Integration of internal linguistic information with sensory input via weakly entrained oscillations

Rong Ding (Rong.Ding@mpi.nl)
Language and Computation in Neural Systems, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

Sanne Ten Oever (sanne.tenoever@maastrichtuniversity.nl)
Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
Language and Computation in Neural Systems, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

Andrea E. Martin (Andrea.Martin@mpi.nl)
Language and Computation in Neural Systems, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

Abstract

Rhythmic neural activity has been widely observed across cognitive domains, including language. Yet, debate continues on how such activity supports speech processing: entraining to external stimuli for optimised tracking, or driving the generation of internal representations. By introducing graded sensory entrainment to a fixed oscillator model of Ten Oever and Martin (2021), our study examined how it influences phase coding—the timing-based differentiation of word nodes from internal feedback—and how these codes bias ambiguous input interpretation. Simulations show that moderate coupling supports reliable phase coding while preserving sensitivity to unexpected inputs. Our model shows how the brain could coordinate top-down linguistic representations with bottom-up sensory processes during speech processing.

Introduction

Neural oscillations, linked to fluctuations in network excitability, are implicated in speech comprehension (e.g., Meyer, 2018), but their exact role remains debated. Some prior models emphasise entrainment to stimuli for efficient tracking (e.g., Giraud & Poeppel, 2012), while others focus on internally driven predictions (e.g., Ten Oever & Martin, 2021). Despite proposals that both mechanisms interact (Meyer, Sun, & Martin, 2019), they are rarely unified in modelling. Our study integrates them by extending an existing STiMCON model (Ten Oever & Martin, 2021) with graded sensory entrainment via a dynamic systems approach to address this divide.

Method

The STiMCON model of Ten Oever and Martin (2021) implemented fixed oscillations to drive internal representations of employed word nodes. In STiMCON, word nodes are sensitised (i.e., their activity heightened) by top-down predictive feedback from an internal language model so that more predictable words activate earlier as oscillatory activity peaks, thereby optimising speech processing in a top-down manner.

Here, we extended the STiMCON model—hereafter referred to as STiMCON+[1]—by explicitly incorporating sensory entrainment to improve sensitivity to bottom-up input. STiMCON+ comprises three layers: STIMULUS (external input), MAIN (oscillating word nodes), and PREDICTION (language-based feedback). All the other model parameters used here were identical to STiMCON. MAIN layer activation is influenced by both other layers. Sensory entrainment was introduced via Stuart–Landau equations (e.g., Doelling & Assaneo, 2021), with coupling strength K determining the influence of sensory input on internal dynamics. External input modulates the oscillator's velocity, producing a phase shift that in turn adjusts the timing of feedback activation.

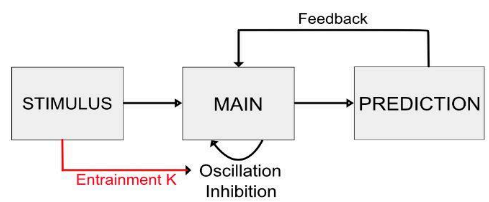

Figure 1. Schema of STiMCON+

Simulations used a five-node language model (*I*, *eat*, *very*, *nice*, and *cake*) with context-driven predictabilities (e.g., after *eat*: *cake* = 0.5, *nice* = 0.3, *very* = 0.2), proportionally to which the PREDICTION layer generates feedback to the MAIN layer. We tested the model's (oscillator frequency = 4 Hz) behaviour under varying K levels using isochronous and non-isochronous 4-Hz input streams for oscillator entraining. Non-isochronous inputs had onsets randomly drawn from a uniform interval around the isochronous timing. The seventh input *eat* engaged the language model to trigger feedback. We assessed three aspects under varying coupling strengths. First, we evaluated whether top-down feedback activation scaled with word predictability. Second, we measured temporal phase codes—the differences in feedback activation timing among word nodes. Third, we tested ambiguous input categorisation by varying both the degrees of ambiguity (between *cake* and *nice*) and onset delays (relative to the offset of a prior input). In this case, an ambiguous

[1] Code can be found: https://github.com/Rong-Ding/Coupled_STiMCON

input was presented after the seven-input stream (for oscillator entrainment). These three measures allow for evaluating the strength of internal representations and their influence on bottom-up processing.

Result

The presentation of both isochronous and non-isochronous stimuli yielded similar trends. Increased sensory coupling amplified feedback activation for more predictable word nodes (e.g., *cake*) and suppressed it for less predictable ones (e.g., *very*) (Figure 2A). Temporal phase codes between word nodes *cake* and *nice* became more stable (i.e., overall smaller temporal differences) under stronger coupling (Figure 2B). Ambiguous input categorisation became more deterministic and strongly dependent on input timing under increased coupling (Figure 2C). Notably, these effects converged more rapidly with isochronous inputs, likely because regular stimuli produce less variable phase shifts.

Discussion and Conclusion

Our simulations indicate that increasing the coupling strength between the internal oscillator and external sensory inputs enhances the reliability of feedback activation for highly predictable word nodes while reducing feedback for less predictable ones. Stronger coupling produces robust phase shifts—which stabilise temporal phase codes—but at the cost of flexibility, potentially leading the model to miss less expected inputs. Conversely, moderate sensory entrainment allows the model to balance top-down predictions with bottom-up variability, enabling it to bias ambiguous input categorisation without becoming overly deterministic. Together, these findings support the argument that moderate coupling is optimal for effective speech tracking, as it preserves reliable phase coding while still accommodating variable and unpredictable input.

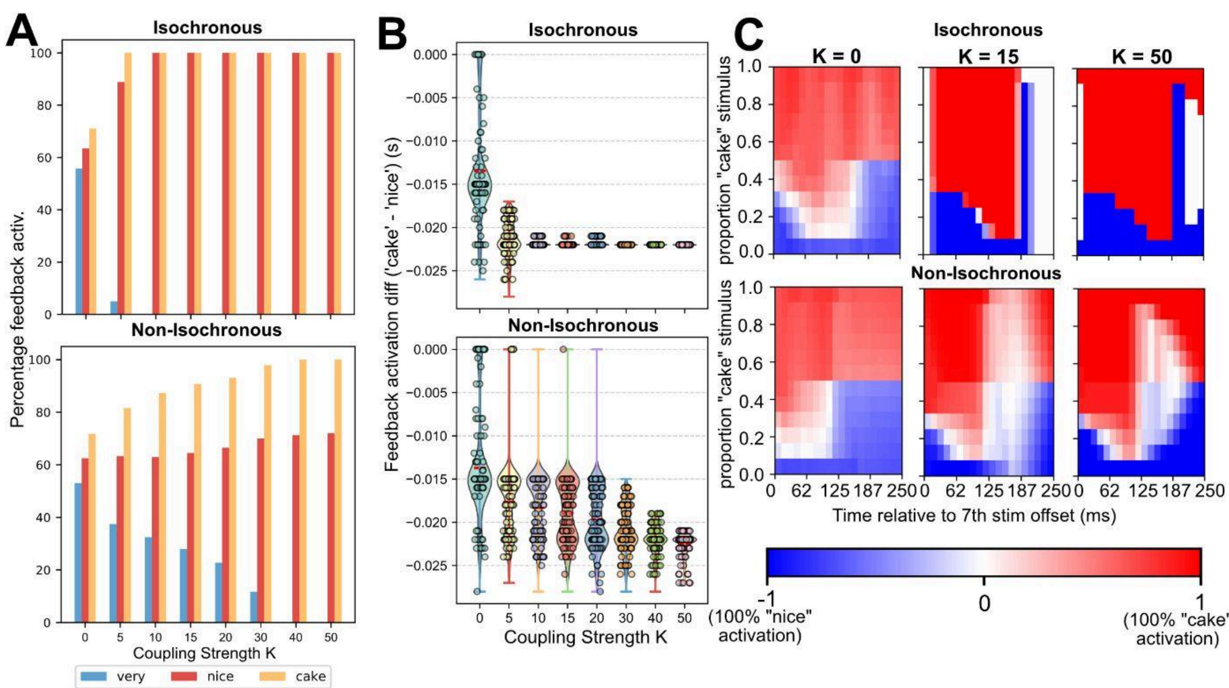

Figure 2. Model performance with 4-Hz isochronous and non-isochronous stimuli. (A) Activation percentages of predicted word nodes across coupling strengths (1000 iterations). (B) Feedback activation time differences between “cake” and “nice” (100 random datapoints plotted per K). (C) Proportion of first-activated nodes across onset delay and stimulus proportion (red: “cake” and blue “nice”; shade: the proportion of a word node as first active). Y-axis: Proportion of *cake* in the ambiguous input as a mix/morph of the stimuli *cake* and *nice*.

Acknowledgments

We gratefully thank all the anonymous reviewers for their useful suggestions.

AEM was supported by an Independent Max Planck Research Group and a Lise Meitner Research Group "Language and Computation in Neural Systems" from the Max Planck Society, by the Netherlands Organization for Scientific Research (NWO; grant 016.Vidi.188.029), and by Big Question 5 (to Prof. dr. Roshan Cools & Dr. Andrea E. Martin) of the Language in Interaction Consortium funded by an NWO Gravitation Grant (024.001.006) awarded to Prof. dr. Peter Hagoort.

SO was supported by the European Research Council (ERC-2023-STG; 101116685) and by the Dutch Research Council (NWO; Starter Grant awarded to starting assistant professors).

References

Doelling, K. B., & Assaneo, M. F. (2021). Neural oscillations are a start toward understanding brain activity rather than the end. *PLoS biology, 19*(5), e3001234. https://doi.org/10.1371/journal.pbio.3001234

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience, 15*(4), 511–517. https://doi.org/10.1038/nn.3063

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience, 48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Meyer, L., Sun, Y., & Martin, A. E. (2020). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience, 35*(9), 1089–1099. https://doi.org/10.1080/23273798.2019.1693050

Ten Oever, S., & Martin, A. E. (2021). An oscillating computational model can track pseudo-rhythmic speech by using linguistic predictions. *eLife, 10*, e68066. https://doi.org/10.7554/eLife.6806