



OPEN ACCESS

EDITED BY

Gian Marco Duma,
Eugenio Medea (IRCCS), Italy

REVIEWED BY

Antje Meyer,
Max Planck Society, Germany
Jarosław Lełonkiewicz,
Pompeu Fabra University, Spain
Eleanor Huizeling,
Max Planck Institute for Psycholinguistics,
Netherlands

*CORRESPONDENCE

Simone Gastaldon
✉ simone.gastaldon@unipd.it

† These authors share first authorship

RECEIVED 12 January 2024

ACCEPTED 06 May 2024

PUBLISHED 21 May 2024

CITATION

Gastaldon S, Bonfiglio N, Vespignani F and
Peressotti F (2024) Predictive language
processing: integrating comprehension
and production, and what atypical
populations can tell us.
Front. Psychol. 15:1369177.
doi: 10.3389/fpsyg.2024.1369177

COPYRIGHT

© 2024 Gastaldon, Bonfiglio, Vespignani and
Peressotti. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Predictive language processing: integrating comprehension and production, and what atypical populations can tell us

Simone Gastaldon^{1,2*†}, Noemi Bonfiglio^{1,3†},
Francesco Vespignani^{1,4} and Francesca Peressotti^{1,2,4}

¹Dipartimento di Psicologia dello Sviluppo e della Socializzazione, University of Padua, Padua, Italy, ²Padova Neuroscience Center, University of Padua, Padua, Italy, ³BCBL–Basque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain, ⁴Centro Interdipartimentale di Ricerca “I-APPROVE–International Auditory Processing Project in Venice”, University of Padua, Padua, Italy

Predictive processing, a crucial aspect of human cognition, is also relevant for language comprehension. In everyday situations, we exploit various sources of information to anticipate and therefore facilitate processing of upcoming linguistic input. In the literature, there are a variety of models that aim at accounting for such ability. One group of models propose a strict relationship between prediction and language production mechanisms. In this review, we first introduce very briefly the concept of predictive processing during language comprehension. Secondly, we focus on models that attribute a prominent role to language production and sensorimotor processing in language prediction (“prediction-by-production” models). Contextually, we provide a summary of studies that investigated the role of speech production and auditory perception on language comprehension/prediction tasks in healthy, typical participants. Then, we provide an overview of the limited existing literature on specific atypical/clinical populations that may represent suitable testing ground for such models—i.e., populations with impaired speech production and auditory perception mechanisms. Ultimately, we suggest a more widely and in-depth testing of prediction-by-production accounts, and the involvement of atypical populations both for model testing and as targets for possible novel speech/language treatment approaches.

KEYWORDS

language prediction, language production, prediction-by-production, speech-motor impairment, auditory impairment, atypical language processing

1 Introduction

Despite the seemingly ubiquitous unpredictability of language in everyday situations, research shows that there is enough regularity in language that we can anticipate (i.e., pre-activate) upcoming linguistic information at various scales, independently of its modality—spoken, written, or signed—before it is available in the environment for processing (Federmeier, 2007, 2021; Van Petten and Luka, 2012; Pickering and Garrod, 2013;

Huettig and Mani, 2016; Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018; Huettig et al., 2022; Radošević et al., 2022; Ryskin and Nieuwland, 2023). Imagine you're at a birthday party, someone is bringing the birthday cake at the table, and you hear them saying "Let's-"; your brain is likely to fill in the sentence with "eat the cake" before the speaker even says it. Linguistic and non-linguistic contextual cues, together with our language- and world-knowledge, guide us in anticipating what is most likely to come next. The information we can predict during processing can range from higher (conceptual) to lower (word orthography, phonology) levels and everything in between (lexical semantics, syntactic features, morphological markings) (but see Ryskin and Nieuwland, 2023, for debates on the nature of predicted representations; in particular for phonological representations, see DeLong et al., 2017; Ito et al., 2017a,b; Nieuwland et al., 2018; Nicenboim et al., 2020; Urbach et al., 2020). A variety of behavioral and neural measures have been associated with predictive language processing. Eye-tracking studies during reading revealed that words that can be predicted are skipped entirely or are fixated for a shorter interval relative to words that cannot be predicted (for a review, see Staub, 2015; for a parafoveal parallel processing interpretation of predictable words skipping, see Staub and Goddard, 2019). In visual world paradigms, where participants listen to sentences while looking at a set of pictures, the eyes fixate the target picture before the referent is mentioned in the sentence in the case of predictable words (for a seminal work pioneering this research, see Altmann and Kamide, 1999; for a review, see Huettig et al., 2011). Electrophysiological measures revealed that words embedded into sentences whose content make such words predictable elicit a different event-related potential (ERP) responses compared to words that are not predictable (i.e., a reduced N400; for reviews, see Kutas and Federmeier, 2011; Van Petten and Luka, 2012; Nieuwland et al., 2020). Crucially, time-frequency analyses of magneto/electroencephalographic (M/EEG) data also showed differences *before* predictable words are read or listened (i.e., a power decrease in the alpha—8–10 Hz—and beta—15–20 Hz—frequency bands), more directly reflecting the prediction stage and specifically the pre-activation of upcoming linguistic information (see e.g., Rommers et al., 2017; Molinaro and Monsalve, 2018; Wang et al., 2018; Gastaldon et al., 2020; León-Cabrera et al., 2022; but see Huizeling et al., 2023 for challenges to such benchmark correlates when testing in a virtual environment). The fact that comprehenders, regardless of modality (e.g., spoken, written, signed) anticipate information yet to be available in the environment is undisputed; however, there is no consensus on the mechanisms and resources employed to this aim (e.g., Huettig, 2015; Kuperberg and Jaeger, 2016; Huettig et al., 2022; Ryskin and Nieuwland, 2023). In this manuscript we aim at unifying two points of view that we believe can provide fruitful advancement in the study of predictive processing in the language domain and that in our opinion have not been sufficiently explored.

Firstly, we focus on a group of models that aim at accounting for predictive processing during language comprehension by adopting an integrated framework in which language comprehension and production are not seen as separate domains. Such models—which for sake of simplicity we will collectively call “prediction-by-production models”, despite some architectural differences—do not propose that prediction is carried out *always and solely* by engaging resources usually devoted to language production; rather, they

suggest that the language production infrastructure *can* be involved for more efficient predictions when the situation and the resources allow for it. Note that not all the models we describe here put language production as the *main* mechanism behind prediction, but they all *incorporate* it.

Secondly, we discuss how the study of sensorimotor and language systems that are dysfunctional or that are atypically developed may be particularly enlightening. Atypical populations with speech-motor and auditory perception impairments offer a fertile ground for clarifying to what extent the (spoken) language system is integrated, and consequently to determine the impact that specific deficits have on linguistic prediction. In conclusion, we bring it all together and suggest how research in predictive language processing may advance by leveraging on these two neglected aspects.

2 An integrated view: the role of speech production and sensorimotor processing in prediction during comprehension

Linguistic prediction is often considered a mechanism deployed during the process of understanding linguistic input, therefore at the service of comprehension. However, in the last decade, researchers have been moving away from the traditional separation between language comprehension and production as distinct domains toward a more integrative view. In this perspective, comprehension and production share representations and processes, at least partially (e.g., AbdulSabur et al., 2014; Dell and Chang, 2014; Pickering and Garrod, 2014; Silbert et al., 2014; Gambi and Pickering, 2017; McQueen and Meyer, 2019; Walenski et al., 2019; Fairs et al., 2021). The idea is not entirely new: in the past, models in which production processes are hypothesized to be implemented for speech perception were proposed, such as the Motor Theory of Speech Perception (Liberman, 1957) and the Analysis-by-Synthesis theory (Halle and Stevens, 1962). In this perspective, perceiving and categorizing speech sounds involve not only perceptual auditory mechanisms/brain networks, but also mechanisms/brain networks devoted to the production of such sounds. While the Motor Theory of Speech Perception has faced significant criticism (see e.g., Galantucci et al., 2006; Massaro and Chen, 2008), the Analysis-by-Synthesis (AxS) framework seems to have regained relevance in recent years (e.g., Poeppel et al., 2008; Poeppel and Monahan, 2011; Skipper et al., 2017; Preisig et al., 2022). AxS envisages top-down contextual inferential processes deployed by the production system that influence the bottom-up perceptual categorization of speech sounds. This view is closely related to the idea of forward models implemented for self-monitoring during speech production. During the act of speaking, the activation of representations flows in a top-down fashion through the linguistic hierarchy. A pre-verbal message is encoded, followed by the selection of lexical items to appropriately encode a linguistic message. Morphosyntactic information is then retrieved, phonological sequences are encoded and transformed into speech-motor sequences for articulation. From the motor sequences, sensory predictions are computed

to monitor the speaker's own output and correct possible errors (Indefrey and Levelt, 2004; Hickok et al., 2011; Indefrey, 2011; Tourville and Guenther, 2011; Hickok, 2012; Strijkers and Costa, 2016). In the rest of the manuscript, when mentioning "production system" (or simply "production"), we refer to the neural, cognitive, motor, and linguistic resources and mechanisms recruited and deployed to implement speech production. Therefore, we do not refer to an encapsulated system, but to a pool of resources and processes which can be partially shared to implement other tasks (McQueen and Meyer, 2019).

This top-down information flow (from higher to lower level representations) has also been proposed by some researchers to underlie predictions during comprehension of other people's speech as well (Pickering and Garrod, 2007, 2013). There are other proposals that ascribe a prominent role to production in speech prediction (Huettig, 2015; Pickering and Gambi, 2018); however, these models are not entirely in agreement regarding which processes and representations are involved. In this section, we present three main models of "prediction-by-production," highlighting points of convergence and divergence between them. We then provide an overview of studies on typical and neurologically healthy participants testing the relationship between speech production and prediction. Importantly, some studies suggest that motor contributions to prediction may be more crucial in challenging listening situations (e.g., Schomers and Pulvermüller, 2016; Skipper et al., 2017); therefore, we also present an overview of studies investigating the contextual top-down effects of prediction when the input is degraded and more difficult to process (e.g., speech-in-noise, vocoded speech).

2.1 Prediction-by-production models

The first attempts at providing an integrated view in which production processes are involved in prediction may be ascribed to Federmeier (2007) PARLO framework, and Pickering and Garrod (2007) emulator proposal. However, more extensively fleshed-out models emerged in the literature only later. Here we present three models that have been extensively developed allowing for the formulation of specific testable hypotheses (for other proposals of links between production and prediction, see also Dell and Chang, 2014; McCauley and Christiansen, 2019).

Pickering and Garrod (2013) (P&G2013) propose an integrated framework that envisions language production and comprehension as forms of action and action perception, respectively. In an attempt to overcome traditional separations, they suggest distinguishing comprehension and production based on the direction of information mapping. Specifically, production processes map from higher to lower levels of the linguistic hierarchy (e.g., from semantics to phonology), while comprehension processes map from lower to higher representational levels (e.g., from phonology to semantics). Given the top-down nature of predictions, they are therefore considered instances of production. The authors draw directly from the literature on action control, where performing an action involves predicting its outcome (Wolpert, 1997; Wolpert and Flanagan, 2001). When performing an action (e.g., moving a hand), an action command is formulated, and two parallel processes are initiated: (1) the action command is sent to the action implementer

in order to perform the plan and construct the percept of the action; (2) an efference copy is generated from the action command and used to generate a forward model, predicting the sensory consequences of the action sequence (predicted percept). Percept and predicted percept are then compared, and any discrepancies are exploited to modify the action plan. As mentioned above, forward models are suggested to underlie self-monitoring during speech production (Hickok et al., 2011; Tourville and Guenther, 2011; Hickok, 2012; Franken et al., 2018; Okada et al., 2018). However, the debate on how self-monitoring is implemented is still open, often with points of view focusing on psycholinguistic (e.g., internal and external loops) or motor aspects (see e.g., Nozari and Novick, 2017; Lind and Hartsuiker, 2020; Nozari, 2020; Roelofs, 2020a,b; but see Hickok, 2012, 2014 for attempts to reconcile the psychoneurolinguistic and the motor perspectives). P&G2013 capitalized on the parallelism between language production as a form of action and language comprehension as a form of action perception for elaborating their proposal. Forward models are generated from a production command representing the situation model and the communicative intent. Crucially, these models are extended at all levels of linguistic representations (i.e., semantics, syntax and phonology), and constitute "impoverished" [sic] representations, enabling rapid generation of predictions. This means that a person can predict the identity of the phonemes of a word but not necessarily encode their order in the sequence or predict the syntactic category of a word but not necessarily other features (e.g., grammatical gender for a noun, or tense for a verb). The key aspect in this proposal is that forward models are not employed solely to predict the speaker's own utterances, but also to predict others' utterances during comprehension (prediction-by-simulation). Specifically, the comprehender covertly imitates the speaker's communicative intent and runs the intention through their own production system to generate forward models and predict the speaker's utterance. While this is proposed as the main mechanism for prediction, the framework assumes that comprehenders can predict also through mere association (prediction-by-association), i.e., semantic priming, based on the experience in comprehending other's sentences. In this case, the production system is not involved.

The model proposed by Huettig (2015) differs from P&G2013 in some respects. In Huettig's (2015) proposal, predictions result from multiple mechanisms that can interact during comprehension (production-, association-, combinatorial-, simulation-based prediction; hence, PACS). Comprehenders sometimes engage their production system to predict what the speaker is likely to say, but they utilize fully-fledged production representations rather than impoverished forward models as in P&G2013 proposal. Additionally, associative mechanisms (priming) at different levels of information (e.g., semantic, phonological, and orthographic, but also at non-linguistic levels) allow for quick pre-activation of information about an incoming input. Combinatorial rules that are likely shared by comprehension and production and that are sensitive to multiple linguistic constraints further guide the pre-activation of information (e.g., determine the syntactic features of upcoming words). Finally, event simulation is exploited as heuristic strategy: the ability to imagine events and simulate their outcome based on previous experience contributes to formulate predictions. Importantly, these mechanisms represent multiple ways to predict and engaging all of them is not a necessary prerequisite for

predicting. They are not encapsulated mechanisms, but they greatly interact with each other. For example, the quick pre-activation through association can serve as input for the combinatorial mechanisms responsible for higher-level structure building, which in turn can generate input for the production system. Furthermore, pre-activation from different mechanisms could sometimes be conflicting. The situational context in which comprehension takes place is essential for appropriate prediction and determines the extent to which each mechanism is exploited. According to [Huettig \(2015\)](#), prediction, while an important aspect, is not fundamental in language comprehension (see also [Huettig and Mani, 2016](#)), and the author stresses the need to investigate more thoroughly the mediating factors that contribute to predictive behavior (e.g., working memory resources, age, literacy).

[Pickering and Gambi \(2018\)](#) (P&G2018) proposed a model that is somewhere in between P&G2013 and PACS model. They suggest two routes for predicting: prediction-by-association (PA) and prediction-by-production (PP). PA automatically triggers during the incremental construction of comprehension representations and is based on the automatic spreading of activation between linguistic levels (e.g., semantic and phonological priming). While not demanding many resources, PA is somewhat inefficient as it activates all connected representations regardless of their relevance in the specific context. This is where PP comes into play. The comprehension representations computed up to that point, including those pre-activated by PA, are transformed into production representations through covert imitation. In other words, the utterance is converted into the message that the comprehender would have formulated, leading to the inference of communicative intention. This process takes into account the non-linguistic context and the comprehender's knowledge about the speaker. The intention is then fed to the production implementer, corresponding to the stages of language production (retrieval of lexical items, lemmas, and phonological information). PP is assumed to be optional and to require more resources than PA. Under conditions of low cognitive resources or time constraints, comprehenders may not engage the production system or go through all the stages of the production implementer, as it is time and resource consuming. This implies that slower input presentations are associated with greater contextual facilitation, which might reflect a more precise representation of the predicted word. In the framework proposed by [Pickering and Gambi \(2018\)](#), this association is explained by the time required to the comprehender to reproduce the sentence in their production system: when the rate of word presentation is faster, early stages of production (involving conceptual formulation and lexical selection) are more likely to be completed than later stages (such as phonological encoding). In such situations the comprehender does not have enough time to predict the form of the word (its phonology or orthography) before encountering it, although they are able to predict its semantics. The role of time constraints on "prediction-by-production" also implies that slower producers are less likely to predict words at later stages of representation or to engage in active prediction at all. P&G2018 also note that the goal of the comprehender is an additional factor influencing the likelihood of prediction, as shown by experiments manipulating task instructions ([Kuperberg and Jaeger, 2016](#)). In everyday situations, the goal of comprehension is usually to infer the message conveyed by the speaker, and in such contexts both high cognitive

load and individual characteristics make active prediction less likely.

In summary, all three proposals assume an important role of priming (or prediction-by-association) as a separate and automatic route for prediction. Associative links between representations, consolidated over time, quickly pre-activate connected information at multiple levels. This activation can then be further constrained. Event simulation – and a specific form of simulation, i.e., covert imitation, – is another central aspect shared by all three proposals. However, there are differences in how simulation is viewed in each model. For P&G2013 and P&G2018 simulation is an integral part of the act of production itself and can be likened to the conceptual formulation of traditional psycholinguistic models of word production ([Indefrey and Levelt, 2004](#); [Indefrey, 2011](#)). On the other hand, the PACS model sees simulation as a separate mechanism that optionally interacts with production representations and other mechanisms. Another distinction lies in the production mechanisms involved. P&G2013 assigns a prominent role to impoverished representations in the form of forward models, which are generated from a conceptual stage ("production command") encoding communicative intention and situation model. In contrast, both PACS and P&G2018 propose the involvement of the fully-fledged production representations ("production implementer"). Importantly, all three models propose that the production route for prediction is optional due to its cognitive cost, even though it leads to more efficient and contextually appropriate predictions.

2.2 Evidence on the relationship between speech production and prediction

The increasing interest in prediction-by-production models has been accompanied by a growing literature highlighting the close relationship between production and predictive processes. Several studies have reported indirect and correlational evidence of such a relationship. For example, production abilities have been shown to partly explain individual differences in prediction: adults who are faster in categorical fluency are more likely to show patterns of neural activity associated with prediction during comprehension ([Federmeier et al., 2010](#)), while children with wider production vocabulary show higher proportion of predictive gaze shifts while listening to sentences ([Mani and Huettig, 2012](#)). This latter finding is in line with a study showing that infants are better at identifying syllables that are part of their babbling repertoire ([Vilain et al., 2019](#)), highlighting the importance of individual production capabilities in consolidating speech perception and, likely, prediction. Furthermore, from a physiological point of view, tongue evoked motor potentials have been found to be modulated by phoneme expectation during word listening ([D'Ausilio et al., 2011](#)), ERPs in a phonological priming study highlighted the involvement of speech motor regions in generating phonology-specific predictive negative slow waves ([Grisoni and Pulvermüller, 2022](#)), and neural oscillatory modulations in the alpha-beta ranges in prediction during speech comprehension show spatiotemporal correlations with modulations during speech planning ([Gastaldon et al., 2020](#)). This body of evidence suggests a contribution of speech-motor processes in prediction. Indirect evidence comes

also from the so-called production effect in memory. Rommers et al. (2020) showed that the memory disadvantage for words read silently relative to words read aloud is reduced for predictable relative to unpredictable words. This suggests that in predictable contexts silent reading becomes more similar to aloud reading, thus producing similar facilitatory effects in later memory. Additionally, Hadley et al. (2020) showed that listeners are more accurate to predict turn ends when they listen to recordings of themselves or of a speaker rated similar to themselves relative to a speaker rated as dissimilar, who might be more difficult to imitate in their own production system. However, all these studies did not manipulate speech production in prediction during comprehension.

More direct evidence of prediction-by-production mechanisms comes from studies manipulating the engagement of the production system and measuring the related effect on prediction during comprehension. The study by Martin et al. (2018) studied ERPs in EEG data in participants reading sentences in Spanish while taxing the production system through articulatory suppression. The experiment manipulation capitalized on the effect by which predictable words elicit a less negative N400 (a negative deflection of the brain potentials peaking around 400 ms after word onset and emerging in centro-parietal sensors). Participants in the critical experimental group had to produce syllables while reading sentence contexts; in the control groups they either had to just move the tongue or to listen to the syllables. Critically, expected and unexpected target words differed in gender and the relative gender-specific article. The ERP results showed that the N400 effect at pre-noun articles was not significant in the syllable production group, while it was significant in the two control groups. In contrast, the N400 effect at the target noun did not vary across groups. Consistently with the importance placed on covert imitation by the prediction-by-production approach, this pattern demonstrates that production mechanisms are exploited to build lexical predictions, particularly during context reading. The early articulatory suppression and the subsequent lack of covert imitation *online*, while the sentence was unfolding over time, might have prevented participants from inferring the intended message, which in turn would have been necessary to build a forward model (Pickering and Garrod, 2013) or to pass such intention through the production system for prediction (Pickering and Gambi, 2018).

While the presence of the control groups in Martin et al. (2018) checked for the role of motor action and feedback perception, it is still possible that syllable production might have interfered with verbal working memory specifically in the experimental group. In this case, lack of prediction would be explained by the decreased availability of verbal working memory rather than by the unavailability of the production system itself. Given this potential confound, Lelonkiewicz et al. (2021) tested the role of production in prediction by exploiting a complementary approach, i.e., examining the effects of *enhancing* rather than suppressing the engagement of the production system during comprehension. In this study, participants were instructed to read high- or low constraining sentence contexts either aloud or silently, thus engaging the production system to different extents. Tasks at target word differed across experiments: word recognition, word naming, picture naming. Results showed that the facilitation given by high constraining contexts relative to low constraining ones was more prominent in the reading aloud condition, suggesting that enhancing the engagement of the production system during

context reading enhances predictive processes. This result can be easily interpreted in the light of the integrated framework proposed by P&G2013. Reading aloud requires performing the action of producing the words through the articulatory system, and this action is accompanied by forward models used for self-monitoring. The availability of forward models might in turn make it easier to predict the possible continuation of the sentence, relative to a condition in which forward models are not explicitly required to perform the task. The results are also consistent with the PACS models and P&G2018, which both propose the involvement of the fully-fledged production representations. The overt articulation required by the reading aloud condition might have rendered such representations easily accessible for prediction-by-production. The study by Lelonkiewicz et al. (2021) also shed light on the nature of such predictions. Reading the sentence aloud was associated with shorter response times but did not increase accuracy, suggesting that production leads to stronger but not necessarily more accurate predictions. This pattern aligns with a prediction-by-production mechanism: when using representations from our own production system to infer the underlying message of the sentence, predictions may not always perfectly reflect the actual continuation intended by the interlocutor. Nevertheless, the production mechanism still helps to constrain the number of predicted continuations (relative to other mechanisms like association), resulting in smoother response selection and shorter response times. In the study, the strength of the interaction between predictability and reading mode (silent vs. aloud) was different according to the type of task performed on the target word: advantage for reading aloud was present in the word and picture naming tasks (at the end of a sentence frame) but absent in the lexical decision task. This pattern is consistent with the proposal that production is not the only possible mechanism for prediction, but its engagement is modulated flexibly according to the situation at hand. Specifically, when production is explicitly required for the task to perform, forward models and production representations might be more readily available to exploit prediction-by-production. In contrast, when the task is limited to comprehension (e.g., lexical decision), other forms of prediction might be implemented, possibly quicker but representationally less detailed. However, in an omnibus analysis the authors find an interaction between predictability and reading mode, but no interaction with the task/experiment. The authors therefore suggest that there is no strong evidence for differential facilitation depending on the degree to which the production system is engaged. In line with the former interpretation, Hintz et al. (2016) found that the facilitatory effects of constraining contexts on reading times are enhanced when reading trials are interleaved with picture naming trials, relative to a condition in which the two tasks are performed in separate blocks. This result suggests that prediction is favored in situations in which the production system is overall more activated. Importantly, the authors note that everyday conversations are constituted by a similar alternation of comprehension and production. Thus, prediction-by-production mechanisms might be more prominent in everyday life than in experimental settings testing solely comprehension.

It is important to note that at least one study claimed to have found evidence *against* the involvement of language production in prediction. Brothers et al. (2023) studied ERPs elicited when reading sentences that could constraint toward two possible

candidates (expected and second-best). The main finding of interest here is that there was no difference in the ERPs of these two possible sentence continuations. From this empirical finding, the authors conclude that language prediction works in a different manner than language production, because, in the latter, lexical selection happens by competition via lateral inhibition (i.e., lexical items inhibit each other; Levelt et al., 1999). If language prediction worked as language production (by means of competitive pre-activation), a cost should have emerged in processing the second-best candidate, a cost that does not emerge in the data. While valuable, these findings do not automatically and incontrovertibly prove that language production has no role in prediction, because this interpretation does not automatically follow from the findings, but depends on the kind of models one assumes for production (lexical competition by lateral inhibition vs. response selection) (Spalek et al., 2013). In a model where there is no competition by lateral inhibition in word production (Dell, 1986; Mahon et al., 2007; Navarrete et al., 2014), the findings by Brothers et al. (2023) would not be taken as evidence against the deployment of production processes during prediction. Furthermore, in the study, no measure of individual language production abilities is put in relation with the ERPs measured during comprehension. Therefore, we believe the interpretation put forward by Brothers et al. (2023) should be taken more cautiously.

Overall, there is both direct and indirect evidence that production processes play a *contributory* role in linguistic prediction. The extent and relevance of such contributions need to be further specified.

2.3 Evidence on the role of prediction in degraded auditory input

Top-down and bottom-up information dynamically interact during speech comprehension, and several studies in the field of prediction have explored these interactions. Here, we will specifically focus on studies examining prediction in conditions where auditory perception is suboptimal. While our consideration is restricted to studies using auditory stimuli, it is essential to note that in everyday life, listeners take into account multiple sources of information, therefore language comprehension is highly visually-situated and context-dependent (Knoeferle, 2019). These factors have been shown to be especially important when the auditory input is suboptimal, due to environmental conditions or hearing difficulties (Peelle and Sommers, 2015; Beauchamp, 2016; Stevenson et al., 2017). In such situations, visual information may modulate the listener's reliance on top-down knowledge, an observation that has not been observed in clear speech situations with ERP measures (Hernández-Gutiérrez et al., 2018). However, measures of cortical tracking of speech showed that audiovisual models better explain brain signals of participants listening to audiovisual narratives relative of audio only or visual only (Crosse et al., 2015).

Top-down processes have been proposed to aid speech comprehension in challenging listening conditions by directly influencing low-level perceptual processing, rather than being used for a post-hoc revision of the perceived word (Davis and Johnsruide, 2007). In line with this view, when listeners know in

advance the content of the sentence, they perceive noisy speech as clearer and estimate the background noise as being quieter, even though this estimation should be based on basic sound perception. Additional evidence of top-down influences on perception comes from perceptual learning phenomena, i.e., the improvement in recognition of degraded speech with exposure to it. Perceptual learning generalizes to words that were not previously heard in degraded form, suggesting that it consists in “re-tuning” phonetic representations, regardless of the specific words in which they are embedded. Importantly, for words embedded in sentences (as opposed to pseudowords), listeners seem to correctly re-tune phonetic representations from the degraded input itself, with no necessity of a concomitant presentation in a clear form. The authors hypothesize that prediction might have a crucial role in this process: by pre-activating relevant representations, prediction could constrain the interpretation of degraded speech, and make the correct phonetic representation available to support perceptual retuning.

Corps and Rabagliati (2020) explored the influence of prediction on recognition of noise-vocoded (NV) speech, differentiating between phonological and semantic predictions. They presented constraining or unconstraining questions in clear form (e.g., *What colors are pandas?* vs. *What colors should I paint the wall?*) before each NV stimulus. In the first case, participants could anticipate a precise answer, including its phonological form (e.g., *black and white*); in the second case they could just anticipate the semantic field of the answer (e.g., colors). Answers were presented in NV form at 50% intelligibility and could be predictable or unpredictable with respect to the questions. Results showed that, relative to the unpredictable condition, form specific predictions did not enhance recognition of NV speech beyond predictions about the semantic field. Furthermore, both types of predictions equally enhanced perceptual learning and generalization to unseen NV stimuli. These results indicate that prediction of high-level information, such as semantics, aids speech perception and perceptual learning. Such facilitation might be especially relevant in everyday life, when listeners lack access to clear phonological forms in the presence of noisy or degraded speech, relative to semantic information from the context, which is more easily available.

The modalities in which comprehenders process the sentence context strongly influence the generation of predictions. This holds true in the auditory domain. Studies utilizing degraded sentences (as opposed to contexts presented in clear or written form, e.g., Corps and Rabagliati, 2020; Van Os et al., 2022) revealed that predictability effects interact with the level of spectral degradation. Facilitation in speech recognition is observed only for moderately degraded stimuli (e.g., Obleser and Kotz, 2010; Bhandari et al., 2021). When sentences are minimally degraded and easily intelligible, prediction has no facilitatory effects on recognition since listeners can comprehend each word regardless of its predictability (Bhandari et al., 2021). This condition can be equated to clear speech conditions, where prediction has facilitatory effects at subsequent stages of processing, increasing speed and efficiency. On the opposite end of the spectrum, high levels of degradation make sentence context hardly intelligible, hindering listeners' ability to use the context to form predictions. In this case, even when listeners can understand the context, the increased listening effort required to process the auditory input leaves few

cognitive resources available for prediction (Bhandari et al., 2021). This view is in accordance with prediction-by-production accounts, which consider production as a costly and optional mechanism for prediction, dependent on the features of the bottom-up input. Furthermore, at moderate levels of intelligibility (i.e., with recognition scores significantly above chance but below ceiling), there is evidence that excitability in the speech-motor areas is higher than when speech is fully intelligible, and such excitability is related to identification accuracy of distorted syllables (Nuttall et al., 2016). These results suggest that motor activation plays a critical role in the comprehension of distorted speech, supporting the notion that prediction can aid speech perception through production mechanisms.

The interaction between predictability and level of intelligibility has also been studied also by exploiting speech masked with noise. Marrufo-Pérez et al. (2019) tested speech recognition, manipulating the stimuli around the individual speech reception threshold (SRT) which is the level of signal-to-noise ratio at which a participant recognizes 50% of the sentences. At this intelligibility level, recognition scores were found to decrease as the sentence unfolded, with initial words recognized better than the words in following positions. This decrease in recognition scores suggests that increase in word predictability is not beneficial at SRT levels, possibly even being detrimental to speech recognition. The authors explored this possibility by modeling the probability of recognizing each word given the recognition (or not recognition) of the previous one. This analysis showed that words were significantly less likely to be recognized when the preceding word was missed or misunderstood, and that this effect overrode the beneficial effect of recognizing the preceding word. However, with increasing intelligibility above the SRT the detrimental effect gradually disappeared, indicating that when speech is clearer listeners are less likely to be misled by the preceding word and rely more on the bottom-up input they perceive (for a similar discussion see also Van Os et al., 2022). Overall, this pattern indicates that beneficial or detrimental effects of predictability strongly depend on the level (and type) of noise as well as on the individual sensitivity to speech-in-noise (indicated here by the individual SRT). In real-life situations, speech-to-noise ratios are sufficiently higher than SRTs for normal hearing listeners, but they might not be so for hearing-impaired listeners; thus, predictability might have different effects in normal and impaired hearing listeners.

It is interesting to note that also older adults exploit contextual information and stored knowledge to facilitate speech comprehension, thus highlighting the relevance of context and prediction in situations in which auditory perception has gradually become less efficient due to normal aging (Pichora-Fuller, 2008). Furthermore, a recent study on audiovisual syllable perception in noise in older adults unveiled the crucial contribution of sensorimotor regions for successful performance: participants who received musical training in their life performed better than those who did not have such training, and this facilitation was due to stronger engagement of sensorimotor regions (Zhang et al., 2023).

In conclusion, experimental studies show that prediction can aid speech perception when the auditory input is sub-optimal; however, the extent by which such facilitation is possible and efficient strongly depends on the features of the bottom-up input. Together with evidence suggesting the involvement of the speech motor system in aiding perception especially in noisy or

auditory degraded conditions (for a review, see Skipper et al., 2017), we believe that a further understanding of how speech production processes contribute to prediction is to be pursued, in particular to empirically test the models presented in section “2.1 Prediction-by-production models.”

3 What atypical and clinical populations can tell us about prediction

In analogy with the classical approach in cognitive neuropsychology (Caramazza and Coltheart, 2006), studying how a particular process emerges in populations with impairments in different relevant mechanisms can help us understand the extent to which the impaired mechanism matters in the intact brain. For instance, studying linguistic prediction in populations with impairments in processes (and brain structures and function) mainly devoted to speech production can help us specify their contribution to prediction during comprehension. Conversely, studying prediction in populations with deficits in encoding the auditory input due to acquired or congenital dysfunctions can shed light on the role of the quality of the sensory input in relation to top-down cognitive processes and demands. In the following section, we discuss the limited literature available, specifically focusing on people with developmental stuttering and people with Parkinson’s disease for production, and cochlear implant mediated speech in deaf people for perception. Such populations provide a unique testing ground for theoretical models. Importantly, findings stemming from such research may turn out to have clinical applications for novel treatment approaches.

3.1 Inefficient speech production

One population that can provide valuable insights on the role of the speech-motor system to prediction during comprehension is people who stutter. Developmental stuttering (DS) is a multifactorial neurodevelopmental disorder that disrupts speech fluency, leading to blocks, repetitions, and prolongations of sounds (Smith and Weber, 2017). The neuroimaging literature consistently showed that DS is associated with atypical gray and white matter patterns in regions involved in speech-motor control (for a review, see Etchell et al., 2018). Specifically, people with DS exhibit an inefficient mapping between motor and sensory predictions, an overly inhibited feedforward motor system, and an inefficient internal timing/sequence organization (e.g., Alm, 2004; Max et al., 2004; Civier et al., 2013; Busan, 2020). Such dysfunctions emerge as atypical modulations of alpha-beta activity in the electro-magnetoencephalographic signal during speech production (e.g., Mersov et al., 2016; Mock et al., 2016; Jenson et al., 2018). Therefore, these processes and their electrophysiological correlates are relevant from the point of view of the prediction-by-production models we discussed above. To explore whether impaired production causes altered predictive processing, Gastaldon et al. (2023) investigated prediction during spoken language comprehension while recording EEG in adults

with DS. The study found that, compared to age-matched fluent speakers, adults who stutter showed atypical patterns of electrophysiological correlates of word prediction during comprehension, namely a reduced N400 effect and an increase of alpha-beta power in premotor and frontal regions. The former finding suggests that the way in which predicted and unpredicted words are integrated is different between controls and adults who stutter. The second finding implies that this difference may be due to less detailed predictions, likely as a result of a lack of (or reduced) involvement of the inefficient speech-motor network. Furthermore, correlations of alpha/beta activity between comprehension and production tasks in cortical regions of interest showed that, while positive correlations were found in left inferior frontal and temporal regions in fluent participants, suggesting possible convergence of mechanisms in the two tasks in those cortical regions (see also [Gastaldon et al., 2020](#)), no consistent correlations were found in adults with DS. This further suggests that when speech-motor processes are sub-optimal, they are not appropriately co-opted for prediction during comprehension. Speculatively, such “impoverished” predictions may emerge as lexico-semantic predictions with no phonological and sensory component.

Another clinical population that can shed light on the role of motor control in linguistic prediction is people with Parkinson’s disease (PD). PD is a neurodegenerative syndrome that initially affects motor control, leading to bradykinesia, rest tremor, and/or rigidity, but later progresses to also impact cognitive and emotional-affective domains ([Bloem et al., 2021](#)). [León-Cabrera et al. \(2021\)](#) studied prediction during reading in people with PD, both with normal cognition and with mild cognitive impairment, compared to matched controls. The authors recorded EEG and focused on two ERPs relevant for prediction: the prediction negative potentials (PNP) before final words and the N400 at the final word. The authors found no group differences between controls and people with PD and normal cognition. However, differences were found between people with PD with and without cognitive impairment: the latter showed no PNP effects and a prolongation of N400 effects, suggesting that general cognitive decline is responsible for impaired predictive processing. Interestingly however, the PNP effect in people with PD (including those with normal cognition) was correlated with verbal fluency. This is also in line with findings on healthy older adults, showing that participants with high verbal fluency generate a pattern of prediction-related ERPs effects similar to that of young adults, differently from older adults with low fluency scores ([Federmeier et al., 2010](#); [Dave et al., 2018](#)). While verbal fluency encompasses a variety of functions, it is considered a measure of speech production efficiency, further linking language production with prediction during language comprehension. This suggests that further investigation into the relationship between speech production in PD and linguistic prediction may be valuable. Note that the authors did not analyze oscillatory activity, which may capture effects that are not easily detected with ERPs ([Gastaldon et al., 2020](#)). Given the relevance of alpha-beta activity also in PD ([Belova et al., 2021](#); [Johari and Behroozmand, 2021](#)), we think this complementary approach could yield fruitful insights.

It is interesting to note that the alpha-beta aberrant activity that is observed both in DS and PD has been linked to a dysregulation of dopamine levels in the cortico-basal ganglia-thalamo-cortical loop that is responsible for initiating speech ([Chang and Guenther, 2020](#);

[Alm, 2021](#)). Studying the role of neurotransmitter levels in key speech production regions may be a promising research direction to further understand the neurophysiological bases of the processes that are hypothesized to be involved in predictive language processing.

To the best of our knowledge, the literature on such topic is still very limited and hopefully developing. Nevertheless, we argue that testing the effects of impaired speech/language production mechanisms on prediction is a first step for delineating a more complete picture of predictive processing and the language system as a whole, beyond the separation between comprehension and production.

3.2 Sub-optimal auditory input: cochlear implant-mediated speech

To investigate the extent to which people rely on predictive processing based on the quality of auditory input, researchers can examine deaf people with cochlear implant (CI). CIs are neural prostheses that are commonly used to treat severe and profound sensorineural deafness, enabling sound perception even in presence of damages or congenital malfunctioning of the cochlea ([Macherey and Carlyon, 2014](#)). CIs support speech perception in deaf people and play a crucial role in the development of spoken language skills in deaf children, particularly if they are implanted before the age of three ([Hunter and Pisoni, 2021](#)). However, while in many cases CIs provide sufficient information for speech comprehension, the auditory input conveyed is qualitatively degraded compared to the natural input of a healthy cochlea (for an in-depth discussion of the acoustic features of the CI input, see [Hunter and Pisoni, 2021](#)). In this section, we will review studies comparing predictive processes in CI users and normal hearing (NH) listeners, to elucidate the role of the auditory input in prediction. We will consider both studies on postlingually and prelingually deaf CI users; however, it is worth noting that speech processing mechanisms might differ between these two populations. Postlingually deaf CI users have acquired language with a normal auditory input and later adapted to the CI input after experiencing hearing loss. On the other hand, prelingually deaf CI users, who represent the majority of participants in studies with children, have developed language abilities solely on the basis of the CI input, potentially leading to differences in linguistic and cognitive development compared to NH peers (see, for example, [Hunter and Pisoni, 2021](#)). Thus, the features of the specific sample tested must not be overlooked.

In line with studies on NH listeners using degraded sentences (e.g., [Obleser and Kotz, 2010](#); [Bhandari et al., 2021](#)), some authors propose that the limited amount of speech cues available to CI users might compromise their capacity to represent context efficiently and generate predictions about upcoming words ([Baskent et al., 2016](#)). Moreover, more effortful speech processing might result in lack of time and cognitive resources sufficient to use semantic cues and predictive processes ([Holt et al., 2021](#); see also [Bhandari et al., 2021](#)). [Winn \(2016\)](#) utilized pupil dilatation as an online measure of cognitive effort during sentence listening. Postlingually deaf CI users were compared with NH controls who listened to clear speech and spectrally degraded speech matching the intelligibility of the CI input. The results indicated that constraining sentence

contexts were associated with a reduction in pupil dilatation, supporting the notion that efficient prediction reduces the overall cognitive resources expense. The effect was found in both groups of participants, but with different timing: pupillary reduction appeared before the end of the sentence for NH listeners when exposed to clear speech, and after the end of the sentence for CI users and for NH listeners exposed to degraded speech. A possible explanation might be that context is not used predictively by CI users, but rather to facilitate a post-hoc restoration of misperceived words in the sentences. Highly constraining contexts allow listeners to infer words that have been misperceived during online processing, potentially leading to a release from the cognitive effort exerted to maintain multiple lexical options active.

A study conducted by Nagels et al. (2020) examined postlingually deaf CI users and NH controls in a visual word paradigm. Participants were shown images of the target word and of distractors, including phonological and semantic competitors of the target word. When predictive strategies were employed, looking times toward semantic competitors increased, while looking times toward phonological competitors decreased compared to a not-constraining condition. This pattern was observed for both the CI and the NH groups, but quantitatively different according to the group. In particular, CI users showed a higher proportion and duration of fixations toward the competitors relative to NH controls. The group of CI users also showed great individual variability in the size of the effect of constraining context. The authors further explored these individual differences by administering an auditory lexical decision task, from which they extracted an individual measure of uncertainty in the interpretation of the speech signal (i.e., “lexical uncertainty”). This measure reflects difficulties in lexical competition and in subsequent lexical access from a speech input. In the visual world paradigm, participants with higher lexical uncertainty displayed more pronounced differences between constraining and not-constraining conditions, suggesting that reliance on context might serve as a compensatory strategy particularly for the individuals who face difficulties in the interpretation of the auditory input. At the same time, higher lexical uncertainty was associated with longer time courses of lexical competition, suggesting that the delay found in the group analysis was primarily driven by this subgroup of CI users.

In addition to time and cognitive resources, a crucial factor in determining the engagement of prediction (and of costly mechanisms as prediction-by-production in particular) is its utility (Kuperberg and Jaeger, 2016). When interpretation of the speech signal is challenging, misunderstandings of previous words might lead to prediction errors (see e.g., Marrufo-Pérez et al., 2019). In such a case, delaying and reducing commitment to predictions might be an adaptive strategy to avoid the costs of revising predictions that turn out to be incorrect (Blomquist et al., 2021). Supporting this notion, a study on single-word recognition revealed that prelingually deaf adolescents with CI adopt a wait-and-see strategy, delaying access to the relevant lexical representation until the auditory input becomes sufficiently informative (McMurray et al., 2017). Importantly, the same study found that NH controls exposed to degraded input exhibited a similar pattern of behavior, suggesting that delaying lexical access might be a strategy driven by the specific features of the input. Blomquist et al. (2021) tested the presence of predictive strategies in children with and without CI

in a visual world paradigm, and found that constraining contexts, compared to unconstraining contexts, elicited faster looks to the target and reduced looks to the phonological competitor before target onset in both groups. However, children with CI showed a smaller proportion of looks to the target and a greater proportion of looks to the competitor relative to controls. These results indicate that children with CI can predict, but they might reduce commitment to the formed predictions relative to their NH peers.

An important consideration when interpreting findings from prelingually deaf CI users is that early exposure to the CI input might lead to significant differences in linguistic development (Hunter and Pisoni, 2021). For example, individual vocabulary measures seem to partly explain group differences in predictive processes, suggesting that children with CI with better linguistic knowledge predict more similarly to their NH peers (Blomquist et al., 2021). Interestingly, a study by Holt et al. (2021) with the visual world paradigm did not find any group difference between NH children and CI users, who exhibited similar timing of lexical access and speech processing in both constraining and not-constraining conditions. This result indicates that there might be a great variability in individual outcomes after early implantation, both in terms of linguistic abilities and neural adaptation (Hunter and Pisoni, 2021). Indeed, for children with reduced difficulty in identifying words, building and committing to predictions might be as convenient as for their NH peers.

Another source of variability might be the specific contextual cues present in the sentence. For example, Davies et al. (2023) tested the use of syntactic cues, especially plural/singular subject-verb agreement, with a variation of the visual world paradigm in which children looked at pictures of singular and plural subjects, one of which was coherent with the sentence they heard. The results showed that children with CI use subject-verb agreement to facilitate lexical access, but they do so more slowly than their NH peers. Importantly, participants in this study greatly overlapped with the ones in Holt et al. (2021), suggesting that the difference in results is driven by the type of information predicted (i.e., semantic vs. syntactic) rather than by differences between participants. Specifically, syntactic cues might be more challenging to recognize when processing speech through a CI.

In conclusion, results from studies on CI users are consistent with studies on NH listeners using degraded speech, suggesting that the quality of the auditory input can impact listeners' use of optional mechanisms of prediction. Results also suggest that prediction is not an all-or-nothing phenomenon: there might be a spectrum of predictive strategies implemented differently between individuals. In the case of CI users, not only the auditory input might play a role, but also its integration with information from other sensory modalities, e.g., visual cues. CI users have been shown to have better integration abilities, especially in presence of higher linguistic information (Rouger et al., 2007; Strelnikov et al., 2009; Stevenson et al., 2017). Currently ongoing research from our lab is going to explore predictive processes in CI users with less artificial stimuli, exploring possible interactions with audiovisual integration processes, and the role of the speech-motor network in aiding prediction and integration.

4 Discussion and future perspectives

The role of prediction in language comprehension has long been recognized and it has been a hot topic for many years now. In this review and perspective we first focused on prediction-by-production frameworks which have been developed in recent years in an attempt to reconcile the traditional and arbitrary divide between language comprehension and production research (McQueen and Meyer, 2019). One crucial aspect that emerges from the proposals summarized in section “2.1 Prediction-by-production models” is that the “production route” is costly and optional: under this proposal, production processes (ranging from intention formulation, lexical selection to phonological encoding and motor programming) are deployed to reinforce prediction only when it is convenient (i.e., when there is an advantage in doing so) and according to the specific situation at hand, in contrast to passive spreading of activation, which happens automatically and without costs. This picture envisages at least a “dual route” (or “multi route” as in Huettig’s PACS model) view of prediction: one is more automatic (spreading of activation / priming / prediction-by-association), the other is costlier (action-/simulation-based prediction / prediction-by-production). Dual-routing cognitive processes is a classical “easy” solution for cognitive scientists to explain complex nonlinear or apparently contradictory behaviors, successfully applied in different fields of cognition such as decision making (heuristic vs. analytic; Samson and Voyer, 2012) and word reading (lexical vs. phonological routes; Humphreys and Evett, 1985). It might be even more plausible to consider the presence of multiple prediction pathways, such as those proposed in the PACS model, under the assumption that various input features can be predicted using distinct mechanisms. For instance, one can easily predict that a noun will follow the English word “the” (maybe after some adjectives), or that after the word “an” a word-initial vowel will follow. In other instances, one can predict the specific phonemes in a word after the uniqueness point has been reached. Other kinds of predictions require rather complex inferential processes and specific world-knowledge (e.g., in the sentence “Caterina is planning to write a grant proposal on vocal communication in . . .” one can predict the word “crows” if they know the specific Caterina it’s been talking about and her specific scientific interests). Other predictions require a more linguistically informed long-term memory information for idioms (e.g., “nine” comes after the sentence fragment “After receiving the results of the exam, Giovanni was on cloud. . .”). Intuitively, it appears clear that all these forms of linguistic predictions may rely on different processes and resources, linguistic and non-linguistic, but research still has not tackled this.

Given the picture above, where at least two routes are outlined, one question that arises then is how this view can be reconciled with the predominant framework that envisages the brain as a “predictive machine,” namely predictive coding (PC) (Friston, 2005, 2010; Bar, 2009; Clark, 2013). Despite being often cited in the literature in language processing, PC is often and wrongly used as a synonym for “predictive processing” in general (as pointed out in Nour Eddine et al., 2022). Instead, PC proposes very specific neural computations as a fundamental algorithm for brain functioning, according to which the brain constantly generates predictions to infer the hidden causes of sensory phenomena in

the environment, by exploiting (and updating) internal generative models. The aim of the brain is to minimize prediction error (the discrepancy between top-down predicted representation and bottom-up signal). This algorithm is proposed to be instantiated in a very specific cortical hierarchy, with state units that encode predictions sending information to lower-levels, and separate error units sending information upward to update the internal model. This formulation makes prediction an automatic process that happens constantly because it is a necessity of the biological system and its core mechanism for encoding information. An increasing number of studies have highlighted that predictive processing seems to be a key element of language processing emerging beyond controlled experimental paradigms, but also in more naturalistic listening and reading situations (e.g., Donhauser and Baillet, 2020; Schrimpf et al., 2021; Brodbeck et al., 2022; Heilbron et al., 2022; Tuckute et al., 2024). Furthermore, recent computational models that strictly adhere to the PC algorithm successfully simulated ERP modulations (especially of the N400) for a variety of effects, although not all of them and with architectural limitations (e.g., some layers are directly fed with values instead of being the results of previous computations, thus failing to replicate other aspects of language processing that are exploited for prediction) (Nour Eddine et al., 2022, 2024). This body of evidence consistently supports the notion that top-down predictions are regularly generated, at least in literate and typical populations, and potentially in alignment with PC algorithms.

How do prediction-by-production models fit in this perspective? According to Pickering and Garrod (2013) and Pickering and Gambi (2018), the production route is the one that provides the best predictions in terms of appropriateness of context and specificity (i.e., down to phonological and sensory information); however, it may not be always possible to implement production-based predictions. Multiple routes—like in the PACS model—seem to be even more difficult to reconcile with a limited set of cortical operations put forward by PC. We believe this is a crucial aspect that research in psycholinguistics and cognitive neuroscience of language need to tackle, that is, to better understand (or confute) the contribution of production processes in prediction in order to also better understand the relation with other proposed mechanisms of prediction and general brain functioning.

Here is why we propose that the study of atypical populations, which is currently underdeveloped as we highlighted in Section “3 What atypical and clinical populations can tell us about prediction,” should be further pursued, since it could contribute to “complete the puzzle.” In particular, investigating prediction in the presence of inefficient language/speech production processes can shed light on what is the nature of predicted representations, possibly identifying a grading of preactivated information. For instance, while providing evidence that developmental stuttering negatively impacts prediction and tentatively providing an explanation on the basis of the neural and behavioral emergence of stuttering, the study by Gastaldon et al. (2023) leaves open the question of what the exact impact on predicted representations is. It is possible that inefficiency of forward model mapping in DS yields unspecified predictions from a sensory point of view. Further research should clarify this point. One could further envisage a research enterprise in which a spectrum of language and speech-motor deficits could be studied, ranging from central linguistic

processes (impairments in lexical selection and phonological encoding, such as in certain profiles of aphasia or developmental language disorders; Denes and Pizzamiglio, 1998; Schwartz, 2010) to more peripheral deficits (such as lack of facial/oral muscle control). For instance, studies on facial paralysis seem to suggest an impairment in emotional faces processing (Sessa et al., 2022; Japee et al., 2023). As a parallelism, one could argue that paralysis of muscle districts responsible for sound production may impact sensory aspects of linguistic processing. A study on people with Moebius syndrome—a congenital neurological disorder causing complete face paralysis which causes impairments in articulating speech sounds (Pamplona et al., 2020)—did not find support for the hypothesis that lack of muscle control impacts lip reading, an aspect closer to face-to-face linguistic communication (Vannuscorps et al., 2021). This suggests that it is unlikely that damage to or atypical development of more peripheral mechanisms of speech and language production affects relevant aspects of speech processing. This remains an empirical question to explore. Conversely, further studying speech comprehension in people with varying degree of hearing loss and deaf people with cochlear implant can clarify the role of top-down preactivation. Importantly, what can be learnt from research on production deficits may be important from a clinical standpoint for populations with hearing impairments. For instance, given the inherently suboptimal nature of the acoustic input conveyed by the implant and the involvement of speech-motor simulation in difficult listening situations (Skipper et al., 2017), it may be particularly relevant for people with CI to exploit production-based mechanisms for prediction. For instance, Sherafati et al. (2022), by using functional near-infrared spectroscopy, found that the left prefrontal cortex was more activated in CI users relative to controls when listening to speech. The authors interpreted this activation as domain-general processes supporting perception and comprehension. However, given the extension of their region of interest into the inferior frontal gyrus, they cannot exclude other language-related processes. Future research should clarify whether CI exploits their production machinery to support comprehension, and under which circumstances. Overall, we believe that the study of atypical/clinical populations, despite the inherent difficulties (e.g., arduous participant recruitment, small sample sizes, etc.), can be fruitful both from a theoretical standpoint, but importantly also from a societal and applied perspective, by fostering public awareness and a deeper understanding of the speech/language deficits, and leading to new approaches to therapy.

In conclusion, we propose that the prediction-by-production perspective needs to be taken into consideration more systematically, faced in a theoretically rigorous way, and more widely empirically tested, also by including atypical populations that can help clarify the picture. This can contribute both to a better understanding of the language system and its relation with

mechanisms of brain functioning, and suggest new venues for speech and language intervention.

Author contributions

SG: Conceptualization, Project administration, Writing – original draft, Writing – review and editing. NB: Conceptualization, Writing – original draft, Writing – review and editing. FV: Conceptualization, Supervision, Writing – review and editing. FP: Conceptualization, Funding acquisition, Supervision, Writing – review and editing.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by a PRIN grant from the Italian Ministry of University and Research awarded to FP (project PRIN 2017 protocol no. 20177894ZH, “The role of cochlear implantation and bimodal bilingualism in early deafness: a window into the neurofunctional mechanisms of human language”). SG was supported by a grant from the Fondazione CARIPARO through the call PHD@UNIPD at the University of Padova (grant no. CUP_C93C23003190005, “Predictive Brain in Audiovisual Speech Comprehension”). Open Access funding provided by Università degli Studi di Padova | University of Padua, Open Science Committee.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., et al. (2014). Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex* 57, 107–127. doi: 10.1016/j.cortex.2014.01.017
- Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *J. Commun. Disord.* 37, 325–369. doi: 10.1016/j.jcomdis.2004.03.001
- Alm, P. A. (2021). The dopamine system and automatization of movement sequences: A review with relevance for speech and stuttering. *Front. Hum. Neurosci.* 15:663. doi: 10.3389/FNHUM.2021.661880
- Altmann, G. T. M., and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73, 247–264. doi: 10.1016/S0010-0277(99)00059-1

- Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1181–1182. doi: 10.1098/rstb.2008.0321
- Baskent, D., Clarke, J., Pals, C., Benard, M. R., Bhargava, P., Saija, J., et al. (2016). Cognitive compensation of speech perception with hearing impairment, Cochlear implants, and aging. *Trends Hear.* 20:2331216516670279. doi: 10.1177/2331216516670279
- Beauchamp, M. S. (2016). “Audiovisual speech integration: Neural substrates and behavior,” in *Neurobiology of language*, eds G. Hickok and S. L. Small (Cambridge, MA: Academic Press), 515–526. doi: 10.1016/B978-0-12-407794-2.00042-0
- Belova, E. M., Semenova, U., Gamaleya, A. A., Tomskiy, A. A., and Sedov, A. (2021). Is there a single beta oscillation band interfering with movement in Parkinson's disease? *Eur. J. Neurosci.* 54, 4381–4391. doi: 10.1111/EJN.15257
- Bhandari, P., Demberg, V., and Kray, J. (2021). Semantic predictability facilitates comprehension of degraded speech in a graded manner. *Front. Psychol.* 12:3769. doi: 10.3389/fpsyg.2021.714485/BIBTEX
- Bloem, B. R., Okun, M. S., and Klein, C. (2021). Parkinson's disease. *Lancet* 397, 2284–2303. doi: 10.1016/S0140-6736(21)00218-X
- Blomquist, C., Newman, R. S., Huang, Y. T., and Edwards, J. (2021). Children with cochlear implants use semantic prediction to facilitate spoken word recognition. *J. Speech Lang. Hear. Res.* 64, 1636–1649. doi: 10.1044/2021_JSLHR-20-00319
- Brodbeck, C., Bhattasali, S., Cruz Heredia, A. A. L., Resnik, P., Simon, J. Z., and Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. *Elife* 11:e72056. doi: 10.7554/ELIFE.72056
- Brothers, T., Morgan, E., Yacovone, A., and Kuperberg, G. (2023). Multiple predictions during language comprehension: Friends, foes, or indifferent companions? *Cognition* 241:105602. doi: 10.1016/J.COGNITION.2023.105602
- Busan, P. (2020). Developmental stuttering and the role of the supplementary motor cortex. *J. Fluency Disord.* 64:105763. doi: 10.1016/j.jfludis.2020.105763
- Caramazza, A., and Coltheart, M. (2006). Cognitive neuropsychology twenty years on. *Cogn. Neuropsychol.* 23, 3–12. doi: 10.1080/02643290500443250
- Chang, S., and Guenther, F. H. (2020). Involvement of the cortico-basal ganglia-thalamocortical loop in developmental stuttering. *Front. Psychol.* 10:28. doi: 10.3389/fpsyg.2019.03088
- Civier, O., Bullock, D., Max, L., and Guenther, F. H. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain Lang.* 126, 263–278. doi: 10.1016/j.bandl.2013.05.016
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477
- Corps, R. E., and Rabagliati, H. (2020). How top-down processing enhances comprehension of noise-vocoded speech: Predictions about meaning are more important than predictions about form. *J. Mem. Lang.* 113:104114. doi: 10.1016/j.jml.2020.104114
- Crosse, M. J., Butler, J. S., and Lalor, E. C. (2015). Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *J. Neurosci.* 35, 14195–14204. doi: 10.1523/JNEUROSCI.1829-15.2015
- D'Ausilio, A., Jarmolowska, J., Busan, P., Bufalari, I., and Craighero, L. (2011). Tongue corticospinal modulation during attended verbal stimuli: Priming and coarticulation effects. *Neuropsychologia* 49, 3670–3676. doi: 10.1016/J.NEUROPSYCHOLOGIA.2011.09.022
- Dave, S., Brothers, T. A., Traxler, M. J., Ferreira, F., Henderson, J. M., and Swaab, T. Y. (2018). Electrophysiological evidence for preserved primacy of lexical prediction in aging. *Neuropsychologia* 117, 135–147. doi: 10.1016/j.neuropsychologia.2018.05.023
- Davies, B., Holt, R., and Demuth, K. (2023). Children with hearing loss can use subject-verb agreement to predict during spoken language processing. *J. Exp. Child Psychol.* 226:105545. doi: 10.1016/J.JECP.2022.105545
- Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283
- Dell, G. S., and Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20120394. doi: 10.1098/rstb.2012.0394
- DeLong, K. A., Urbach, T. P., and Kutas, M. (2017). Is there a replication crisis? Perhaps. Is this an example? No: A commentary on Ito, Martin, and Nieuwland (2016). *Lang. Cogn. Neurosci.* 32, 966–973. doi: 10.1080/23273798.2017.1279339
- Denes, G., and Pizzamiglio, L. (1998). *Handbook of clinical and experimental neuropsychology*. London: Psychology Press, doi: 10.4324/9781315791272
- Donhauser, P. W., and Baillet, S. (2020). Two distinct neural timescales for predictive speech processing. *Neuron* 105:385–393.e9. doi: 10.1016/j.neuron.2019.10.019
- Etchell, A. C., Civier, O., Ballard, K. J., and Sowman, P. F. (2018). A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *J. Fluency Disord.* 55, 6–45. doi: 10.1016/j.jfludis.2017.03.007
- Fairs, A., Michelas, A., Dufour, S., and Strijkers, K. (2021). The same ultra-rapid parallel brain dynamics underpin the production and perception of speech. *Cereb. Cortex Commun.* 2, 1–17. doi: 10.1093/TEXCOM/TGAB040
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505. doi: 10.1111/j.1469-8986.2007.00531.x
- Federmeier, K. D. (2021). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology* 59:e13940. doi: 10.1111/PSYP.13940
- Federmeier, K. D., Kutas, M., and Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain Lang.* 115, 149–161. doi: 10.1016/j.bandl.2010.07.006
- Franken, M. K., Eisner, F., Acheson, D. J., McQueen, J. M., Hagoort, P., and Schoffelen, J.-M. (2018). Self-monitoring in the cerebral cortex: Neural responses to small pitch shifts in auditory feedback during speech production. *Neuroimage* 179, 326–336. doi: 10.1016/j.neuroimage.2018.06.061
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychon. Bull. Rev.* 13, 361–377. doi: 10.3758/BF03193857
- Gambi, C., and Pickering, M. J. (2017). “Models linking production and comprehension,” in *The handbook of psycholinguistics*, eds E. M. Fernández and H. S. Cairns (Hoboken, NJ: John Wiley & Sons, Inc), 157–181. doi: 10.1002/9781118829516.ch7
- Gastaldon, S., Arcara, G., Navarrete, E., and Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex* 133, 328–345. doi: 10.1016/j.cortex.2020.09.026
- Gastaldon, S., Busan, P., Arcara, G., and Peressotti, F. (2023). Inefficient speech-motor control affects predictive speech comprehension: Atypical electrophysiological correlates in stuttering. *Cereb. Cortex* 33, 6834–6851. doi: 10.1093/cercor/bhad004
- Grisoni, L., and Pulvermüller, F. (2022). Predictive and perceptual phonemic processing in articulatory motor areas: A prediction potential & mismatch negativity study. *Cortex* 155, 357–372. doi: 10.1016/J.CORTEX.2022.06.017
- Hadley, L. V., Fisher, N. K., and Pickering, M. J. (2020). Listeners are better at predicting speakers similar to themselves. *Acta Psychol.* 208:103094. doi: 10.1016/j.actpsy.2020.103094
- Halle, M., and Stevens, K. (1962). Speech recognition: A model and a program for research. *IRE Trans. Inf. Theory* 8, 155–159. doi: 10.1109/TIT.1962.1057686
- Heilbron, M., Armeni, K., Schoffelen, J. M., Hagoort, P., and De Lange, F. P. (2022). A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 119:e2201968119. doi: 10.1073/PNAS.2201968119/SUPPL_FILE/PNAS.2201968119.SAPP.PDF
- Hernández-Gutiérrez, D., Abdel Rahman, R., Martín-Loeches, M., Muñoz, F., Schacht, A., and Sommer, W. (2018). Does dynamic information about the speaker's face contribute to semantic speech processing? ERP evidence. *Cortex* 104, 12–25. doi: 10.1016/J.CORTEX.2018.03.031
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145. doi: 10.1038/nrn3158
- Hickok, G. (2014). Towards an integrated psycholinguistic, neurolinguistic, sensorimotor framework for speech production. *Lang. Cogn. Neurosci.* 29, 52–59. doi: 10.1080/01690965.2013.852907
- Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* 69, 407–422. doi: 10.1016/j.neuron.2011.01.019
- Hintz, F., Meyer, A. S., and Huettig, F. (2016). Encouraging prediction during production facilitates subsequent comprehension: Evidence from interleaved object naming in sentence context and sentence reading. *Q. J. Exp. Psychol.* 69, 1056–1063. doi: 10.1080/17470218.2015.1131309
- Holt, R., Bruggeman, L., and Demuth, K. (2021). Children with hearing loss can predict during sentence processing. *Cognition* 212:104684. doi: 10.1016/j.cognition.2021.104684
- Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Res.* 1626, 118–135. doi: 10.1016/J.BRAINRES.2015.02.014
- Huettig, F., and Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Lang. Cogn. Neurosci.* 31, 19–31. doi: 10.1080/23273798.2015.1072223
- Huettig, F., Audring, J., and Jackendoff, R. (2022). A parallel architecture perspective on pre-activation and prediction in language processing. *Cognition* 224:105050. doi: 10.1016/J.COGNITION.2022.105050

- Huetting, F., Rommers, J., and Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychol.* 137, 151–171. doi: 10.1016/j.actpsy.2010.11.003
- Huizinga, E., Alday, P. M., Peeters, D., and Hagoort, P. (2023). Combining EEG and 3D-eye-tracking to study the prediction of upcoming speech in naturalistic virtual environments: A proof of principle. *Neuropsychologia* 191:108730. doi: 10.1016/j.neuropsychologia.2023.108730
- Humphreys, G. W., and Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behav. Brain Sci.* 8, 689–705. doi: 10.1017/S0140525X00045684
- Hunter, C. R., and Pisoni, D. B. (2021). “Some neuromyths and challenging questions about cochlear implants,” in *The handbook of speech perception*, 2nd Edn, eds J. S. Pardo, L. C. Nygaard, R. E. Remez, and D. B. Pisoni (Hoboken, NJ: Wiley-Blackwell), 540–569.
- Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Front. Psychol.* 2:255. doi: 10.3389/fpsyg.2011.00255
- Indefrey, P., and Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001
- Ito, A., Martin, A. E., and Nieuwland, M. S. (2017a). Why the A/AN prediction effect may be hard to replicate: A rebuttal to DeLong, Urbach, and Kutas (2017). *Lang. Cogn. Neurosci.* 32, 974–983. doi: 10.1080/23273798.2017.1323112
- Ito, A., Martin, A. E., and Nieuwland, M. S. (2017b). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Lang. Cogn. Neurosci.* 32, 954–965. doi: 10.1080/23273798.2016.1242761
- Japee, S., Jordan, J., Licht, J., Lokey, S., Chen, G., Snow, J., et al. (2023). Inability to move one’s face dampens facial expression perception. *Cortex* 169, 35–49. doi: 10.1016/j.cortex.2023.08.014
- Jenson, D., Reilly, K. J., Harkrider, A. W., Thornton, D., and Saltuklaroglu, T. (2018). Trait related sensorimotor deficits in people who stutter: An EEG investigation of μ rhythm dynamics during spontaneous fluency. *Neuroimage Clin.* 19, 690–702. doi: 10.1016/j.nicl.2018.05.026
- Johari, K., and Behroozmand, R. (2021). Neural correlates of speech and limb motor timing deficits revealed by aberrant beta band desynchronization in Parkinson’s disease. *Clin. Neurophysiol.* 132, 2711–2721. doi: 10.1016/j.clinph.2021.06.022
- Knoefler, P. (2019). Predicting (variability of) context effects in language comprehension. *J. Cult. Cogn. Sci.* 3, 141–158. doi: 10.1007/s41809-019-00025-5
- Kuperberg, G. R., and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci.* 31, 32–59. doi: 10.1080/23273798.2015.1102299
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Lonkiewicz, J. R., Rabagliati, H., and Pickering, M. J. (2021). The role of language production in making predictions during comprehension. *Q. J. Exp. Psychol.* 74, 2193–2209. doi: 10.1177/17470218211028438
- León-Cabrera, P., Pagonabarraga, J., Moris, J., Martínez-Horta, S., Marín-Lahoz, J., Horta-Barba, A., et al. (2021). Neural signatures of predictive language processing in Parkinson’s disease with and without mild cognitive impairment. *Cortex* 141, 112–127. doi: 10.1016/j.cortex.2021.03.032
- León-Cabrera, P., Piai, V., Moris, J., and Rodríguez-Fornells, A. (2022). Alpha power decreases associated with prediction in written and spoken sentence comprehension. *Neuropsychologia* 173:108286. doi: 10.1016/j.neuropsychologia.2022.108286
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776
- Lieberman, A. M. (1957). Some results of research on speech perception. *J. Acoust. Soc. Am.* 29, 117–123. doi: 10.1121/1.1908635
- Lind, A., and Hartsuiker, R. J. (2020). Self-monitoring in speech production: Comprehending the conflict between conflict- and comprehension-based accounts. *J. Cogn.* 3:16. doi: 10.5334/joc.118
- Macherey, O., and Carlyon, R. P. (2014). Cochlear implants. *Curr. Biol.* 24, R878–R884. doi: 10.1016/j.cub.2014.06.053
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., and Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 503–535. doi: 10.1037/0278-7393.33.3.503
- Mani, N., and Huetting, F. (2012). Prediction during language processing is a piece of cake-But only for skilled producers. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 843–847. doi: 10.1037/a0029284
- Marrufo-Pérez, M. I., Eustaquio-Martín, A., and Lopez-Poveda, E. A. (2019). Speech predictability can hinder communication in difficult listening conditions. *Cognition* 192:103992. doi: 10.1016/j.cognition.2019.06.004
- Martin, C. D., Branzi, F. M., and Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Sci. Rep.* 8:1079. doi: 10.1038/s41598-018-19499-4
- Massaro, D. W., and Chen, T. H. (2008). The motor theory of speech perception revisited. *Psychon. Bull. Rev.* 15, 453–457. doi: 10.3758/PBR.15.2.453/METRICS
- Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., and Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemp. Issues Commun. Sci. Disord.* 31, 105–122. doi: 10.1044/cicsd_31_s_105
- McCauley, S. M., and Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. *Psychol. Rev.* 126, 1–51. doi: 10.1037/REV0000126
- McMurray, B., Farris-Trimble, A., and Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition* 169, 147–164. doi: 10.1016/j.cognition.2017.08.013
- McQueen, J., and Meyer, A. S. (2019). “Key issues and future directions: Toward a comprehensive cognitive architecture for language use,” in *Human language: From genes and brains to behavior*, ed. P. Hagoort (Cambridge, MA: The MIT Press), 85–96.
- Mersov, A.-M., Jobst, C., Cheyne, D. O., and De Nil, L. (2016). Sensorimotor oscillations prior to speech onset reflect altered motor networks in adults who stutter. *Front. Hum. Neurosci.* 10:443. doi: 10.3389/fnhum.2016.00443
- Mock, J. R., Foundas, A. L., and Golob, E. J. (2016). Cortical activity during cued picture naming predicts individual differences in stuttering frequency. *Clin. Neurophysiol.* 127, 3093–3101. doi: 10.1016/j.clinph.2016.06.005
- Molinario, N., and Monsalve, I. F. (2018). Perceptual facilitation of word recognition through motor activation during sentence comprehension. *Cortex* 108, 144–159. doi: 10.1016/j.cortex.2018.07.001
- Nagels, L., Bastiaanse, R., Başkent, D., and Wagner, A. (2020). Individual differences in lexical access among cochlear implant users. *J. Speech Lang. Hear. Res.* 63, 286–304. doi: 10.1044/2019_JSLHR-19-00192
- Navarrete, E., Del Prato, P., Peressotti, F., and Mahon, B. Z. (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *J. Mem. Lang.* 76, 253–272. doi: 10.1016/j.jml.2014.05.003
- Nicenboim, B., Vasishth, S., and Röslér, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a Bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia* 142:107427. doi: 10.1016/j.neuropsychologia.2020.107427
- Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., et al. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philos. Trans. R. Soc. B Biol. Sci.* 375:20180522. doi: 10.1098/rstb.2018.0522
- Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., et al. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *Elife* 7:e33468. doi: 10.7554/eLife.33468.001
- Nour Eddine, S., Brothers, T., and Kuperberg, G. R. (2022). The N400 in silico: A review of computational models. *Psychol. Learn. Motiv. Adv. Res. Theory* 76, 123–206. doi: 10.1016/BS.PLM.2022.03.005
- Nour Eddine, S., Brothers, T., Wang, L., Spratling, M., and Kuperberg, G. R. (2024). A predictive coding model of the N400. *Cognition* 246:105755. doi: 10.1016/j.cognition.2024.105755
- Nozari, N. (2020). A comprehension- or a production-based monitor? Response to Roelofs (2020). *J. Cogn.* 3, 1–21. doi: 10.5334/joc.102
- Nozari, N., and Novick, J. (2017). Monitoring and control in language production. *Curr. Dir. Psychol. Sci.* 26, 403–410. doi: 10.1177/0963721417702419
- Nuttall, H. E., Kennedy-Higgins, D., Hogan, J., Devlin, J. T., and Adank, P. (2016). The effect of speech distortion on the excitability of articulatory motor cortex. *Neuroimage* 128, 218–226. doi: 10.1016/j.neuroimage.2015.12.038
- Obleser, J., and Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb. Cortex* 20, 633–640. doi: 10.1093/cercor/bhp128
- Okada, K., Matchin, W., and Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. *Psychon. Bull. Rev.* 25, 423–430. doi: 10.3758/s13423-017-1284-x
- Pamplona, M., del, C., Ysunza, P. A., Telich-Tarriba, J., Chávez-Serna, E., Villate-Escobar, P., et al. (2020). Diagnosis and treatment of speech disorders in children with Moebius syndrome. *Int. J. Pediatr. Otorhinolaryngol.* 138:110316. doi: 10.1016/j.ijp.2020.110316
- Peelle, J. E., and Sommers, M. S. (2015). Prediction and constraint in audiovisual speech perception. *Cortex* 68, 169–181. doi: 10.1016/j.cortex.2015.03.006
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47(Suppl. 2), S72–S82. doi: 10.1080/14992020802307404
- Pickering, M. J., and Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychol. Bull.* 144, 1002–1044. doi: 10.1037/bul0000158
- Pickering, M. J., and Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends Cogn. Sci.* 11, 105–110. doi: 10.1016/j.tics.2006.12.002

- Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495
- Pickering, M. J., and Garrod, S. (2014). Neural integration of language production and comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15291–15292. doi: 10.1073/pnas.1417917111
- Poeppl, D., and Monahan, P. J. (2011). Feedforward and feedback in speech perception: Revisiting analysis by synthesis. *Lang. Cogn. Process.* 26, 935–951. doi: 10.1080/01690965.2010.493301
- Poeppl, D., Idsardi, W. J., and Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1071–1086. doi: 10.1098/rstb.2007.2160
- Preisig, B. C., Riecke, L., and Hervais-Adelman, A. (2022). Speech sound categorization: The contribution of non-auditory and auditory cortical regions. *Neuroimage* 258:119375. doi: 10.1016/j.neuroimage.2022.119375
- Radošević, T., Malaia, E. A., and Milković, M. (2022). Predictive processing in sign languages: A systematic review. *Front. Psychol.* 13:805792. doi: 10.3389/fpsyg.2022.805792/BIBTEX
- Roelofs, A. (2020a). On (correctly representing) comprehension-based monitoring in speaking: Rejoinder to Nozari (2020). *J. Cogn.* 3, 1–7. doi: 10.5334/joc.112
- Roelofs, A. (2020b). Self-monitoring in speaking: In defense of a comprehension-based account. *J. Cogn.* 3, 1–13. doi: 10.5334/joc.61
- Rommers, J., Dell, G. S., and Benjamin, A. S. (2020). Word predictability blurs the lines between production and comprehension: Evidence from the production effect in memory. *Cognition* 198:104206. doi: 10.1016/j.cognition.2020.104206
- Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., and Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Lang. Cogn. Neurosci.* 32, 576–589. doi: 10.1080/23273798.2016.1183799
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7295–7300. doi: 10.1073/PNAS.0609419104
- Ryskin, R., and Nieuwland, M. S. (2023). Prediction during language comprehension: What is next? *Trends Cogn. Sci.* 27, 1032–1052. doi: 10.1016/J.TICS.2023.08.003
- Samson, A., and Voyer, B. G. (2012). Two minds, three ways: Dual system and dual process models in consumer psychology. *AMS Rev.* 2, 48–71. doi: 10.1007/S13162-012-0030-9
- Schomers, M. R., and Pulvermüller, F. (2016). Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Front. Hum. Neurosci.* 10:435. doi: 10.3389/FNHUM.2016.00435/BIBTEX
- Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., et al. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2105646118. doi: 10.1073/PNAS.2105646118
- Schwartz, R. G. (2010). *Handbook of child language disorders*, 2nd Edn. London: Psychology Press, doi: 10.4324/9780203837764
- Sessa, P., Schiano Lomoriello, A., Duma, G. M., Mento, G., De Stefani, E., and Ferrari, P. F. (2022). Degenerate pathway for processing smile and other emotional expressions in congenital facial palsy: An hEEG investigation. *Philos. Trans. R. Soc. B* 377:20210190. doi: 10.1098/RSTB.2021.0190
- Sherafati, A., Dwyer, N., Bajracharya, A., Hassanpour, M. S., Eggebrecht, A. T., Firszt, J. B., et al. (2022). Prefrontal cortex supports speech perception in listeners with cochlear implants. *Elife* 11:e75323. doi: 10.7554/ELIFE.75323
- Silbert, L. J., Honey, C. J., Simony, E., Poeppl, D., and Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc. Natl. Acad. Sci. U.S.A.* 111, E4687–E4696. doi: 10.1073/pnas.1323812111
- Skipper, J. I., Devlin, J. T., and Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang.* 164, 77–105. doi: 10.1016/J.BANDL.2016.10.004
- Smith, A., and Weber, C. (2017). How stuttering develops: The multifactorial dynamic pathways theory. *J. Speech Lang. Hear. Res.* 60, 2483–2505. doi: 10.1044/2017_JSLHR-S-16-0343
- Spalek, K., Damian, M. F., and Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the special issue on lexical competition in language production. *Lang. Cogn. Process.* 28, 597–614. doi: 10.1080/01690965.2012.718088
- Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Lang. Linguist. Compass* 9, 311–327. doi: 10.1111/lnc3.12151
- Staub, A., and Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 45, 110–127. doi: 10.1037/xlm0000561
- Stevenson, R. A., Sheffield, S. W., Butera, I. M., Gifford, R. H., and Wallace, M. T. (2017). Multisensory integration in cochlear implant recipients. *Ear Hear.* 38, 521–538. doi: 10.1097/AUD.0000000000000435
- Strelnikov, K., Rouger, J., Barone, P., and Deguine, O. (2009). Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scand. J. Psychol.* 50, 437–444. doi: 10.1111/J.1467-9450.2009.00741.X
- Strijkers, K., and Costa, A. (2016). The cortical dynamics of speaking: Present shortcomings and future avenues. *Lang. Cogn. Neurosci.* 31, 484–503. doi: 10.1080/23273798.2015.1120878
- Tourville, J. A., and Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Lang. Cogn. Process.* 26, 952–981. doi: 10.1080/0169096093498424
- Tuckute, G., Sathe, A., Srikant, S., Taliaferro, M., Wang, M., Schrimpf, M., et al. (2024). Driving and suppressing the human language network using large language models. *Nat. Hum. Behav.* 8, 544–561. doi: 10.1038/s41562-023-01783-7
- Urbach, T. P., DeLong, K. A., Chan, W.-H., and Kutas, M. (2020). An exploratory data analysis of word form prediction during word-by-word reading. *Proc. Natl. Acad. Sci. U.S.A.* 117, 20483–20494. doi: 10.1073/pnas.1922028117
- Van Os, M., Kray, J., and Demberg, V. (2022). Rational speech comprehension: Interaction between predictability, acoustic signal, and noise. *Front. Psychol.* 13:8064. doi: 10.3389/fpsyg.2022.914239/BIBTEX
- Van Petten, C., and Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *Int. J. Psychophysiol.* 83, 176–190. doi: 10.1016/j.jpsycho.2011.09.015
- Vannuscorps, G., Andres, M., Carneiro, S. P., Rombaux, E., and Caramazza, A. (2021). typically efficient lipreading without motor simulation. *J. Cogn. Neurosci.* 33, 611–621. doi: 10.1162/JOCN_A_01666
- Vilain, A., Dole, M., Loevenbruck, H., Pascalis, O., and Schwartz, J. L. (2019). The role of production abilities in the perception of consonant category in infants. *Dev. Sci.* 22:e12830. doi: 10.1111/DESC.12830
- Walenski, M., Europa, E., Caplan, D., and Thompson, C. K. (2019). Neural networks for sentence comprehension and production: An ALE-based meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 40, 2275–2304. doi: 10.1002/hbm.24523
- Wang, L., Hagoort, P., and Jensen, O. (2018). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *J. Cogn. Neurosci.* 30, 432–447. doi: 10.1162/jocn_a_01190
- Winn, M. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends Hear.* 20:2331216516669723. doi: 10.1177/2331216516669723
- Wolpert, D. M. (1997). Computational approaches to motor control. *Trends Cogn. Sci.* 1, 209–216. doi: 10.1016/S1364-6613(97)01070-X
- Wolpert, D. M., and Flanagan, J. R. (2001). Motor prediction. *Curr. Biol.* 11, R729–R732. doi: 10.1016/S0960-9822(01)00432-8
- Zhang, L., Wang, X., Alain, C., and Du, Y. (2023). Successful aging of musicians: Preservation of sensorimotor regions aids audiovisual speech-in-noise perception. *Sci. Adv.* 9:eadg7056. doi: 10.1126/SCIADV.ADG7056