). Internal modeling of upcoming speech: A causal role of the right posterior cerebellum in non-motor aspects of language production. *Cortex*, *81*, 203–214. https://doi.org/10.1016/j.cortex.2016.05.008

Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language*, *75*, 27–44. https://doi.org/10.1016/j.jml.2014.04.006

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227. https://doi.org/10.1016/j.jml.2005.10.002

Seyfeddinipur, M., Kita, S., & Indefrey, P. (2008). How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition*, *108*(3), 837–842.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences*, *4*(4), 147–155.

Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *Journal of Neuroscience*, *30*(46), 15643–15653.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning, second edition: An Introduction*. MIT Press.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings* *of the National Academy of Sciences*, *94*(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

Titone, C. M. C. D. (1996). Phoneme Monitoring. *Language and Cognitive Processes*, *11*(6), 635–646. https://doi.org/10.1080/016909696387042

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trude, A., & Nozari, N. (2017). Inhibitory control supports referential context use in language production and comprehension. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, 1218–1224.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (2001). What do double dissociations prove? *Cognitive Science*, *25*(1), 111–172. https://doi.org/10.1207/s15516709cog2501_5

Vanderhasselt, M.-A., De Raedt, R., & Baeken, C. (2009). Dorsolateral prefrontal cortex and Stroop performance: Tackling the lateralization. *Psychonomic Bulletin & Review*, *16*(3), 609–612. https://doi.org/10.3758/PBR.16.3.609

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, *128*(3), 442–472. https://doi.org/10.1037/0033-2909.128.3.442

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering Meaning: Left Prefrontal Cortex Guides Controlled Semantic Retrieval. *Neuron*, *31*(2), 329–338. https://doi.org/10.1016/S0896-6273(01)00359-2

Wessel, J. R. (2012). Error awareness and the error-related negativity: Evaluating the first decade of evidence. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00088

Wessel, J. R., & Anderson, M. C. (2024). Neural mechanisms of domain-general inhibitory control. *Trends in Cognitive Sciences*, *28*(2), 124–143. https://doi.org/10.1016/j.tics.2023.09.008

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin & Review*, *19*(4), 699–707. https://doi.org/10.3758/s13423-012-0262-6

Zarr, N., & Brown, J. W. (2016). Hierarchical error representation in medial prefrontal cortex. *NeuroImage*, *124*, 238–247. https://doi.org/10.1016/j.neuroimage.2015.08.063

Zhu, D., Wang, X., Zhao, E., Nozari, N., Notebaert, W., & Braem, S. (2024). *Cognitive control is task specific: Further evidence against the idea of domain-general conflict adaptation*. OSF. https://doi.org/10.31234/osf.io/kgrvd