



On putative shortcomings and dangerous future avenues: response to Strijkers & Costa

Peter Indefrey

To cite this article: Peter Indefrey (2016) On putative shortcomings and dangerous future avenues: response to Strijkers & Costa, *Language, Cognition and Neuroscience*, 31:4, 517-520, DOI: [10.1080/23273798.2015.1128554](https://doi.org/10.1080/23273798.2015.1128554)

To link to this article: <https://doi.org/10.1080/23273798.2015.1128554>

 Published online: 25 Jan 2016.

 Submit your article to this journal [↗](#)

 Article views: 135

 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 3 View citing articles [↗](#)

COMMENTARY

On putative shortcomings and dangerous future avenues: response to Strijkers & Costa

Peter Indefrey^{a,b}

^aDepartment of Linguistics, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; ^bDonders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

In their opinion article, Strijkers and Costa (*in press*, henceforth S&C) raise a number of concerns about the neurocognitive model of word production that Pim Levelt and I proposed (Indefrey & Levelt, 2004; Indefrey, 2011). I agree on one or two points, find our model misrepresented in others and altogether reject their attempt in the second half of their article to undermine the sequentiality that is at the core of the model.

Let me first support S&C's concerns about how our model is sometimes used to interpret the results of studies when their design does not allow for a clear interpretation itself. This type of reference to the model ("we find something in time-window X so it must be lemma access") is indeed highly doubtful and often wrong. It is obvious that observations from such studies cannot support the model and I have said so before (Indefrey, 2011). It is also obvious that this type of reference to the model cannot be held against the model.

Another point in which I agree with S&C is the necessity of top-down modulations in a neural model of word production. The model as presented in I&L (2004) and Indefrey (2011) does not contain any reference to such modulations simply because such modulations were not among the processing components that we could identify by comparing the neural responses elicited by the most frequently used tasks (picture naming, word generation, word and pseudoword reading). In fact, word production starts with a top-down modulation. In the simple case of picture naming the first processing stage, conceptual preparation, means that a speaker must decide on which concept to express (DOG, DALMATIAN, or ANIMAL). At some point the speaker zooms in on the concept he or she wants to express, and it is precisely at this point that this concept receives extra activation, that is, a top-down upregulation. It is this upregulation that results in stronger activation of the corresponding lemma compared to the lemmas of other related concepts and eventually in the selection of this lemma. S&C's assumption that "an activated concept will

automatically spread its activation to the lexical representation with which it is connected, regardless whether there is the intention to name that concept or not" (p. 11) is thus not entirely correct: the activation strength will differ depending on whether there is the intention to name that concept or not. The finding by Strijkers, Holcomb, and Costa (2011) that the intention to speak at all or not has a similar but much more general modulating effect on the degree of activation reaching the lexical representations is thus highly interesting but not at all incompatible with spreading activation models, such as the model of Levelt, Roelofs, and Meyer (1999).

I will now turn to the points in which S&C misrepresent our model. On page 3 they state that I&L assigned conceptual processing to a specific region, namely the occipito-temporal cortex. We do no such thing. Instead we explicitly state that:

Considering the evidence about widespread and possibly category-specific areas being involved in prelexical conceptual processing (...) it seems plausible that tasks like picture naming and word generation, which probably activate quite different concepts, should only converge and enter a common pathway from the point of lexical selection onwards. (Indefrey & Levelt, 2004, p. 123)

Please note that widespread and category-specific conceptual neural representations may include, for example, motor cortex representations related to action semantics. Thus, in contrast to what S&C assume on p. 22, I&L do not intend to claim that motor cortex is only functionally involved near the end of speech preparation. The same holds for other sensorimotor components of concepts. I&L are simply silent about conceptual representations, because the studies they analysed did not allow for the identification of the neural correlates of concepts of any category. In fact, the precise nature of conceptual representations, their components, and their links to the lemma level is somewhat

underspecified in current models of word production and needs to be spelled out in more detail, for example based on richer frame-theoretic representations of concepts (cf. Barsalou, 1982, 2008; Löbner, 2014).

I also find S&C's discussion in Section 1 somewhat misleading. Admittedly that is in part our own fault. S&C correctly state that the set of studies used to identify the anatomical correlates of picture naming contained magnetoencephalography (MEG) studies that were later used to test the model. Indeed there were two MEG studies (Salmeiln, Hari, Lounasmaa, & Sams, 1994; Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998) among the 23 studies on picture naming. However, removing those two studies does not alter the set of regions reliably activated during picture naming, so that a general invalidation of our procedure is not warranted. We identified independently (a) the time course of the processing stages of word production based on reaction time and EEG studies, and (b) the anatomical regions related to the word production stages mostly based on hemodynamic studies. We then used the few available MEG studies that provided both temporal and spatial information to test the time windows for every region predicted from the first two steps. S&C criticise that the MEG studies "had to rely on a-priori assumptions about when the different representational codes of word production are thought to become activated in the course of speech planning" (p. 8). This is not correct. The MEG studies were merely used to test whether the brain regions in question were activated in the predicted time windows. Insofar as the predicted time windows had been incorrect, for example, because the chronometric data did "not stem from speech production behavior under normal conditions" (p. 6) this test would have failed.

I will now turn to what I believe to be the main issue. S&C are somewhat inconsistent with respect to their view on the serial order of brain activations. Initially (p. 4), they acknowledge that there is good evidence for sequentiality in the brain activation patterns. In later parts of their article they come back on that (a) by saying that the temporal order of the earliest (occipitotemporal) and latest (Broca's area and motor cortex) observed MEG activations is trivial ("not so surprising", p. 8) for a picture naming task and (b) by casting doubt on a serial order as such based on evidence from three articles reporting spatiotemporal data that according to S&C speak against sequentiality or suggest a different order of the activation of brain regions (Schuhmann, Schiller, Goebel, & Sack, 2012; Sahin, Pinker, Cash, Schomer, & Halgren, 2009; Miozzo, Pulvermüller, & Hauk, 2014). Schuhmann and colleagues investigated the time windows during which repetitive transcranial magnetic stimulation (TMS) to three target regions

interfered with object naming. They found stimulation of the left mid MTG to affect naming latencies when applied in time windows around 225 and 400 ms after picture onset, stimulation of the left inferior frontal gyrus (LIFG) to affect naming latencies when applied in a time window around 300 ms, and stimulation of the left posterior superior temporal gyrus (STG) to affect naming latencies when applied in a time window around 400 ms. S&C conclude that "The finding that middle temporal gyrus (MTG) activation comes prior to STG and LIFG activation is consistent with I&L's predictions (2004). However, the fact that the LIFG is functionally relevant prior to the STG and that the MTG displays causal involvement at two different moments in time are not consistent with their predictions. Especially the latter result suggests that the time course of cortical area activations is not restricted to a single time frame, but shows patterns consistent with the notions of feedback projections and reentrant activity" (p. 19). S&C do not mention that Schuhmann et al. (2012) themselves interpret the effects of stimulation in IFG as pointing to a functional role of this region in syllabification, and the effects of stimulation in the time window around 400 ms as pointing to a functional role of both mid MTG and posterior STG in self-monitoring. This interpretation is entirely consistent with I&L (2004). Schuhmann et al. (2012) used rather short picture names resulting in mean naming latencies of 455 ms in the no TMS condition. Considering that the temporal estimate of 400 ms for IFG activation in I&L is based on naming latencies of 600 ms, IFG activation around 300 ms is predicted for shorter words. The word production model of Levelt et al. (1999) assumes self-monitoring feedback loops and I&L proposed the left STG to be involved in self-monitoring (I&L, Figure 5). Considering lexical biases in speech errors (Baars, Motley, & MacKay, 1975; Motley & Baars, 1976) it makes a lot of sense that self-monitoring should also involve the lemma level and hence mid MTG. All TMS interference effects observed by Schuhmann et al. (2012) are thus in line with I&L. A possible problematic point is the absence of an earlier interference effect in posterior STG. According to I&L there should be an interference effect for stimulation around 200–300 ms in the posterior STG and MTG. Considering the size of this region and the fact that Schuhmann et al. (2012) only targeted the dorsalmost part of it, it is conceivable that enough of the region remained unaffected to ensure timely word form retrieval.

Sahin et al. (2009) recorded electrophysiological responses from intracranial electrodes in neurosurgical patients during word production. In different conditions, the participants produced isolated words ("lexical processing"), words in a syntactic context ("lexical and grammatical

processing”) or words requiring overt inflection in a syntactic context (“lexical, grammatical, and phonological processing”). Sahin et al. (2009) observed word-frequency dependent lexical responses around 200 ms, “grammatical” responses around 320 ms, and “phonological” responses around 450 ms after stimulus presentation. Crucially, all responses were observed in Broca’s area. According to S&C these findings are problematic for I&L, because they suggest that “brain regions might not have a single time course linked to a single linguistic function” (p. 19) and because I&L “argue that lexical and grammatical processing should be restricted to temporal areas”(p. 19). We never said anything like that. It is true that I&L assume a specialisation of brain regions for the processes involved in word production in that the mid MTG subserves lemma access but not word form retrieval, Wernicke’s area word form retrieval but not lemma access and so forth. However S&C confuse “specialisation” with what I would like to call “specificity”. “Specialisation” means that a Region X supports a function Y that other regions do not support. In other words, this is what in the history of neuroscience is known as the localisationist as opposed to the holistic view of the brain. Other examples for specialised areas in this sense are the motor cortex, the visual cortex and so on. “Specificity” by contrast means that Region X supports a function Y and no other function. I&L never claim that Broca’s area has no other linguistic (or nonlinguistic, see e.g. Binkofski & Buccino, 2004) function besides phonological encoding in word production. In fact, I have always argued and provided evidence for a role of Broca’s area in syntactic processing in both speech production and comprehension (cf. Indefrey et al., 2001, 2004; Indefrey, 2011, 2012, more recently also for a role in sentence-level semantic combinatorial processing, Hagoort & Indefrey, 2014). The same holds for a role of Broca’s area in reading (e.g. Hagoort et al., 1999). Note that the tasks used by Sahin et al. (2009) involved the presentation of a written target word rather than a picture. As I have argued earlier (Indefrey, 2011), the activation observed 200 ms after the target word might well reflect lexical access in word reading (i.e. from a graphemic code) rather than the concept-based lemma access in word production.

S&C discuss the study of Miozzo et al. (2014) as providing evidence in favour of a neural assembly model which, they argue, in contrast to the I&L model not only is able to account for but in fact predicts “near-simultaneous” activation of different properties of words, such as lexico-semantic and phonological properties. Miozzo et al. (2014) modulated the visual complexity of to-be-named pictures, specific semantic features and action features of the depicted concepts, and properties of the word forms of the picture names (length, number of phonological neighbours). They used these variables as regressors in a multiple regression analysis of MEG

data recorded during picture naming. The crucial finding according to the authors and to S&C was an early (150 ms after picture onset) influence of word form properties in posterior MTG that occurred simultaneously with an influence of semantic features in posterior inferior temporal gyrus, (pITG). Upon closer inspection of the data there are some issues that make this particular finding somewhat unconvincing. The region-of-interest that was sensitive to phonological properties (pMTG) was (a) also sensitive to visual complexity in the same time window and (b) immediately adjacent to the region sensitive to semantic features (pITG). Furthermore, the authors report that the phonological variables were highly correlated with word frequency. As a result the phonological variable compared pictures with short, high frequency names such as “Door, car, chair, tie” with pictures showing objects with long, low frequency names such as “Caterpillar, screwdriver, helicopter, strawberry” (see Miozzo et al., 2014, Table 1). Note that word frequency is typically correlated with semantic or visual properties such as familiarity and the authors do not report having controlled for familiarity differences. It seems thus conceivable that the phonological variable might have been confounded.

Although I felt it necessary to comment in detail on the studies that S&C present as problematic for the I&C model, I think that, more generally speaking, a model that is based on a meta-analysis and hence on robust findings across a larger number of studies should not be easily dismissed on the basis of a few as yet unconfirmed findings. A much better test bed addressing the issues that also S&C are interested in is a recent excellent and comprehensive meta-analysis of MEG studies on speech production by Munding, Dubarry, and Alario (2015). Based on the most frequently observed activations across 17 MEG studies, the results (see Munding et al. (2015), Figure 4) on the one hand show a sequentiality of the onset of activation for lemma access in left mid MTG, phoneme retrieval in left posterior STG/MTG and syllabification in left IFG that is well in line with I&L (2004). Activation of motor cortex, however, is frequently observed much earlier than assumed by I&L (2004) and also much earlier than speech onset. In addition, the authors point out that activation in left mid MTG is frequently observed for much longer than predicted by I&L (2004) and that for the regions assumed to be involved in lemma access (mid MTG) and syllabification (IFG) there are also some observations suggesting an almost immediate onset of activation. The prolonged activation of mid STG is in line with the observations by Schuhmann et al. (2012) suggesting a role of this region in self-monitoring (see discussion above). The reliable early

activation of motor cortex and occasional observations of early activations in the other regions are indeed not compatible with a strictly cascading model. One way to accommodate both the observed sequentiality and the early onsets is to distinguish between the information flow captured in the I&L model and additional top-down modulations preparing all involved areas for the task at hand as suggested by Munding and colleagues and also S&C.

I am much more sceptical about the alternative interpretation in terms of a neural assembly model. As S&C mention, such a model predicts simultaneous activation of semantic, lexical, phonological and even articulatory information. The results of Munding et al. (2015), however, show that the most frequently found time windows of activation do not support early simultaneous activation onsets of all regions. By S&C's own standards, this should constitute a serious problem for a neural assembly model. Note furthermore that the undeniable sequentiality observed for the robust time windows is predicted by a psycholinguistic cascading model in combination with brain regions that are specialised for the processing components of the psycholinguistic model. By contrast, according to S&C a neural assembly model allows for sequentiality because "activity can reverberate in local parts of the assembly" (p. 21). Fine, but based on what does such a model predict it?

To conclude, I would like to take up a fundamental remark that S&C make on page 16: "Here we would like to focus on one aspect of this problem namely whether modern advances in neuroscience can indeed be aligned (mapped) with serial processing hierarchies of psycholinguistic function." Do S&C really think that the neuroscientific investigation of word production should decouple from the functional evidence provided by psycholinguistics? I hope not.

Funding

This article was funded by the Deutsche Forschungsgemeinschaft (DFG) [SFB 991]

References

- 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, 382–391.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, 10, 82–93.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Binkofski, F., & Buccino, G. (2004). Motor functions of the Broca's region. *Brain and Language*, 89, 362–369.
- Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, 37, 347–362.
- Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., & Seitz, R.J. (1999). The neural circuitry involved in the reading of German words and pseudowords: A PET study. *Journal of Cognitive Neuroscience*, 11, 383–398.
- Indefrey, P. (2011). Neurobiology of syntax. In P.C. Hogan (Eds.), *The Cambridge encyclopedia of the language sciences* (pp. 835–838). Cambridge, NY: Cambridge University Press.
- Indefrey, P. (2012). Hemodynamic studies of syntactic processing. In M. Faust (Ed.), *The handbook of the neuropsychology of language* (Vol. 1, pp. 209–228). Malden, MA: Blackwell.
- Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., & Hagoort, P. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences of the USA*, 98, 5933–5936.
- Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language*, 89, 312–319.
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92, 101–144.
- Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, 10, 553–567.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.
- Löbner, S. (2014). Evidence for frames from human language. In T. Gamerschlag, D. Gerland, R. Osswald & W. Petersen (Eds.), *Frames and concept types. Applications in language and philosophy* (Vol. 94, Studies in Linguistics and Philosophy, pp. 23–67). Dordrecht: Springer.
- Miozzo, M., Pulvermüller, F., & Hauk, O. (2014). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, 25, 3343–3355.
- Motley, M. T., & Baars, B. J. (1976). Semantic bias effects on the outcomes of verbal slips. *Cognition*, 4, 177–187.
- Munding, D., Dubarry, A.-S., & Alario, F.-X. (2015). On the cortical dynamics of word production: A review of the MEG evidence. *Language, Cognition and Neuroscience*. Advance online publication. doi:10.1080/23273798.2015.1071857
- Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science*, 326, 445–449.
- Salmelin, R., Hari, R., Lounasmaa, O. V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature*, 368, 463–465.
- Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2012). Speaking of which: Dissecting the neurocognitive network of language production in picture naming. *Cerebral Cortex*, 22, 701–709.
- Strijkers, K., & Costa, A. (in press). The cortical dynamics of speaking: Present shortcomings and future avenues. *Language, Cognition and Neuroscience*. doi:10.1080/23273798.2015.1120878
- Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language*, 65, 345–362.