An Integrated Neurogenetic and Linguistic Framework for Understanding Sentence Processing

Chaoyue Lu*

School of Psychology and Neuroscience, University of Bristol,
Beacon House, Queens Rd, Bristol BS8 1QU, UK

*Corresponding author. Email: chaoyue.lu.331@gmail.com

Abstract. The sentence processing mechanism is a complex cerebral function and involves multiple factors at the genetic and experiential levels. This article presents a coherent model (namely, the FOXP2 and ROBO1 model) that integrates the genetic architecture, comprising FOXP2 and ROBO1 variants, and in analogy, the linguistic experience that accounts for the variability in neural mechanisms of sentence comprehension. We synthesize evidence that (i) genetic variants influence the structural and connectional substrate of language-relevant regions (e.g., posterior inferior frontal gyrus and planum temporale), and (ii) linguistic experience shapes the functional recruitment and coupling of cortical networks during comprehension. This article, gathering the empirical evidence from neurogenetic and cross-linguistic neurosciences, claims that sentence comprehension takes place in the network created by the exiting interplay of genetic predisposition and the influence of environment, which leads to a more comprehensive model intended to account for the differences in language processing.

Keywords: Sentence Processing, Broca's Area (pIFG), FOXP2, ROBO1, Cross-Linguistic

1 Introduction

Sentence processing is a cognitive mechanism that structures word sequences syntactically and semantically to derive meaning representations, attracting computational and memory resources for the maintenance and prediction of linguistic structures in comprehension[1][2][3]. A central issue in neurolinguistics concerns the functional role of left inferior frontal regions—particularly the posterior inferior frontal gyrus (pIFG, commonly subsumed under "Broca's area")—in such processing. Two dominant accounts persist: (i) a syntax-specific account, assigning to pIFG a core role in hierarchical, syntax-specific computations[4]; and (ii) a domain-general account, which attributes pIFG activation to increased working-memory and cognitive control demands that are often concomitant with complex syntactic processing[5][6].

Modern science increasingly discovers the significance of genetic composition variations[7][8] and the experience of language[9]is an important factors that affect neural activity during the comprehension of sentences. Such variability is not purely statistical

T. Chamaratana et al. (eds.), *Proceedings of the 2026 12th International Conference on Digital Humanities and Frontiers in Social Sciences (DHFSS 2026)*, Advances in Social Science, Education and Humanities Research 1031,
https://doi.org/10.2991/978-2-38476-611-6_5

nonsense but reflects the way sentences are formed. Neurocomputational individual differences across participants usually reflect the mechanistic variety rather than the neural variability linked to task performance, and it has been detected that different individuals may have different computational strategies, although their task outcomes are identical. Among these, by far the biggest divergence is in the heterogeneity seen in both neural and behavioral. They discovered that the mentioned variability was attributable to the different mixes of three computational units, all of which were predicted to cause the same behavioral correspondence. These computational techniques, which have a specified and widely substantiated set of neural and behavioral essays, may use a different motor command.

Genetic factors like FOXP2 and ROBO1 affect neural tissues that are involved in sentence comprehension. In particular, the FOXP2 mutations disrupt frontostriatal and frontocerebellar circuits controlling speech motor control, and the particular affected brain parts are Broca's area, caudate nucleus, and cerebellum[10]. However, common variations may have minimal impact on the brain structure with respect to their allele frequency[11]. ROBO1 pathways include neuropilin signaling, which promotes the development of the cortical circuit.

In addition to the contribution of genes to neural circuits, bilingual life experience is a vital determinant of inter-individual differences in sentence processing mechanisms. The left inferior frontal cortex (LIFC) in early bilinguals is activated at a level much higher than the equivalent for monolinguals despite their performance. This increased participation level extends to more superior and middle temporal regions, which is evidence of the stronger cognitive workload comparatively both in Chinese and English within the fundamentally shared neural circuitry[12][13][14]. This article examines how hereditary and environmental factors influence sentence processing, specifically genetic factors such as FOXP2 and ROBO1, and experiential factors including multilingual learning.

Despite growing recognition of these factors, current research mainly examines them in isolation. The subject of genetic studies tends to be structural variations, yet it is rarely discussed regarding their connection to processing dynamics in real time. Cross-language studies have shown the activation of variations but have paid insufficient attention to the genetic composition of the driving factors of variations. Through this fragmented line of thinking, a totally empty space is created. We have yet to draw a picture of how innate factors, like FOXP2 and ROBO1 genetic variations, and acquired factors, like native language typology and second language learning, interact with each other to influence the brain's sentence processing. Understanding sentence processing requires moving beyond isolated factor analysis toward a model that captures the interaction between innate (genetic endowment) and acquired (linguistic experience).

The article aims to establish an all-encompassing framework to reveal the neural process of sentence comprehension through the lens of individual differences. Particularly, the goal is to explain the operational structure and roles for important genes, specifically in FOXP2 and ROBO1, for the modulation of language-relevant brain areas like the posterior inferior frontal gyrus (pIFG) and planum temporale, and for the influence on the phonological memory systems that serve syntactic processing. The purpose of this article is also to clarify how varying linguistic environments, for example in

terms of native language typology and bilingual exposure, add to the differential patterns of neural engagement, especially in China and Germany, w where the syntactic categories are completely different. More crucially, the current work attempts to emphasize the way in which genetic and environmental concerns influence the development of the sentence processing systems. In other words, the aspect being studied here is how genetic variants FOXP2 and ROBO1, which are responsible for the elevation of language-defining areas such as the pIFG and planum temporale, and human communications, can combine with linguistic experience (native language typology and bilingual exposure) to obtain unequal patterns of neural activation and processing capacities during sentence comprehension.

2 Influences of Hereditary and Environmental Factors on Sentence Processing

2.1 Hereditary Factors

The range of the effect that the genetic variants have on the different brain structures that are important for language processing varies tremendously based on the frequency and the degree of effect of the variants in the overall population. Genetic studies have found that FOXP2, the gene that serves a critical role in language formation, causes structural alterations in the planum temporale and the pIFG[11], thereby leading to an increase in syntactic processing efficiency problems.

Genomic studies have indicated that the FOXP2 gene, which is fundamental for language development, causes alterations in the structures of pIFG and planum temporale[11], thereby regulating the efficiency of syntactic processes. The voxel-based morphometry (VBM) and volumetric techniques used by Hoogman and co-authors (2014) to examine the effect of the common variants of FOXP2 on neuroanatomy in over one thousand individuals from the general population[11]. There was found no such relation of the presence of single nucleotide polymorphisms to neuroanatomy in the general population. Accordingly, this gene may not pose nearly as much risk to brain structure as is popularly believed, because of the fact that its disruptive alleles are quite rare. Alternatively, however, another situation is that the effects of common variations are obvious, but standard volume measurement techniques fail to capture the ancient low mechanisms[11].

The fMRI study of the KE family that was conducted through a comparison method of a hidden and a saturation temporal verb-generation and word repetition task is an example. Liégeois et al. (2003) study discovered that there was a reduced activation of the broca's area and the putamen with a wider, posterior activation profile for the affected members[15]. Therefore, the FOXP2 dysfunction alters the development and function of the corticobasal ganglia circuit that supports speech and language processing.

Similarly, ROBO1, which is related to language development, also acts on the phonological working memory in a similar way by increasing axon connections[16], ultimately affecting sentence deduction under high memory load. This article adopted a

family-based association design in 538 twin/sibling families, and the ROBO1 genetic variant was tested via 144 ROBO1 tag SNPs, and the phonological memory was indexed with a standardized nonword-repetition composite and digit span. It discovered multiple ROBO1 variations that affect non-word repetition, and two single nucleotide polymorphisms have similar effects; however, this effect on reading/spelling and working memory is distinct and not strong, which suggests that ROBO1 variation is associated with the phonological buffer subordinate to mental load that can limit sentence comprehension[16].

Sun et al. (2017) genotyped the two ROBO1 variants (rs4535189/rs6803202) in 115 children and combined diffusion MRI with structural MRI to estimate the midline corpus callosum microstructure alongside standardized reading tests[17]. They found that genotypes have an impact on word list reading and corpus callosum diffuser, mediate from axial diffuser, carrying the Robo1-reading association in the genus. They established that the ROBO1 variation is a pathway from white matter to behavior, and that it affects the interhemispheric frontal by the access to regions that support phonological operations, which are used for paragraph level operations, under a high memory load, that it is under negative influence.

Therefore, these data confirm the replication of the pathway from gene to white matter to behavior, where the variation of ROBO1 determines the interhemispheric frontal connection that is involved in phonological operations that are pertinent to the paragraph level under a high memory load. Therefore, these observations reveal that the language-relevant gene variants affect the neuronal architecture in more tangible structural variations of specific brain regions, with the behavioral outcomes most dramatically affected in more severe variant carriers.

2.2 Environmental Factors

Through languages, cross-linguistic research indicates that the native language typology and the differential sentence processing may be a result of adaptive brain processing, which is influenced by the cumulative linguistic experience. For instance, the fMRI study[18] demonstrated that native Mandarin speakers passively listen to different words, and both the left middle/anterior temporal cortex and bilateral STG are engaged during sentence comprehension. This is consistent with the fact that Chinese is more dependent on the temporal lob[19][20]. During this process, linguistic learning experiences can influence the way the brain adapts.

Furthermore, imaging techniques have demonstrated language differences not only in brain activity patterns but also in how language networks are functionally interconnected. To illustrate, when reading English, Chinese–English bilinguals displayed more significant connectivity between the areas involved in visuo-orthographic, the left postcentral gyrus, whereas when reading Chinese, the bilateral central gyrus was observed; essentially, this increased use of the cognitive system is consistent with recent data that supports the concept that bilingualism can enhance remote processing such as integration and organization of information through network pathways[21][22][23]. We compared English pseudoword rhyming with Chinese word rhyming in a matched group of

native and second-language bilinguals using task fMRI and psychophysiological interaction. There was a spatial shift in the seed-to-sensorimotor coupling by the script. For proficiency, rMOG-left postcentral connectivities were have suggested both the script-specific and proficiency-dependency in the reading network.

Even at the level of L2 activation, there are differences due to the varying language distances. An example of this is that Korean speakers exhibited more significant activation of syntactic brain regions when processing English, whereas Chinese speakers showed increased activation during Japanese. The

fMRI was used in which Chinese and Korean participants comprehended the sentences in L1 versus two other second languages (English/Japanese), and the fMRI results of the sentence comprehension were significantly greater in LIFG and bilateral STG for the Korean than the other two languages, whereas there was an anterior STG for Chinese, yet itewas significantly lower. The distance between the languages and how well the languages are known to the individual would determine to a certain degree the usage of the brain compute resource for language processing are other variables that affect the brain[24][25][26].

Leaving the genetic and environmental shared factors which contribute towards the determination and modification of sentence processing mainly on the dynamic inherited successive neural architecture and the person's linguistic experience being the guide. This comprehensive picture demonstrates that genetic predisposition and environmental exposure are not separate but constantly interact throughout the entire developmental process.

Through the genetic aspect, the variants of language genes like FOXP2 and ROBO1 construct the connectivity scaffolding essential to the language, encompassing the lateral prefrontal cortex, planum temporale, and corpus callosum. This is a brain anatomy-structure difference the fund produces individual variability associated with syntactic and semantic processing. Nevertheless, the functionality of these resources is critically discerned by language experiences of the individual. Translations of language types in the cross-cultural studies rarely are similar and that can be demonstrated to be different in various languages. For example, Chinese speakers show distinct patterns of semantic engagement in the temporal lobe, while users of structured languages such as Arabic exhibit differentiated involvement of the pIFG in syntactic processing[27][28][29].

Critically, the role of these genetics and environmental factors is much more interactive than most people expect. During childhood, the same neural substrate can achieve achievement orientation through different types of language input. And that clears out the runners if the origin of genetically similar populations might be revealed that they have strategies because their analyzed languages are widely different, and genetic variants might or might not be multiple according to their native languages. Therefore, it is precisely these complicated neurocognitive abilities, including both cerebral-biological and learned, which are limited by their nature as well as exposure to language in the environment. The balance between the two is the imbalance of genetic susceptibility, which is shaped to the demands of language in the environment.

3 Conclusion

Research shows that interindividual differences in genetics and linguistics, or the genetic and linguistic background of different people around the world, are the main reasons that contribute to different brain responses to sentences. Results of genetic studies show that structural and functional alternations related to FOXP2 and ROBO1 can affect the efficiency of syntax and phonology. Language typology shapes both the patterns of language-relevant brain networks and the connections between them iare found through cross-linguistic research.

This major concern exemplified the necessity for the model to reach beyond generalizations to reflect stable, individual-contextual characteristics. This article helps to strengthen what we call an interaction-based view, integrating genetic and language facts, which contributes to the understanding of how they work together to achieve a bit of mastery in the sentence processing area by the local brain. The integration of such an approach is not only crucial for language processing models but also has potential to influence personalized technologies for language learning and rehabilitation.

The aim of future research is to test comprehensive models integrating genetic neuroimaging methods and enable the use of language and behavioral features in multimodal programs[30]. In this respect, connecting genotype data (e.g., FOXP2, ROBO1) with functional connections during real-life sentence comprehension would uncover the extent to which genotypes influence neural processing strategies in cross-language comparisons. The long-term studies that follow learners who are uniquely exposed to different grammatical systems help to define the relationship between the genetic predispositions and the experience which rewire the neural language networks over time. The improvement in computational modeling and imaging techniques (for example, for instance, dynamic causal modeling) would be necessary as they help us understand the ways neural circuits constrain and learning practices modify the elements of phonology, semantics, and syntax. Extending this research to educational and clinical populations may clarify the roles of genes and language in the classic scenario of gene-language interactions, which provides new opportunities for the development of personalized therapies and educational programs tailored to the individual neurogenetic profiles.

Reference

1. Linzen, T., & Baroni, M. (2021). Syntactic Structure from Deep Learning. Annual Review of Linguistics, 7(1), 195–212. https://doi.org/10.1146/annurev-linguistics-032020-051035
2. Ness, T., Langlois, V. J., Kim, A. E., & Novick, J. M. (2023). The State of Cognitive Control in Language Processing. Perspectives on Psychological Science. https://doi.org/10.1177/17456916231197122
3. Sun, K., & Wang, R. (2025). Computational Sentence-Level Metrics of Reading Speed and Its Ramifications for Sentence Comprehension. Cognitive Science, 49(7). https://doi.org/10.1111/cogs.70092
4. Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. Behavioral and Brain Sciences, 23(1), 1–21. https://doi.org/10.1017/s0140525x00002399

5. Baddeley, A. (2003). Working memory and language: an overview. Journal of Communication Disorders, 36(3), 189–208. https://doi.org/10.1016/s0021-9924(03)00019-4
6. Rogalsky, C., & Hickok, G. (2011). The Role of Broca's Area in Sentence Comprehension. Journal of Cognitive Neuroscience, 23(7), 1664–1680. https://doi.org/10.1162/jocn.2010.21530
7. Kong, X.-Z., Tzourio-Mazoyer, N., Joliot, M., Fedorenko, E., Liu, J., Fisher, S. E., & Francks, C. (2020). Gene Expression Correlates of the Cortical Network Underlying Sentence Processing. Neurobiology of Language, 1(1), 77–103. https://doi.org/10.1162/nol_a_00004
8. Wen, H., Wang, D., & Bi, Y. (2024). Processing Language Partly Shares Neural Genetic Basis with Processing Tools and Body Parts. ENeuro, 11(8), ENEURO.0138-24.2024. https://doi.org/10.1523/eneuro.0138-24.2024
9. Román, P., & Gómez-Gómez, I. (2022). Changes in Native Sentence Processing Related to Bilingualism: A Systematic Review and Meta-Analysis. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.757023
10. Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. Nature Reviews Neuroscience, 6(2), 131–138. https://doi.org/10.1038/nrn1605
11. Hoogman, M., Guadalupe, T., Zwiers, M. P., Klarenbeek, P., Francks, C., & Fisher, S. E. (2014). Assessing the effects of common variation in the FOXP2 gene on human brain structure. Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00473
12. DeLuca, V., Toms Voits, Ni, J., Carter, F., Rahman, F., Mazaheri, A., Krott, A., & Katrien Segaert. (2024). Mapping individual aspects of bilingual experience to adaptations in brain structure. Cerebral Cortex, 34(2). https://doi.org/10.1093/cercor/bhae029
13. Quiñones, I., Gisbert-Muñoz, S., Amoruso, L., Manso-Ortega, L., Mori, U., Bermudez, G., Robles, S. G., Pomposo, I., & Carreiras, M. (2024). Unveiling the neuroplastic capacity of the bilingual brain: insights from healthy and pathological individuals. Brain Structure and Function, 229(9), 2187–2205. https://doi.org/10.1007/s00429-024-02846-9
14. Sheng, Y., Yang, S., Rao, J., Zhang, Q., Li, J., Wang Dian-jian, & Zheng, W. (2023). Age of Bilingual Onset Shapes the Dynamics of Functional Connectivity and Laterality in the Resting-State. Brain Sciences, 13(9), 1231–1231. https://doi.org/10.3390/brainsci13091231
15. Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. Nature Neuroscience, 6(11), 1230–1237. https://doi.org/10.1038/nn1138
16. Bates, T. C., Luciano, M., Medland, S. E., Montgomery, G. W., Wright, M. J., & Martin, N. G. (2010). Genetic Variance in a Component of the Language Acquisition Device: ROBO1 Polymorphisms Associated with Phonological Buffer Deficits. Behavior Genetics, 41(1), 50–57. https://doi.org/10.1007/s10519-010-9402-9
17. Sun, X., Song, S., Liang, X., Xie, Y., Zhao, C., Zhang, Y., Shu, H., & Gong, G. (2017). ROBO1polymorphisms, callosal connectivity, and reading skills. Human Brain Mapping, 38(5), 2616–2626. https://doi.org/10.1002/hbm.23546
18. Liu, H., & Chen, S. A. (2020). The Neural Correlates of Spoken Sentence Comprehension in the Chinese Language: An fMRI Study. Psychology Research and Behavior Management, Volume 13, 641–652. https://doi.org/10.2147/prbm.s251935
19. Guo, W., Geng, S., Cao, M., & Feng, J. (2022). The Brain Connectome for Chinese Reading. Neuroscience Bulletin, 38(9), 1097–1113. https://doi.org/10.1007/s12264-022-00864-3
20. Sun, Z., Shi, Y., Guo, P., Yang, Y., & Zhu, Z. (2021). Independent syntactic representation identified in left front-temporal cortex during Chinese sentence comprehension. Brain and Language, 214, 104907–104907. https://doi.org/10.1016/j.bandl.2021.104907

21. Gracia-Tabuenca, Z., Barbeau, E. B., Kousaie, S., Chen, J.-K., Chai, X., & Klein, D. (2024). Enhanced efficiency in the bilingual brain through the inter-hemispheric cortico-cerebellar pathway in early second language acquisition. Communications Biology, 7(1). https://doi.org/10.1038/s42003-024-06965-1
22. Fedeli, D., Del Maschio, N., Sulpizio, S., Rothman, J., & Abutalebi, J. (2021). The bilingual structural connectome: Dual-language experiential factors modulate distinct cerebral networks. Brain and Language, 220, 104978. https://doi.org/10.1016/j.bandl.2021.104978
23. Stasenko, A., Kaestner, E., Arienzo, D., Schadler, A., Reyes, A., Shih, J. J., Helm, J. L., Połczyńska, M., & McDonald, C. R. (2023). Bilingualism and Structural Network Organization in Temporal Lobe Epilepsy. Neurology, 100(18), e1887–e1899. https://doi.org/10.1212/wnl.0000000000207087
24. Gallo, F., Andriy Myachykov, Nelyubina, M., Yury Shtyrov, Kubiak, J., Liliia Terekhina, & Jubin Abutalebi. (2023). Linguistic distance dynamically modulates the effects of bilingualism on executive performance in aging. Bilingualism: Language and Cognition, 1–10. https://doi.org/10.1017/s1366728923000743
25. Kepinska, O., Caballero, J., Oliver, M., Marks, R. A., Haft, S. L., Zekelman, L., Kovelman, I., Uchikoshi, Y., & Hoeft, F. (2023). Language combinations of multilinguals are reflected in their first-language knowledge and processing. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-27952-2
26. Leivada, E., Kelly-Iturriaga, L., Masullo, C., Westergaard, M., Rothman, J., Leivada, E., Kelly-Iturriaga, L., Masullo, C., Westergaard, M., & Rothman, J. (2025). The unpredictable role of language distance in bilingual cognition: A systematic review from brain to behavior. Bilingualism Language and Cognition, 1–14. https://doi.org/10.1017/s1366728925100746
27. Wei, X., Adamson, H., Schwendemann, M., Goucha, T., Friederici, A. D., & Anwander, A. (2023). Native language differences in the structural connectome of the human brain. NeuroImage, 270, 119955. https://doi.org/10.1016/j.neuroimage.2023.119955
28. Zhang, Y., Taft, M., Tang, J., & Li, L. (2024). Neural correlates of semantic-driven syntactic parsing in sentence comprehension. NeuroImage, 289, 120543. https://doi.org/10.1016/j.neuroimage.2024.120543
29. Zhu, Y., Xu, M., Lu, J., Hu, J., Kwok, V. P. Y., Zhou, Y., Yuan, D., Wu, B., Zhang, J., Wu, J., & Tan, L. H. (2022). Distinct spatiotemporal patterns of syntactic and semantic processing in human inferior frontal gyrus. Nature Human Behaviour, 6(8), 1104–1111. https://doi.org/10.1038/s41562-022-01334-6
30. Chung, J., Kim, S., Won, J. H., & Park, H. (2024). Integrating Multimodal Neuroimaging and Genetics: A Structurally-Linked Sparse Canonical Correlation Analysis Approach. IEEE Journal of Translational Engineering in Health and Medicine, 12, 659–667. https://doi.org/10.1109/jtehm.2024.3463720