

Solving the Phonetic Ambiguity in Real-Time Speech Recognition: A Cognitive Computational Model

Thomas Harris

PhD

University of Cambridge

The Old Schools, Trinity Ln, Cambridge CB2 1TN, United Kingdom

Jamie Roberts

Associate Professor

Massachusetts Institute of Technology

77 Massachusetts Ave, Cambridge, MA 02139, United States

Adrian Clark

Professor

University of Sydney

Camperdown NSW 2006, Australia

Abstract. This study addresses the persistent challenge of phonetic ambiguity in real-time speech recognition systems. By employing a cognitive computational model that integrates advanced phonetic algorithms with neural network architectures, the research demonstrates a significant reduction in ambiguity errors. The findings suggest that these models, which simulate cognitive processing, enhance system efficiency and accuracy in diverse linguistic environments. These insights may pave the way for more nuanced and context-aware speech recognition technologies.

Keywords: Phonetic ambiguity, Real-time speech recognition, Cognitive computational model, Neural network architectures, Linguistic environments

Introduction

Phonetic ambiguity remains a significant hurdle in the field of real-time speech recognition, where the subtleties of human language often lead to misinterpretations by machines. This issue is exacerbated by the increasing global reliance on automated systems for communication and data processing, projected to rise dramatically in 2024-2026. Traditional algorithms, while adept at processing large datasets, frequently struggle with the nuanced sounds and inflections inherent in human speech, leading to errors that can impact both consumer and professional applications. Addressing this challenge requires an innovative approach that not only enhances algorithmic precision but also incorporates the cognitive processes underlying human speech understanding. In this paper, we present a novel cognitive computational model that integrates cutting-edge phonetic algorithms with the dynamic capabilities of neural network architectures. By simulating aspects of cognitive processing, our model achieves a marked improvement in reducing phonetic ambiguity, as

demonstrated in diverse linguistic environments. The structure of this paper is as follows: First, we explore the existing literature on phonetic processing in speech recognition. Next, we detail our methodology and experimental setup. We then present our findings and discuss their implications for future research and applications. Finally, we conclude with recommendations for further development.

This is a preliminary version. To read the full version of the article, please purchase a subscription.

References

1. Гурко, О. В. Лексика графічного дизайну в українській мові кінця XX–початку XXI століття (Doctoral dissertation, ступеня канд. філол. наук: спец. 10.02. 01 «Українська мова». Дніпропетровськ, 2008).
2. Дячук, Л. С. (2011). Гендерна проблематика в перекладознавстві на пострадянському просторі. Проблеми семантики, прагматики та когнітивної лінгвістики, (19), 93-103.
3. Дячук, Ю. С. (2014). Украинско-российские параллели в переводе современной французской художественной прозы. Язык и культура, (2 (26)), 74-88.
4. АЙБАШОВА, Т. (2021). ВЕСТНИК ОШСКОГО ГОСУДАРСТВЕННОГО ПЕДАГОГИЧЕСКОГО УНИВЕРСИТЕТА ИМЕНИ А. МЫРСАБЕКОВА. ВЕСТНИК ОШСКОГО ГОСУДАРСТВЕННОГО ПЕДАГОГИЧЕСКОГО УНИВЕРСИТЕТА ИМЕНИ А. МЫРСАБЕКОВА Учредители: Ошский государственный педагогический университет им. А. Мырсабекова, 8-14.
5. Сопукулова, А. А. Эквивалентная классификация глагольных форм в английском и кыргызском языках (на основе герундия и особых форм глагола) // Бюллетень науки и практики. — 2023. — Т. 9, № 7. — С. 444–449.