SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 11, No. 3.1, (2025), pp. 1245-1256
Open Access. Online & Print

DOI: 10.5281/zenodo.11325116

NEUROCOGNITIVE SIGNATURES OF LEXICAL-CLASS PERCEPTION IN NOISE AMONG L2 LISTENERS

Awad Alshehri[1*]

[1]Imam Mohammad Ibn Saud Islamic University (IMSIU), CLT, Saudi Arabia.

Received: 28/08/2025
Accepted: 17/09/2025

Corresponding Author: Awad Alshehri
(email@somewhere.com)

ABSTRACT

This study investigated the neurocognitive foundations of lexical category resilience in second language (L2) speech perception under adverse listening conditions. Forty Saudi Arabic learners of English at beginner and advanced proficiency levels completed word-recognition and sentence-recognition tasks in speech-shaped noise while behavioral accuracy and EEG measures were recorded. Mixed-effects models revealed a consistent hierarchy of perceptual resilience adjectives ranked highest, followed by nouns, adverbs, and verbs across proficiency levels, with stops and affricates providing stronger perceptual anchors than approximants and trills. Neurophysiologically, content words elicited larger PMN and P300 amplitudes, indicating more robust phonological mapping and attentional resource allocation. Advanced learners showed stronger N100 and PMN responses, reflecting earlier and more automatic processing, whereas beginner learners relied on later attention-mediated components (P300, N400). Time–frequency analyses identified theta and gamma synchronization as neural markers of successful perception, especially for high-resilience categories. Working memory and phonological awareness correlated strongly with both behavioral and neural measures, particularly for low-resilience categories and less proficient learners. These findings integrate phonetic, linguistic, and cognitive factors to provide a comprehensive account of how lexical category, phonetic salience, and proficiency interact in L2 speech perception under noise, offering implications for L2 pedagogy, automatic speech recognition systems, and neurocognitively informed language training. Nonetheless, the focus on Saudi Arabic learners and the lack of control for certain psycholinguistic variables (e.g., imageability, concreteness) limit generalizability, highlighting the need for studies with diverse populations and more comprehensive stimuli measures.

KEYWORDS: Attention and Working Memory, Auditory Processing, Electroencephalography (EEG), Lexical Categories, Lexical Processing, Neurolinguistics, Phonetic Resilience, Second Language Acquisition, Speech-in-Noise, Speech Perception.

Copyright: © 2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License. (https://cre-ativecommons.org/licenses/by/4.0/).

1. INTRODUCTION

Speech perception is a highly dynamic neurocognitive process in which listeners convert rapidly unfolding acoustic signals into meaningful linguistic units, often under suboptimal listening conditions. When background noise disrupts auditory input, the brain recruits compensatory mechanisms to reconstruct degraded speech signals. This challenge is amplified for second language (L2) learners, whose phonological representations and processing routines differ from those of native speakers. Although the detrimental effects of noise on speech intelligibility are well documented (Miller & Nicely, 1955; Cutler, 2012), the extent to which different lexical categories withstand noise interference remains poorly understood, particularly in L2 contexts.

Processing speech in adverse conditions underpins effective communication and carries significant implications for language learning, academic success, and technological innovation. Neuroscientific research increasingly highlights that speech perception engages distributed neural systems encompassing both auditory cortices and higher-order language regions (Wong et al., 2009; Golestani et al., 2013). Yet, how these networks respond to varying grammatical categories in noisy environments especially for L2 learners whose cognitive resources are already taxed by non-native processing remains an open question.

Lexical categories such as nouns, verbs, and adjectives differ markedly in their acoustic salience, semantic weight, and syntactic roles. Content words typically bear greater prosodic prominence and semantic load, potentially enhancing their robustness under masking conditions (Field, 2004; Morrill et al., 2015). In contrast, function words often occur in reduced, unstressed forms and depend heavily on syntactic context, making them particularly susceptible to noise interference (Kerlin et al., 2010). Exploring how these linguistic differences interact with neurocognitive mechanisms can clarify the architecture of both first- and second-language speech processing.

The present study adopts a multidimensional approach that integrates phonetic analysis, psycholinguistic measures, and electroencephalographic (EEG) data to examine how lexical categories differ in their resilience to noise. Rather than treating speech perception as a uniform process, we hypothesize that distinct lexical classes engage specialized neural mechanisms and exhibit variable perceptual robustness based on their phonological and grammatical characteristics. By recording real-time EEG responses during speech perception tasks, we aim to reveal the neural dynamics underlying successful and unsuccessful reconstruction of degraded speech across lexical categories.

To capture developmental effects, we compare beginner and advanced Saudi learners of English. This population provides a compelling test case because Arabic and English differ substantially in phonological inventory, syllable structure, and prosodic organization. Examining L2 learners at contrasting proficiency levels enables us to trace how perceptual strategies and neural responses evolve with increasing linguistic competence under adverse listening conditions.

1.1. Research Questions and Theoretical Framework

This study addresses three interrelated research questions that collectively explore the neurocognitive mechanisms of speech perception under adverse listening conditions

1. How does neural processing and the reconstruction of speech signals vary across lexical categories (nouns, verbs, adjectives, function words) when speech is presented in background noise?
2. Which neurophysiological markers differentiate successful from unsuccessful sound reconstruction in L2 learners, and how do these markers change with increasing proficiency?
3. How do phonetic features (e.g., manner of articulation, voicing) interact with lexical category to influence speech perception accuracy and its underlying neural correlates?

To address these questions, we draw on three complementary theoretical perspectives. The Distributed Cohort Model (Gaskell & Marslen-Wilson, 2002) highlights the parallel activation of lexical candidates during speech perception, showing how contextual cues rapidly constrain word recognition in real time. Adaptive Resonance Theory (Grossberg, 2003) explains how perception arises from a resonance between sensory input and stored representations, offering insight into how listeners reconcile degraded acoustic signals with existing lexical knowledge. Finally, the Automatic Selective Perception Model (Strange, 2011) emphasizes the role of selective attention in L2 speech perception, especially under conditions of high perceptual load. Taken together, these models illuminate how bottom-up acoustic cues and top-down linguistic knowledge interact dynamically during real-time speech reconstruction.

Guided by this framework and prior empirical findings, we propose three hypotheses. First, lexical category effects: content words particularly nouns and adjectives are expected to exhibit greater perceptual resilience than function words, reflected in both higher behavioral accuracy and enhanced P300 amplitudes, the latter indexing attentional resource allocation during processing. Second, proficiency effects: advanced L2 learners should demonstrate more native-like neural profiles, including stronger early sensory responses (N100) and phonological mapping signals (PMN), whereas beginning learners are predicted to rely more heavily on later integrative processes such as the N400. Third, phonetic interactions: the manner of articulation is expected to modulate lexical effects, with stops and affricates serving as more robust perceptual anchors than approximants and rhotic consonants, particularly for learners at lower proficiency levels.

Theoretically, this work advances our understanding of the neurocognitive architecture underlying speech perception by clarifying how linguistic category, proficiency, and phonetic features jointly shape auditory processing. Practically, the findings can inform L2 listening pedagogy, guide assessment of auditory processing difficulties, and enhance speech recognition technologies for multilingual users.

2. LITERATURE REVIEW

2.1. Neural Foundations of Speech Perception in Noise

Speech perception in noise represents a complex neurocognitive challenge engaging multiple auditory and cognitive systems. Neuroimaging studies reveal that speech-in-noise tasks activate not only canonical auditory regions but also recruit additional cortical networks compared to quiet conditions (Wong et al., 2009; Du et al., 2014). Functional MRI evidence highlights increased activation in the left inferior frontal gyrus, bilateral anterior insula, and premotor cortex when speech is degraded, suggesting compensatory reliance on articulatory-based mechanisms (Hervais-Adelman et al., 2012). Electrophysiological findings corroborate this view: noise reduces the temporal precision of auditory brainstem encoding, cascading into disruptions in higher-level linguistic processing (Anderson & Kraus, 2010).

Event-related potential (ERP) research documents systematic modulation of multiple components under adverse conditions. The N100, indexing early auditory encoding, shows diminished amplitude and delayed latency in noise, reflecting compromised sensory processing (Billings et al., 2013). The phonological mapping negativity (PMN), associated with phonological categorization, is highly sensitive to acoustic degradation (Newman & Connolly, 2009). Critically, the P300, a marker of attentional allocation and context updating, exhibits reduced amplitude during speech-in-noise perception, pointing to resource limitations in challenging environments (Parbery-Clark et al., 2011).

For L2 learners, these neural challenges are magnified by less robust phonological representations in the target language. Compared to native listeners, L2 processing engages more diffuse cortical networks and exhibits delayed semantic integration, as reflected in later N400 responses in noisy conditions (Rüschemeyer et al., 2006; FitzPatrick & Indefrey, 2014). These differences help explain why L2 learners experience disproportionate difficulty in adverse acoustic settings.

2.2. Lexical Category Processing and Neural Specificity

Evidence increasingly suggests that lexical categories recruit distinct neural mechanisms. MEG studies demonstrate that nouns and verbs activate different cortical networks, with nouns engaging visual-temporal regions and verbs recruiting frontal motor areas, emerging as early as 150–200 ms post-stimulus (Pulvermüller et al., 2012). Similarly, oscillatory analyses reveal stronger theta and gamma responses for open-class words (nouns, verbs, adjectives) than for closed-class words, likely reflecting richer semantic networks (Bastiaansen et al., 2008). ERP studies also consistently show larger N400 responses for open-class words, indicating greater semantic integration demands (Brown et al., 1999).

Yet the interaction between lexical category and noise on neural processing remains underexplored. Given that content words exhibit greater acoustic-phonetic salience and semantic load than function words, they may provide stronger perceptual anchors under masking conditions. In contrast, function words, with reduced stress and frequent vowel reduction, may be more susceptible to noise interference (Kerlin et al., 2010).

2.3. Phonetic Resilience and Perceptual Anchors in L2 Processing

Certain phonetic cues appear to support perception under adverse listening conditions. Consonants with abrupt spectral transitions such as stops and affricates serve as reliable perceptual anchors compared to sounds with gradual onsets like nasals and approximants (Miller & Nicely, 1955; Wright, 2004). The Phonetic Prominence Hypothesis (Cutler & Butterfield, 1992) posits that segments with high acoustic salience resist misperception and facilitate lexical access, a property disproportionately associated with content words.

For L2 learners, cross-linguistic similarities further modulate resilience. The Perceptual Assimilation Model (Best & Tyler, 2007) argues that L2 sounds assimilated into existing L1 categories are perceived more accurately, whereas novel or ambiguous sounds remain vulnerable to noise masking. These effects likely interact with lexical category, as function words already phonetically reduced offer fewer robust cues for L2 listeners under masking conditions.

2.4. Working Memory, Cognitive Load, and Attention in L2 Speech Perception

Adverse listening conditions impose heavy demands on working memory and attention, particularly for L2 listeners managing both acoustic degradation and non-native processing demands. Neuroimaging studies show stronger recruitment of executive control networks during L2 speech perception in noise than in quiet, suggesting increased cognitive load (Abutalebi & Green, 2007).

Individual differences in working memory capacity predict speech-in-noise comprehension, with higher spans linked to greater resistance to interference (Skoe et al., 2019). According to Cognitive Load Theory (Mattys et al., 2009), listeners reallocate limited cognitive resources depending on situational demands: under noise, resources shift toward acoustic-phonetic decoding at the expense of lexical-semantic processing. This reallocation likely penalizes complex lexical categories, such as morphologically rich verbs, under high load.

Attention also plays a pivotal role: selective attention enhances the neural representation of target speech in noise (Kerlin et al., 2010). The Automatic Selective Perception Model (Strange, 2011) proposes that attentional mechanisms become more efficient with L2 proficiency, potentially narrowing the performance gap between content and function words as learners advance.

2.5. Methodological Advances in Studying L2 Speech Perception

Recent methodological innovations integrate behavioral, neurophysiological, and computational approaches to capture real-time speech processing in noise. Time-frequency EEG analyses reveal oscillatory signatures associated with phonological, syntactic, and semantic processing (Bastiaansen et al., 2012). Eye-tracking provides millisecond-level insights into lexical competition dynamics, demonstrating delayed resolution in L2 listeners under adverse conditions (McQueen & Viebahn, 2007).

Computational models, including Bayesian frameworks (Norris & McQueen, 2008), formalize how prior knowledge and sensory evidence interact to reconstruct degraded signals, offering tools for modeling L2-specific constraints. Finally, mixed-effects models and growth curve analyses now account for participant- and item-level variability, improving statistical precision and generalizability critical advances given the heterogeneity of L2 populations.

2.6. Synthesis

Together, these findings underscore that speech-in-noise perception reflects a dynamic interplay among neurocognitive mechanisms, lexical category properties, phonetic salience, and cognitive resources. However, the combined effects of lexical class, phonetic features, and L2 proficiency on neural speech processing in noise remain insufficiently understood. Addressing this gap, the present study integrates behavioral accuracy, EEG measures, and lexical-phonetic manipulations to provide a comprehensive account of speech reconstruction under adverse listening conditions.

3. METHODOLOGY

3.1. Research Design

This study employed a mixed-methods design integrating behavioral measures with neurophysiological recordings to investigate lexical category resilience in noise. A quasi-experimental approach (Campbell & Stanley, 2015) compared L2 learners at two proficiency levels (Level 2: beginner; Level 8: advanced) during speech perception tasks under controlled noise conditions (Broersma & Cutler, 2011).

The research followed a 2 × 4 × 7 factorial design with three independent variables

1. Proficiency level: beginner, advanced
2. Lexical category: noun, verb, adjective, adverb
3. Manner of articulation: stop, nasal, fricative, affricate, rhotic approximant (/ɹ/), approximant (w/y), lateral

Both word-level and sentence-level stimuli were included to capture lexical effects in isolation and in syntactic context, allowing comprehensive assessment under adverse listening conditions.

3.2. Participants

Participants were 40 Saudi tertiary students (20 per proficiency group) enrolled in an English language program at a major university. Sample size was determined through power analysis using GPower (Faul et al., 2009) based on pilot effect sizes (d = 0.65, α = 0.05, power = 0.80).

- Level 2 (Beginner): n = 20; 12 male, 8 female; M age = 19.4 years (SD = 1.2); 150 hours of English instruction; Oxford Placement Test scores: 15–25
- Level 8 (Advanced): n = 20; 11 male, 9 female; M age = 21.2 years (SD = 1.5); 700+ hours of English instruction; Oxford Placement Test scores: 55–65

All participants were right-handed native Arabic speakers with normal hearing (confirmed via pure-tone audiometry) and normal or corrected vision. None reported neurological disorders, speech/language impairments, or extensive English exposure outside the classroom. All procedures were approved by the university's Institutional Review Board, and written informed consent was obtained from all participants in accordance with the 2013 Helsinki Declaration.

3.3. Stimuli Development

Lexical and Phonetic Control. Stimuli were selected using the Academic Word List (Coxhead, 2000) and British National Corpus frequency data (Leech et al., 2001). A total of 28 target words (7 per lexical category) were

- Matched for syllable length (monosyllabic) and frequency (20–50 occurrences per million).
- Within the known vocabulary of beginner participants, verified against their course materials.

Each lexical category contained words representing all seven manners of articulation in word-initial position for maximal perceptual salience (Table 1).

Table 1: Word-Level Stimuli by Lexical Category and Manner of Articulation.

Lexical Class	Stop (p/b)	Affricate (ch/j)	Fricative (f/v)	Nasal (m/n)	Rhotic (/ɹ/)	Approximant (w/y)	Lateral (l)
Nouns	pen, bat	chip, jam	fish, van	man, net	road	wall, yard	lamp
Verbs	put, buy	check, join	fix, visit	meet, need	run	wait, yell	look
Adjectives	poor, big	cheap, just	fast, vivid	main, new	real	warm, young	loud
Adverbs	past, back	cheaply, justly	fast, very	mainly, nearly	rarely	wildly, yearly	lately

3.4. Sentence Construction and Validation

We constructed 28 sentences (4 per manner of articulation), each containing 6–8 words, a simple SVO structure, and moderate semantic predictability. Ten native English speakers validated the sentences, achieving >95% accuracy in quiet listening conditions.

Sample sentences included

- Stops: The cat is sitting on the mat.
- Nasals: Mike and Nina met Monday.
- Fricatives: Sophie feels safe outside.
- Affricates: Charlie jumped in joy.
- Rhotics: Randy rode a red bike.
- Approximants: You and William waited patiently.
- Laterals: Lily loves lemon lollipops.

3.5. Audio Recording and Noise Calibration

Recordings were produced by a female native speaker of American English in a sound-attenuated booth (Shure SM58 microphone, 44.1 kHz, 16-bit). Speech-shaped noise was added to achieve –3 dB SNR for word stimuli and 0 dB SNR for sentence stimuli. Pilot testing calibrated SNRs to yield 50–60% accuracy in advanced learners, ensuring adequate performance variability for statistical analysis.

3.6. EEG Data Acquisition and Processing

EEG signals were recorded using a 64-channel ActiCap system (Brain Products GmbH) at 1000 Hz with FCz reference and AFz ground. Data were preprocessed with ICA-based artifact rejection (±75 μV), 0.1–30 Hz band-pass filtering, and baseline correction (–200 to 0 ms).

ERP components of interest included:

- N100 (80–150 ms): early sensory processing
- PMN (250–350 ms): phonological mapping
- P300 (300–500 ms): attention/context updating
- N400 (350–550 ms): semantic integration
- Mean amplitude values were extracted for each component for statistical analysis.

3.7. Experimental Procedure

Participants completed two tasks

1. Word recognition: typing the word they heard.
2. Sentence recognition: typing the full sentence.

The experiment consisted of four blocks: two-word blocks (70 trials each) and two sentence blocks (35 trials each). Practice trials were administered before the main experiment.

EEG was recorded in an electrically shielded booth with participants fixating on a central cross. Stimulus presentation was synchronized with EEG event markers, and responses were collected after stimulus offset. After the EEG session, participants completed working memory (Unsworth et al., 2005) and phonological awareness (Wagner et al., 1999) assessments to evaluate individual differences.

3.8. Data Analysis Plan

Behavioral accuracy was analyzed using mixed-effects logistic regression with the model specification: Accuracy~Proficiency×Lexical Category×Manner+(1|Subject)+(1|Item). This model included Proficiency, Lexical Category, Manner of Articulation, and all their interactions as fixed effects, with by-subject and by-item intercepts as random effects. To control for multiple comparisons, post-hoc contrasts were adjusted using Bonferroni correction. Additionally, error types were categorized descriptively into substitution, omission, and metathesis patterns for further analysis.

ERP data were analyzed via mixed ANOVAs with Greenhouse–Geisser correction for sphericity violations. Brain–behavior correlations examined links between ERP amplitudes and accuracy.

Time–frequency analyses compared theta, alpha, beta, and gamma oscillations for successful vs. unsuccessful trials.

Finally, structural equation modeling tested whether neural indices mediated lexical and phonetic effects, and support vector machine (SVM) classification identified the minimal neural–behavioral feature set predicting accurate perception.

All analyses were performed in R (4.1.0), MATLAB (R2021a), and JASP (0.16) with $\alpha = .05$ and corrections for multiple comparisons.

4. RESULTS

4.1. Behavioural Findings

4.1.1. Word-Level Recognition Performance

Mixed-effects logistic regression revealed significant main effects of lexical category ($\chi^2(3) = 47.82$, $p < .001$), proficiency level ($\chi^2(1) = 38.54$, $p < .001$), and manner of articulation ($\chi^2(6) = 61.37$, $p < .001$), along with interactions between lexical category and proficiency ($\chi^2(3) = 9.45$, $p = .024$) and between lexical category and manner of articulation ($\chi^2(18) = 33.92$, $p = .013$).

Table 2 summarizes recognition accuracy by lexical category and proficiency level. Adjectives achieved the highest overall accuracy (51.4%), followed by nouns (47.1%), adverbs (44.3%), and verbs (37.1%). Pairwise contrasts showed significant differences among all lexical categories ($p < .05$) except between nouns and adverbs ($p = .28$). Accuracy gains from Level 2 to Level 8 were consistent across categories (≈20 percentage points) but largest for verbs, suggesting proficiency mitigates, but does not eliminate, lexical category effects.

Table 2: Word Recognition Accuracy (%) by Lexical Category and Proficiency Level.

Proficiency Level	Nouns	Verbs	Adjectives	Adverbs	Overall Mean
Level 8 (Advanced)	57.1	47.1	61.4	54.3	55.0
Level 2 (Beginning)	37.1	27.1	41.4	34.3	35.0
Mean Difference	20.0	20.0	20.0	20.0	20.0

Manner of articulation effects showed stops (63.2%) and affricates (58.7%) yielded the highest accuracy, whereas approximants (27.5%) and trills (32.1%) performed poorest, especially in verbs where approximant-initial tokens fell to 10% accuracy for beginners.

4.1.2. Sentence-Level Recognition Performance

Sentence-level accuracy exceeded word-level accuracy, confirming the facilitatory role of context. Mixed-effects regression revealed significant effects of sentence type ($F(6,33) = 18.72$, $p < .001$) and proficiency level ($F(1,38) = 42.56$, $p < .001$). Sentences dominated by stop consonants showed the highest accuracy (84% for Level 8; 66% for Level 2), whereas approximant-dominated sentences showed the lowest (49% for Level 8; 29% for Level 2).

Function words (articles, prepositions, conjunctions) had significantly lower recognition (41.7%) than content words (nouns = 61.3%, verbs = 53.2%, adjectives = 64.5%, adverbs = 56.8%), with larger proficiency gaps for function words (26.8 points) than for content words (17.4 points).

4.1.3. Error Pattern Analysis

Chi-square analysis confirmed lexical category differences in error types ($\chi^2(18) = 82.63$, $p < .001$). Nouns showed more consonant substitutions (31.2%) and omissions (27.5%), verbs exhibited vowel reduction (36.7%) and consonant confusion (29.3%), adjectives showed metathesis (24.1%) and vowel shifts (28.6%), while adverbs were marked by syllable deletion (38.2%).

Table 3: Distribution of Error Types (%) by Lexical Category.

Error Type	Nouns	Verbs	Adjectives	Adverbs
Substitution	31.2	18.4	20.3	15.7
Omission	27.5	13.9	15.6	10.2
Insertion	5.8	3.5	7.2	11.4
Metathesis	11.3	15.3	24.1	9.8
Vowel Reduction	12.4	36.7	28.6	38.2
Consonant Confusion	8.9	29.3	16.8	22.3
Allophonic Variation	2.9	4.9	7.4	8.4

Approximants and trills were most prone to consonant confusion, while fricatives and affricates frequently triggered metathesis errors. Stops, though generally resilient, showed predictable voicing substitutions (e.g., /p/ → /b/).

4.2. Neurophysiological Findings

N100 (80–150 ms). Advanced learners showed significantly larger N100 amplitudes (–3.21 μV) than beginners (–2.47 μV; $F(1,38) = 7.24$, $p = .011$, $\eta^{2p} = 0.16$), indicating more efficient early auditory processing. Manner of articulation influenced N100 amplitude ($F(6,228) = 4.93$, $p < .001$, $\eta^{2p} = 0.11$), with stops and affricates producing stronger responses than approximants and nasals.

4.2.1. PMN (250–350 ms)

PMN amplitude was modulated by lexical category ($F(3,114) = 5.38$, $p = .002$, $\eta^{2p} = 0.12$), proficiency ($F(1,38) = 9.72$, $p = .003$, $\eta^{2p} = 0.20$), and their interaction ($F(3,114) = 3.46$, $p = .019$, $\eta^{2p} = 0.08$). Content words, especially adjectives and nouns, elicited larger PMN responses than verbs ($p < .05$), particularly in advanced learners, suggesting stronger phonological mapping processes.

4.2.2. P300 (300–500 ms)

Significant main effects of lexical category ($F(3,114) = 11.27$, $p < .001$, $\eta^{2p} = 0.23$), proficiency ($F(1,38) = 5.63$, $p = .023$, $\eta^{2p} = 0.13$), and manner of articulation ($F(6,228) = 3.94$, $p = .001$, $\eta^{2p} = 0.09$) were observed.

Content words elicited larger P300 amplitudes than verbs across both proficiency groups, indicating greater attentional resource allocation. Importantly, P300 amplitude correlated strongly with behavioral accuracy ($r = 0.68$, $p < .001$), more so for advanced learners ($r = 0.74$) than beginners ($r = 0.59$).

4.2.3. N400 (350–550 ms)

Content words produced larger N400 responses than verbs ($F(3,114) = 8.75$, $p < .001$, $\eta^{2p} = 0.19$), particularly in beginners. Function words showed attenuated N400 responses ($F(1,38) = 19.63$, $p < .001$, $\eta^{2p} = 0.34$), especially in noise, indicating disrupted semantic integration under adverse conditions.

4.3. Time–Frequency Analyses

Time–frequency analyses revealed distinct oscillatory patterns across frequency bands. Theta synchronization (4–7 Hz) was stronger for correctly recognized items ($F(1,38) = 12.47$, $p < .001$), particularly for nouns and adjectives ($F(3,114) = 4.18$, $p = .008$). Alpha desynchronization (8–12 Hz) was greater for content words compared to verbs ($F(3,114) = 5.72$, $p = .001$), with effects most pronounced among advanced learners. Finally, gamma synchronization (30–80 Hz) was stronger for adjectives and nouns than for verbs and adverbs ($F(3,114) = 3.84$, $p = .012$), especially when words were correctly recognized.

4.4. Integrative Analyses

Structural equation modeling (CFI = 0.94, RMSEA = 0.056) showed P300 amplitude partially mediated the relationship between lexical category and accuracy, accounting for 42% of variance.

- Advanced learners relied more on early components (N100, PMN).
- Beginners depended on later components (P300, N400), indicating a shift toward automatic processing with proficiency.

Machine learning classification (SVM) achieved 78.3% accuracy, with P300 amplitude, theta power, and gamma synchronization emerging as top predictors of successful perception. Accuracy was highest for adjectives (83.1%) and nouns (80.7%), lowest for verbs (72.4%).

Working memory correlated with perception of verbs ($r = 0.56$, $p < .001$) and function words ($r = 0.61$, $p < .001$), while phonological awareness correlated with PMN amplitude ($r = 0.48$, $p = .002$) and theta power ($r = 0.43$, $p = .005$).

4.5. Summary of Key Findings

Adjectives and nouns demonstrated greater resilience to noise than verbs and function words, a hierarchy mirrored in PMN and P300 amplitudes, linking lexical class to phonological mapping and attention allocation. Stops and affricates provided stronger perceptual anchors than approximants and trills, especially within verbs, while advanced learners engaged earlier processing stages (N100, PMN) compared to beginners, who relied more on later, attention-driven components (P300, N400). Across conditions, P300, theta, and gamma activity predicted behavioral success and mediated lexical effects on perception, with working memory and phonological awareness contributing most strongly to processing low-resilience categories and supporting early-stage neural mechanisms.

5. DISCUSSION

This study investigated the neurocognitive foundations of differential resilience to noise across lexical categories in L2 speech perception. By integrating behavioral accuracy with electrophysiological measures, we identified distinct neural signatures underlying the successful reconstruction of degraded speech. Across tasks, adjectives and nouns consistently showed greater resistance to noise than verbs and function words, a hierarchy reflected in both performance metrics and neural processing patterns.

5.1. Lexical Category Effects on Speech Perception

Our first research question addressed how speech processing differs across lexical categories under adverse conditions. Recognition accuracy revealed a clear hierarchy: adjectives > nouns > adverbs > verbs, with adjectives reaching the highest accuracy (51.4%) and verbs the lowest (37.1%). This pattern aligns with studies suggesting that content words carry greater semantic weight and prosodic prominence than function words (Field, 2004; Morrill et al., 2015) but extends prior work by showing systematic differences within content words themselves.

Neurophysiologically, adjectives and nouns elicited larger PMN responses than verbs, suggesting more robust phonological mapping and activation of lexical representations. Similarly, P300 amplitudes, indexing attentional resource allocation, were larger for adjectives and nouns, indicating that these word classes may inherently capture more processing resources. Linguistically, adjectives and nouns often carry richer semantic content and less morphosyntactic complexity than verbs (Cutler, 2012), facilitating top-down lexical activation when bottom-up cues are degraded.

The comparatively poorer performance of verbs, despite their status as content words, suggests that grammatical function and morphosyntactic demands influence perceptual resilience beyond the traditional content/function distinction.

5.2. Neurophysiological Markers of Successful Perception

Our second research question examined the neural signatures distinguishing successful from unsuccessful perception. P300 amplitude emerged as a robust predictor of accurate word recognition ($r = 0.68$), confirming its role in attention allocation and context updating under challenging conditions.

Time–frequency analyses revealed complementary neural mechanisms underlying speech perception. Theta synchronization (4–7 Hz), linked to phonological working memory (Bastiaansen et al., 2008), was stronger for correctly recognized adjectives and nouns, while gamma synchronization (30–80 Hz), associated with feature binding, also differentiated successful recognition of these word classes, suggesting more efficient acoustic–phonetic integration. Proficiency effects further clarified these patterns: advanced learners exhibited larger N100 and PMN responses, reflecting more efficient early sensory encoding and phonological mapping, whereas beginners relied more heavily on later attention-driven processes (P300, N400), consistent with a shift from controlled to automatic processing as proficiency develops (Segalowitz & Hulstijn, 2005). Together, these findings indicate that automatization emerges earlier for certain lexical categories, particularly adjectives and nouns, marking a developmental trajectory toward more efficient speech processing.

5.3. Interaction of Phonetic Features and Lexical Category

Our third research question explored whether manner of articulation interacts with lexical category in shaping speech perception under noise. Stops (63.2%) and affricates (58.7%) served as perceptual anchors, whereas approximants (27.5%) and trills (32.1%) were most vulnerable especially in verbs.

Neural evidence showed reduced PMN responses for verbs beginning with approximants compared to those with stops, indicating weaker phonological mapping when both lexical and phonetic factors posed challenges.

These findings support the Phonetic Prominence Hypothesis (Cutler & Butterfield, 1992) but further suggest that lexical function modulates acoustic salience: the same phonetic feature can produce different perceptual outcomes depending on grammatical category and cognitive load.

5.4. Working Memory and Phonological Awareness

Integrative analyses revealed that working memory capacity correlated strongly with recognition accuracy for verbs (r = 0.56) and function words (r = 0.61), suggesting these word classes impose greater cognitive demands, particularly for beginners.

Similarly, phonological awareness correlated with PMN amplitude (r = 0.48) and theta power (r = 0.43), especially for perceptually salient segments (e.g., stops, affricates). These findings indicate that metalinguistic skills facilitate both behavioral accuracy and neural efficiency in challenging listening conditions.

Together, these results support Cognitive Load Theory (Mattys et al., 2009): beginners allocate greater resources to processing difficult word classes, while advanced learners rely on more automatized mechanisms requiring fewer cognitive resources.

5.5. Theoretical Implications

This study advances speech perception and L2 acquisition theory in several ways. First, the findings challenge models treating lexical categories as homogeneous, demonstrating that grammatical function and phonetic structure jointly shape perceptual resilience, revealing a lexical hierarchy in perception. Second, proficiency-dependent differences in attention allocation suggest that selective attention develops at different rates across lexical categories, aligning with the Automatic Selective Perception model (Strange, 2011). Third, neural evidence links acoustic salience to both phonological mapping and perceptual resilience, with effects modulated by lexical function, supporting the Phonetic Prominence Hypothesis (Cutler & Butterfield, 1992). Finally, the results indicate that lexical activation strength in noise may vary by grammatical category, with adjectives and nouns receiving stronger activation than verbs, consistent with the Distributed Cohort Model (Gaskell & Marslen-Wilson, 2002).

5.6. Pedagogical and Technological Applications

Pedagogically, the findings underscore the importance of targeting vulnerable lexical and phonetic categories, particularly verbs beginning with approximants or trills, which may benefit from explicit, focused training. Incorporating metalinguistic components such as phonological awareness and working memory exercises could further enhance listening comprehension in noisy or adverse conditions. Additionally, adaptive task designs that gradually reduce background noise for challenging word classes may facilitate perceptual learning and long-term retention. Technologically, the results can inform the development of automatic speech recognition systems that weight lexical and phonetic features according to their perceptual salience, as well as assistive listening devices capable of integrating neural markers (e.g., P300 amplitudes, theta synchronization) to provide real-time, neuro-informed speech enhancement.

6. LIMITATIONS AND FUTURE DIRECTIONS

This study has several limitations. First, although word frequency and syllable length were controlled, other psycholinguistic variables such as imageability, concreteness, and lexical neighborhood density were not systematically addressed. These factors are known to affect lexical access and auditory word recognition, especially under noisy conditions, and their absence may have influenced the observed differences across lexical categories. Future research should incorporate standardized measures of these variables to better isolate lexical-class effects. Second, the exclusive focus on Saudi Arabic learners of English limits the generalizability of the findings across language backgrounds. Cross-linguistic differences in phonology, prosody, and lexical structure may influence how listeners perceive speech in noise. Future research should therefore include learners from typologically diverse L1 backgrounds to determine whether the observed lexical-class effects and neural signatures extend across languages and proficiency profiles. Third, while EEG provided high temporal resolution, it lacked spatial precision. Combining EEG with fMRI could capture both the timing and localization of lexical-category effects more comprehensively. Finally, the cross-sectional design cannot fully capture developmental trajectories. Longitudinal studies tracking learners over time would clarify how proficiency shapes perceptual resilience. Future work should also examine semantic predictability, syntactic complexity, and naturalistic listening environments to enhance ecological validity.

7. CONCLUSION

This study examined the neurocognitive foundations of phonetic resilience across lexical categories in L2 speech perception under adverse listening conditions. By integrating behavioral measures with EEG recordings, we identified distinct neural signatures associated with successful reconstruction of degraded speech signals. Across tasks, adjectives and nouns consistently demonstrated greater resistance to noise than verbs and function words, a behavioral hierarchy mirrored in PMN and P300 responses, reflecting more robust phonological mapping and attentional resource allocation.

Our findings indicate that phonetic composition, grammatical function, and cognitive constraints jointly shape the perceptual resilience of lexical categories. Manner of articulation interacted with word class, with verbs showing particular vulnerability when combined with approximants and trills. Advanced learners exhibited more efficient early sensory processing and phonological mapping, whereas beginning learners relied more on later, attention-mediated processes. Moreover, working memory capacity and phonological awareness correlated strongly with both behavioral accuracy and neural indices, especially for low-resilience categories and less proficient learners.

Theoretically, these results advance models of L2 speech perception by demonstrating that lexical class, phonetic salience, and cognitive resources interact dynamically, shaping processing hierarchies that evolve with proficiency. Pedagogically, they suggest that L2 listening instruction should target the most vulnerable categories particularly verbs and phonetically complex words while supporting metalinguistic skills such as phonological awareness to enhance processing efficiency.

Future research should extend these findings to naturalistic listening contexts, longitudinal designs, and typologically diverse languages to clarify how linguistic, phonetic, and neurocognitive factors jointly drive the development of perceptual resilience in L2 learners.

Acknowledgements: This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2502)

REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. Journal of Neurolinguistics, 20(3), 242-275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. Journal of the American Academy of Audiology, 21(9), 575-585. https://doi.org/10.3766/jaaa.21.9.3

Bastiaansen, M., Magyari, L., & Hagoort, P. (2008). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. Journal of Cognitive Neuroscience, 20(7), 1162-1174. https://doi.org/10.1162/jocn.2008.20080

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), The Oxford handbook of event-related potential components (pp. 31-49). Oxford University Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), Language experience in second language speech learning: In honor of James Emil Flege (pp. 13-34). John Benjamins Publishing.

Billings, C. J., McMillan, G. P., Penman, T. M., & Gille, S. M. (2013). Predicting perception in noise using cortical auditory evoked potentials. Journal of the Association for Research in Otolaryngology, 14(6), 891-903. https://doi.org/10.1007/s10162-013-0415-y

Boersma, P., & Weenink, D. (2021). Praat: Doing phonetics by computer [Computer program]. Version 6.1.53. Retrieved from http://www. praat.org/

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. The Quarterly Journal of Experimental Psychology, 64(1), 74-95. https://doi.org/10.1080/17470218. 2010.499174

Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. Journal of Cognitive Neuroscience, 11(3), 261-281. https://doi.org/10.1162/089892999563382

Campbell, D. T., & Stanley, J. C. (2015). Experimental and quasi-experimental designs for research. Ravenio Books.

Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213-238. https://doi.org/10.2307/3587951

Cutler, A. (2012). Native listening: Language experience and the recognition of spoken words. MIT Press.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. Journal of Memory and Language, 31(2), 218-236. https://doi.org/10.1016/0749-596X(92)90012-M

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proceedings of the National Academy of Sciences, 111(19), 7126-7131. https://doi.org/10.1073/pnas. 1318738111

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Field, J. (2004). Psycholinguistics: The key concepts. Routledge.

FitzPatrick, I., & Indefrey, P. (2014). Head start for target language in bilingual listening. Brain Research, 1542, 111-130. https://doi. org/10.1016/j.brainres.2013.10.014

Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. Cognitive Psychology, 45(2), 220-266. https://doi.org/10.1016/S0010-0285(02)00003-8

Golestani, N., Hervais-Adelman, A., Obleser, J., & Scott, S. K. (2013). Semantic versus perceptual interactions in neural processing of speech-in-noise. NeuroImage, 79, 52-61. https://doi.org/10.1016/j.neuroimage.2013.04.049

Grossberg, S. (2003). Resonant neural dynamics of speech perception. Journal of Phonetics, 31(3- 4), 423-445. https://doi.org/10.1016/S0095-4470(03)00051-2

Hervais-Adelman, A., Carlyon, R. P., Johnsrude, I. S., & Davis, M. H. (2012). Brain regions recruited for the effortful comprehension of noise-vocoded words. Language and Cognitive Processes, 27(7-8), 1145-1166. https://doi.org/10.1080/01690965.2012.662280 JASP Team. (2021). JASP (Version 0.16) [Computer software].

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party". Journal of Neuroscience, 30(2), 620-628. https://doi.org/10.1523/JNEUROSCI.3631-09.2010

Leech, G., Rayson, P., & Wilson, A. (2001). Word frequencies in written and spoken English: Based on the British National Corpus. Longman.

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. Cognitive Psychology, 59(3), 203-243. https://doi.org/10.1016/j.cogpsych.2009.04.001

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. Quarterly Journal of Experimental Psychology, 60(5), 661-671. https://doi.org/10.1080/17470210601183890

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. The Journal of the Acoustical Society of America, 27(2), 338-352. https://doi.org/10.1121/1.1907526

Morrill, T. H., Dilley, L. C., McAuley, J. D., & Pitt, M. A. (2015). Distal rhythm influences whether or not listeners hear a word in continuous speech: Support for a perceptual grouping hypothesis. Cognition, 144, 87-100. https://doi.org/10.1016/j.cognition.2015.07.012

Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom-up and top-down effects on spoken word processing. Biological Psychology, 80(1), 114-121. https://doi.org/10.1016/j.biopsycho.2008.04.008

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. Psychological Review, 115(2), 357-395. https://doi.org/10.1037/0033-295X.115.2.357

Parbery-Clark, A., Marmel, F., Bair, J., & Kraus, N. (2011). What subcortical-cortical relationships tell us about processing speech in noise. European Journal of Neuroscience, 33(3), 549-557. https://doi.org/10.1111/j.1460-9568.2010.07546.x

Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., & Nimmo-Smith, I. (2012). Distributed cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. Human Brain Mapping, 33(5), 992-1007. https://doi.org/10.1002/hbm.21254

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rüschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2006). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. Human Brain Mapping, 27(2), 144-157. https://doi.org/10.1002/hbm.20176

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), Handbook of bilingualism: Psycholinguistic approaches (pp. 371-388). Oxford University Press.

Skoe, E., Brody, J., & Theodore, R. M. (2019). Bilingualism and speech understanding in noise: Auditory and linguistic factors. Journal of the American Academy of Audiology, 30(4), 322-332. https://doi.org/10.3766/jaaa.17099

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. Journal of Phonetics, 39(4), 456-466. https://doi.org/10.1016/j.wocn.2010.09.001

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. Behavior Research Methods, 37(3), 498-505. https://doi.org/10.3758/BF03192720

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive Test of Phonological Processing: CTOPP. Pro-Ed.

Wong, P. C., Uppunda, A. K., Parrish, T. B., & Dhar, S. (2009). Cortical mechanisms of speech perception in noise. Journal of Speech, Language, and Hearing Research, 52(5), 1122-1133. https://doi.org/10.1044/1092-4388(2009/08-0131)

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), phonetically based phonology (pp. 34-57). Cambridge University Press.