CrossMark click for updates

COMMENTARY

# Neural integration of language production and comprehension

**Martin J. Pickering[a,1] and Simon Garrod[b]**

[a]*Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom; and* [b]*Institute for Neuroscience and Psychology, University of Glasgow, Glasgow G12 8QB, United Kingdom*

Two key assumptions underpin the cognitive neuroscience of language. First, there is a clear-cut split between the processes involved in understanding an utterance (recognizing a word, resolving ambiguity) and the processes involved in crafting that utterance (translating an idea into sound or writing). For example, the "classic" Lichtheim–Broca–Wernicke model proposes distinct anatomical pathways associated with production and comprehension, primarily on the basis of deficit–lesion correlations in aphasia (1). Second, researchers assume that the linguistic mechanisms are lateralized, with production processes (e.g., lexical selection, articulation) and, to some extent, comprehension processes primarily occurring in the left hemisphere. Silbert et al. (2) report a neuroimaging study based on the production and comprehension of naturalistic narrative that challenges these two assumptions.

Using a novel "time-warping" technique, Silbert et al. (2) analyzed the correlation between blood-oxygen level-dependent (BOLD) responses as three participants repeatedly produced and other participants subsequently comprehended a single 15-min narrative. Silbert et al.'s approach contrasts with almost all neuroimaging studies of language processing, which depend on event-related designs and averaging of the neural responses, and which make use of very limited stimuli or responses (e.g., single phonemes, words, or decontextualized sentences). The authors reported all of the brain regions concurrently activated when a speaker repeated the narrative (production) and when a different listener heard the narrative (comprehension). In addition, they identified the brain regions in which the neural responses were coupled between production and comprehension of the same narrative. To accomplish this, Silbert et al. identified regions in which BOLD changes correlated between the speaker's brain and the listener's brain at exactly the same point in the narrative.

Silbert et al.'s (2) most striking findings are as follows. First, an extensive and substantially bilateral network of brain regions active during production and comprehension. These regions comprise both traditional language related networks and traditional nonlinguistic networks (e.g., regions commonly associated with mentalizing). Second, a considerable degree of both overlap and coupling between the regions associated with production and comprehension. And third, some indication that there are regions associated exclusively with either production or comprehension. Silbert et al.'s novel time-warping and intrasubject correlation techniques make a significant contribution to neuroimaging methodology and their findings are relevant for the neural organization of language processing, but perhaps the most important implications are for cognitive theories of linguistic communication.

**Silbert et al.'s study shows tight coupling between much of production and comprehension, which clearly contrasts with the traditional assumption that these processes are separate.**

## Theoretical Importance of Coupling

Silbert et al.'s (2) study shows tight coupling between much of production and comprehension, which clearly contrasts with the traditional assumption that these processes are separate. In fact, the study is compatible with a growing body of theory arguing that tight coupling is necessary to explain many facts about conversation. For example, people find interactive dialogue remarkably straightforward, even though they regularly encounter elliptical utterances, cannot plan their own utterance in advance, and constantly switch between comprehension and production (3). Furthermore, addressees comprehend while providing feedback and speakers comprehend that feedback and use it "on line" in ways that enhance their own contributions. Additionally, addressees take the floor to make their own full contributions with no or minimal delay (4), which is far less time than would be needed to plan a contribution "from scratch." It would be hard to see how such behavior could occur without temporally interwoven production and comprehension. It is, however, striking that the tight coupling and time-locking occurs in Silbert et al. (2), and therefore supports the claim that such interweaving occurs in monologue as well as dialogue.

A number of different recent frameworks emphasize tight coupling between production and comprehension. These frameworks do so by assuming some degree of shared representations (i.e., representational parity): for example, that people draw on a single lexicon during production and comprehension. These frameworks assume that production and comprehension involve some different processes (which underlie the conversion of meaning into sound on the one hand, and sound into meaning on the other hand). However, the frameworks also assume that processes primarily associated with production can be recruited to assist comprehension, and that processes primarily associated with comprehension can be recruited to assist production. Such approaches often explain language learning, acquisition, and change (5, 6) or dialogue (7) rather than isolated acts of production or comprehension. Other studies focus more specifically on speech control (8, 9).

Under many of these accounts, the interweaving of production and comprehension provides critical support for prediction.

Author contributions: M.J.P. and S.G. wrote the paper.

The authors declare no conflict of interest.

See companion article on page E4687.

[1]To whom correspondence should be addressed. Email: martin.pickering@ed.ac.uk.

Downloaded by guest on April 11, 2021

Thus, the P-chain framework (5) assumes that the same process is responsible for prediction during production and prediction during comprehension. Another account proposed that speakers predict linguistic properties (e.g., sound, grammar, meaning) of their own upcoming utterance and then use comprehension mechanisms to determine what the sensory consequences of those utterances should be (7). These sensory estimates can be compared with the actual consequences to provide a control mechanism for spoken language. Moreover, listeners can predict what other speakers are likely to say next by covertly imitating their production processes, deriving the speaker's underlying intention, and using that intention to make the same sorts of predictions that they also make for their own upcoming utterances. This account means that comprehenders are constantly drawing on mechanisms primarily associated with production. If they do have to respond (as in dialogue), comprehenders are ideally placed to do so. This account is therefore compatible with Silbert et al.'s (2) finding of coupled activation of production- and comprehension-based networks.

Importantly, such predictions cannot merely involve using linguistic context to predict upcoming words or grammatical forms. Instead, producers and comprehenders must base prediction on higher-level nonlinguistic social information, such as whether the social context requires an addressee to produce an elaborate response, or whether an utterance should be interpreted literally or not (10); such predictions would therefore result in the activation of areas associated with mentalizing, as Silbert et al. found (2).

Interweaving production and comprehension may have other functions as well as prediction. For example, comprehenders may rapidly construct production-based representations of utterances to facilitate ambiguity resolution (11), memory (perhaps via rehearsal), or inference. Such proposals are also compatible with the reported results (2). However, it may be harder to reconcile the results with the assumption that producers construct a sound-based representation that they then comprehend (12). This is because such "inner-loop monitoring" is relatively slow and should lead to reduced coupling between production and comprehension (as well as additional activation in production relative to comprehension).

## Shared Representations at Different Linguistic Levels

Silbert et al.'s (2) findings are compatible with shared representations at higher linguistic levels, such as the grammar and meaning of both words and longer utterances. For example, picture-naming (a production task) is inhibited by simultaneous comprehension of a spoken word related in meaning (13). Similarly, speakers who have just comprehended an utterance with a particular grammatical form then tend to repeat that grammatical form (14), without being aware they are doing so. Silbert et al.'s (2) study shows how such shared representations lead to shared neural activation in a way that has clear effects during extended language use.

However, although high-level linguistic representations may be shared between production and comprehension, Silbert et al.'s (2) results suggest greater separation of representations at lower levels concerned with sound. In fact, there is much less agreement about whether sound-based representations are shared across production and comprehension. Thus, picture-naming can be facilitated by simultaneous comprehension of a spoken word related in sound to the picture's name (13). Moreover, there is some motor activation during speech perception, some of which is articulator-specific (15). However, such activation may depend on the task used (16). Additionally, some evidence suggests that producers imitate phonetic features that they hear (17) but other evidence does not (18). It seems most likely that production and comprehension share some but not all sound-based representations. This finding would be compatible with motor activation being enhanced when comprehension is difficult (19), or with the possibility that production-based representations are involved in particular processes, such as timing and turn-taking.

In conclusion, Silbert et al.'s (2) study provides strong evidence for a close linking between production and comprehension processes, as well as shared representations, primarily at high linguistic and nonlinguistic levels. We believe that a major challenge for cognitive and neural theories of language processing is to determine precisely how these processes are linked.


**1** Shalom DB, Poeppel D (2008) Functional anatomic models of language: Assembling the pieces. *Neuroscientist* 14(1):119–127.
**2** Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U (2014) Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc Natl Acad Sci USA* 111:E4687–E4696.
**3** Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. *Behav Brain Sci* 27(2):169–190, discussion 190–226.
**4** Stivers T, et al. (2009) Universals and cultural variation in turn-taking in conversation. *Proc Natl Acad Sci USA* 106(26):10587–10592.
**5** Dell GS, Chang F (2014) The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Phil Trans R Soc B* 369(1634):20120394.
**6** Macdonald MC (2013) How language production shapes language form and comprehension. *Front Psychol* 4:226.
**7** Pickering MJ, Garrod S (2013) An integrated theory of language production and comprehension. *Behav Brain Sci* 36(4):329–347.
**8** Hickok G, Houde J, Rong F (2011) Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* 69(3):407–422.
**9** Tourville JA, Guenther FH (2011) The DIVA model: A neural theory of speech acquisition and production. *Lang Cogn Process* 26(7): 952–981.
**10** Bašnáková J, Weber K, Petersson KM, van Berkum J, Hagoort P (2013) Beyond the language given: The neural correlates of inferring speaker meaning. *Cer Cortex* 24(10):2572–2578.
**11** Garrett M (2000) Remarks on the architecture of language production systems. *Language and the Brain: Representation and Processing*, eds Grodzinsky Y, Shapiro LP (Academic, San Diego), pp 31–69.
**12** Levelt WJM (1989) *Speaking: From Intention to Articulation* (MIT Press, Cambridge, MA).
**13** Schriefers H, Meyer AS, Levelt WJM (1990) Exploring the time course of lexical access in language production: Picture-word interference studies. *J Mem Lang* 29(1):86–102.
**14** Branigan HP, Pickering MJ, Cleland AA (2000) Syntactic co-ordination in dialogue. *Cognition* 75(2):B13–B25.
**15** Pulvermüller F, Fadiga L (2010) Active perception: Sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci* 11(5): 351–360.
**16** Scott SK, McGettigan C, Eisner F (2009) A little more conversation, a little less action—Candidate roles for the motor cortex in speech perception. *Nat Rev Neurosci* 10(4):295–302.
**17** Goldinger SD (1998) Echoes of echoes? An episodic theory of lexical access. *Psychol Rev* 105(2):251–279.
**18** Mitterer H, Ernestus M (2008) The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109(1):168–173.
**19** Adank P (2012) The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain Lang* 122(1):42–54.





Downloaded by guest on April 11, 2021