Wah Academia Journal of Social Sciences
Volume 5, Issue 2, Jun 2026, pp. 466-488
ISSN – E 2958-8731 P 2958-8723
https://doi.org/10.63954/WAJSS.5.2.29.2026

Hesitation Pauses at Clause Boundaries in Pakistani English and Urdu: A Cross-Linguistic Study with Speaker Reflections

Dr Kamran Ali[1], Shahzeb Shafi[2] and Muhammad Hamzah Masood[3]

Abstract

This study investigates the distribution of hesitation pauses at clause boundaries in spontaneous Pakistani English and Urdu speech, examined from psycholinguistic and neurolinguistic perspectives. Original speech data were collected from 10 Pakistani English and 10 Urdu adult speakers, yielding approximately 110 minutes of spontaneous narration. Six participants additionally completed structured speaker-reflection interviews to examine their metalinguistic awareness of hesitation behaviour. Each hesitation pause token was coded as either Clause Boundary (CB) or Clause Internal (CI), and the positional distribution was tested against chance expectation using chi-square analysis. Results showed that 71% of hesitation pauses in Pakistani English and 64% in Urdu occurred at clause boundaries, replicating and extending Hawkins' (1971) clause-boundary hypothesis in a cross-linguistic South Asian context for the first time. Both comparisons yielded statistically significant associations between pause position and clause boundary status: Pakistani English $\chi^2(1, N = 263) = 21.4$, $p < .001$, Cramér's V = 0.29; Urdu $\chi^2(1, N = 224) = 14.7$, $p < .001$, Cramér's V = 0.26. Speaker reflections identified lexical search and propositional planning as the two most frequently perceived causes of hesitation. Findings are interpreted within Levelt's (1983, 1993) Incremental Sentence Planning Theory and grounded in neurolinguistic evidence — including ERP and neuroimaging studies of syntactic planning — for elevated processing load at clause boundaries.

Keywords: Hesitation Pauses, Clause Boundary, Disfluency, Speech Planning, Filled Pauses, Pakistani English, Urdu, Cross-Linguistic

[1]Department of Humanities and Social Sciences, Bahria University, Karachi – Pakistan
[2]Department of English, Faculty of Social Sciences and Humanities, Hamdard University, Main Campus, Karachi – Pakistan
[3]Department of Linguistics and Communications, School of Liberal Arts, University of Management and Technology, Lahore – Pakistan

Introduction

The basic feature of spontaneous speech is the process of linguistic planning that must be performed in real time: the conceptualization and access to the lexicon, the preparation of the syntactic structure, and the articulation of the sentence must be accomplished under considerable temporal constraints. In contrast to written discourse, writers can edit their documents before finalizing them, but speakers have little opportunity to edit their discourse before it is actually spoken. In this regard, the spontaneous speaker has to plan his/her utterances step by step during the speech. The temporal pressures lead to a variety of disfluencies: filled pauses, repetitions, false starts and self-corrections etc. While disfluencies have been regarded as performance errors, in recent years psycholinguistic research has viewed them as theoretically rich source for understanding the cognitive architecture of spontaneous speech production (Clark, 2002; Clark & Fox Tree, 2002; Levelt, 1983, 1993).

Hesitation pauses have been considered as a critical phenomenon for inferring the cognitive architecture of spontaneous speech production across all disfluency types. Previous research on hesitation pauses has shown that hesitation behaviour is not random but is systematically associated with phases of higher cognitive processing, such as lexical retrieval, syntactic planning and discourse organization (Beattie, 1979; Goldman-Eisler, 1968; Maclay & Osgood, 1959). One of the major issues in this literature is the syntactic position of hesitation pauses, specifically the frequency of pauses at clause boundaries relative to their distribution within clauses, which has implications for the cognitive and neural organization of speech planning.

A significant contribution to this discussion was made by Hawkins (1971), whose seminal study The Syntactic Location of Hesitation Pauses found that two-thirds of the total number of pauses and nearly three-quarters of the total pause time were found at clause boundaries. This result has long been taken to provide strong support for the idea that speakers plan spontaneous speech in clause-sized units rather than word-by-word. However, Hawkins' original study was conducted on 48 children using an elicited narrative task, employed no inferential statistics, and has never been replicated in a cross-linguistic context involving South Asian languages. These are significant methodological gaps that the present study is designed to address.

The present study directly addresses three gaps in the existing literature. First, no prior study has examined clause-boundary hesitation pauses in Pakistani English, a

legitimate and distinct variety of World English spoken by over 100 million educated speakers. Second, no cross-linguistic comparison of hesitation pause distribution has been conducted between Pakistani English and Urdu, despite the typological and sociolinguistic proximity of these two languages in Pakistani bilingual contexts. Third, no prior study has combined distributional corpus analysis with structured speaker-reflection data to examine the metalinguistic dimension of hesitation behaviour — that is, what speakers themselves perceive as the cause of their pauses.

The present study accordingly collects original spontaneous speech data from 20 adult bilingual participants in Karachi, Pakistan, recording 10 educated bilingual speakers who use Pakistani English as their primary academic language (L1 Urdu/L2 English) and 10 Urdu-dominant speakers recorded in Urdu in a standardised narration task. A subset of six participants additionally completed structured reflection interviews. All data were collected, transcribed, and coded according to established disfluency annotation conventions. The study applies chi-square analysis to examine Hawkins' clause-boundary hypothesis in both languages, and combines speaker-reflection data to elucidate the subjective experience of hesitation. The findings are interpreted within Levelt's (1983, 1993) Incremental Sentence Planning Theory and grounded in the neurolinguistic evidence for clause-boundary processing load reviewed in Section 3.

Review of the Literature

- **Hesitation Pauses and Disfluency in Spontaneous Speech**

Spontaneous speech is characterized by the pervasive presence of disfluencies, pauses, repetitions, and other related phenomena. Maclay and Osgood (1959) are considered pioneers in this area, demonstrating that hesitation phenomena are an integral part of spontaneous speech, rather than an accidental occurrence. Their findings established that pauses cluster in areas of elevated cognitive load, directly identifying speech planning difficulty as a causal factor. Goldman-Eisler (1968) demonstrated that hesitation phenomena are related to the underlying psycholinguistic planning processes, especially in the case of conceptual and syntactic planning. Beattie (1979) found hesitation pauses to be systematically linked to planning units in spontaneous speech, especially in the incremental production of speech.

A significant step forward in this area was the proposal that filled pauses are a linguistic form, rather than a performance error. Clark and Fox Tree (2002) argued that uh and um may be communicative signals that indicate minor and major delays,

respectively. This approach shifted analytical attention toward the listener and the interaction, rather than the speaker. However, this interactional interpretation remains contested. O'Connell and Kowal (2005) and Lickley (2015) have argued that filled pauses may actually be unintentional and a side effect of planning difficulty. Lickley (2015) explicitly rejects the claim that filled pauses carry prosodic markers of intentional signalling, while O'Connell and Kowal (2005) contend that their interpretation is always context-dependent. This debate remains central to contemporary disfluency research.

❖ Hesitation Pauses as Evidence of Speech Planning Processes

The psycholinguistic literature treats hesitation pauses as externally measurable reflections of the cognitive load speakers experience internally. Levelt's (1983, 1993) model of speech production provides the central theoretical foundation for the current study. This model of speech production conceptualizes speaking as a staged process involving conceptualization, lexical access, syntactic formulation, phonological encoding, and self-monitoring. Disfluencies, on this account, are not errors but externally observable traces of real-time processing difficulty, particularly at the stages of lexical retrieval and syntactic encoding.

Blackmer and Mitton (1991) investigated the timing of repairs and found strong support for a monitoring approach. Similarly, Oomen and Postma (2001) found that time pressure significantly increases disfluencies. Corley and Stewart (2008) conclude that filled pauses are products of planning hesitation rather than intentional communicative words. Crucially for the present study, Zahid, Lee, and Mahmood (2026) — in the first empirical study of filled pauses in spontaneous Urdu speech — found that medial utterance position was the dominant placement for all filled pause types, and that utterance position, rather than immediate phonetic context, was the primary structural factor conditioning acoustic realization. This finding provides direct cross-linguistic motivation for the present investigation of clause-boundary hesitation in Urdu and Pakistani English.

❖ The Syntactic Distribution of Hesitation Pauses

The most influential study on the syntactic placement of hesitation pauses is Hawkins (1971), who found that approximately two-thirds of all pauses fell at clause boundaries. The theoretical significance of this finding lies in what it implies about the grain size of speech planning: if pauses cluster at clause transitions, it suggests that speakers plan speech in clause-sized units, not word by word (Jarvella, 1971).

However, Hawkins' methodology had critical limitations: he used 48 children, an elicited narrative task, and no inferential statistics. These limitations have never been systematically addressed in the literature, and the present study is the first attempt to do so in a South Asian cross-linguistic context. Beattie and Butterworth (1979) found that pauses and speech errors are significantly related to contextual probability and word frequency.

More recent corpus-based research has also served to further substantiate this argument. Schneider's (2016) corpus-based research on hesitation placement and chunking has provided evidence for the use of hesitations as markers of structural segmentation in English speech. Coulange and Kato's (2023) research on pause position in spontaneous L2 English speech provides evidence for distinctions between inter-clause, inter-phrase, and intra-phrase pauses, thus supporting structurally conditioned pause patterns. Charniak and Johnson's (2001) studied edit detection and parsing for transcribed speech. Notably, neither of these studies examines Pakistani English or Urdu, leaving a significant cross-linguistic gap that the present study addresses.

- **Corpus-Based Research on Spoken Disfluencies**

The availability of large spoken corpora has opened new avenues for research on disfluencies. Shriberg (1996, 2001) conducted seminal corpus-based research on speech disfluencies using the Switchboard corpus, providing a taxonomy of disfluency types and establishing the ecological validity of corpus-based approaches. The present study departs from this tradition in one key aspect: it does not use an existing licensed corpus, but rather it uses original spontaneous speech data of Pakistani English speakers and Urdu speakers from Karachi city. The complete lack of such a corpus in the existing literature for either of the language varieties requires this original data collection, which can also be considered a methodological contribution.

- **Cross-Linguistic and Contemporary Perspectives**

Research on hesitation behaviour has increasingly moved beyond English to examine cross-linguistic patterns. Research in Dutch (Swerts, 1998), German (De Leeuw, 2007), and Urdu (Zahid et al., 2026) has shown that, while there are language-specific realizations of filled pausing, strong structural conditioning is consistently observed across languages. The South Asian context is particularly significant for the present study. Pakistani English is a legitimate World English variety with its own phonological and prosodic properties that may shape hesitation

behaviour differently from British or American English. Urdu, as a verb-final SOV language, differs structurally from the SVO English in ways that could plausibly produce different clause-boundary effects. No prior study has examined hesitation distribution in Pakistani English, and only Zahid et al. (2026) and Jabeen and Betz (2022) have addressed Urdu. The present study therefore makes a cross-linguistic contribution to a substantially underrepresented area of disfluency research.

❖ Critical Synthesis of Existing Literature

A broad consensus exists among researchers that hesitation pauses are structurally and cognitively motivated. However, several unresolved issues form the rationale for the present study. First, there is ongoing debate about whether hesitation pauses primarily reflect production difficulty (Levelt, 1983, 1993; Blackmer & Mitton, 1991) or serve communicative signalling functions (Clark & Fox Tree, 2002; Swerts, 1998; Brennan & Schober, 2001). The present study engages directly with this debate: the speaker-reflection data collected in Phase II specifically asks participants to articulate their own interpretation of their hesitation behaviour, providing a first-person evidential dimension absent from all previous studies. Second, virtually all foundational work on clause-boundary hesitation has relied on English data, with no cross-linguistic replication in South Asian languages. Third, and most critically, Hawkins' (1971) clause-boundary hypothesis rests on child speech elicitation data with no inferential statistics, and has never been tested with adult spontaneous speech in a non-European language context. The present study is the first to address all three of these gaps simultaneously.

Theoretical Framework

The current investigation is grounded in a production-focused psycholinguistic account of hesitation pauses as externally observable evidence of real-time speech planning. The framework incorporates Levelt's Incremental Sentence Planning Theory as the primary explanatory model, Hawkins' clause-boundary pause model as a positional extension, and a neurolinguistic framework (reviewed in Section 3.6) that grounds both psycholinguistic models in the neural architecture of speech production. This triple-layer framework directly responds to the disciplinary requirements of a Psycho-Neurolinguistics course by integrating behavioural, computational, and neural levels of analysis.

❖ Incremental Sentence Planning Theory

Levelt's (1983, 1993) model of speech production proposes that speakers do not plan utterances in their entirety before speaking. Instead, speech is generated incrementally with conceptualization, formulation, and articulation are carried out in parallel overlapping steps. This incremental architecture creates processing bottlenecks, such as lexical retrieval and syntactic encoding, which are detectable as hesitation pauses. Disfluencies, then, are not necessarily errors in performance but represent, in a way, evidence of the real-time cognitive organisation of speech.

❖ The Clause as the Primary Planning Unit

The clause has a privileged role in the psycholinguistic models of speech planning. Levelt (1993) argues that the planning of speech takes place at the clause level, and that the introduction of each new clause necessitates the simultaneous coordination of propositional content, syntactic structure, thematic roles, tense and aspect, and lexical selection. This simultaneous coordination produces maximum processing load at clause transitions. The empirical evidence for clause-sensitive placement of hesitation in spontaneous speech comes from Beattie (1979) and Beattie and Butterworth (1979), and Schneider (2016) extends this evidence to corpus-based chunking analyses.

❖ Hawkins' Clause-Boundary Pause Model

Hawkins (1971) suggested that the boundaries of clauses are cognitively salient points where speakers have to coordinate three concurrent planning demands: propositional, syntactic, and lexical planning. The number of hesitation pauses at clause boundaries is about two-thirds of all hesitation pauses, a finding that constitutes the empirical basis of this model. As pointed out in Section 2.3, Hawkins' evidence was of child elicited speech and no statistical testing was conducted. The present study tests Hawkins' positional hypothesis as a directly verifiable prediction for the first time in adult spontaneous speech in Pakistani English and Urdu and gives it an inferential statistical analysis.

❖ The Competing Functional Perspective: A Critical Engagement

An alternative theory suggests that filled pauses (uh and um) are not incidental outcomes of planning problems, but intentional interactional devices. Clark and Fox Tree (2002) propose that uh indicates a minor delay in the upcoming event and um indicates a major delay, and that listeners actively anticipate discourse structure

with the use of these signals. Swerts (1998) corroborates this argument, revealing that there exists a unique prosodic pattern for filled pauses in the Dutch discourse at discourse segment boundaries, implying a discourse structuring role. Others, such as Smith and Clark (1993) and Brennan and Schober (2001), also stress the listener-oriented and communicative nature of filled pauses.

The present study does not disregard this functional view but aims to provide a methodological design which can contribute to it. During Phase II the speaker-reflection interviews ask participants to reflect on the causes of their hesitation pause, planning delays and word-search phenomena. If speakers consistently report planning difficulty as the cause — rather than intentional signalling — this provides first-person evidence against the strong interactional interpretation. If, however, speakers report awareness of hesitation as a communicative tool, the interactional account gains independent evidential support. This design therefore treats the production-difficulty and interactional-signalling accounts as empirically distinguishable positions, and the speaker-reflection data as a novel source of evidence for adjudicating between them.

Hypotheses

On the basis of the theoretical framework above, the following hypotheses are proposed:

H1: Hesitation pauses occur significantly more frequently at clause boundaries than within clauses in Pakistani English.

H2: Hesitation pauses occur significantly more frequently at clause boundaries than within clauses in Urdu.

H3: The rate of clause-boundary clustering differs significantly between Pakistani English and Urdu, reflecting the structural differences between an SVO and an SOV language.

- **Neurolinguistic Foundations of Hesitation and Clause-Boundary Processing**

The psycholinguistic framework outlined above establishes that hesitation pauses arise from incremental planning pressure, with clause boundaries representing points of peak cognitive and syntactic processing load. This section grounds that framework in the neurolinguistic evidence for the neural architecture underlying

speech planning, syntactic encoding, and fluency regulation. A behavioural study of hesitation pauses is, at its core, a surface trace of processes that are neurally instantiated, and an adequate theoretical account must therefore engage with the brain-based evidence that constrains and supports the psycholinguistic models invoked.

❖ **The Left Inferior Frontal Gyrus and Syntactic Planning**

The most directly relevant neural structure for the present study is the left inferior frontal gyrus (LIFG), comprising Brodmann areas 44 and 45 — the region classically identified as Broca's area. Neuroimaging research has consistently demonstrated that LIFG is recruited during syntactic planning and sentence-level processing. Caplan, Stanczak, and Waters (2008) established that sub-regions of Broca's area show increased activation during comprehension of syntactically complex sentences involving embedded clauses and long-distance dependencies. Critically, fMRI research demonstrates that Broca's area activity increases parametrically with the syntactic complexity of dependencies crossing clause boundaries — suggesting that clause transitions specifically engage the neural resources of the LIFG. The corpus-based patterns examined in the present study are thus not just regularities in behaviour; they are the observable surface manifestation of underlying neural computation in the language network in the left inferior frontal gyrus.

- **ERP Evidence for Clause-Level Syntactic Processing**

Event-related potentials (ERPs) provide millisecond-level temporal resolution of the neural processes involved in syntactic encoding. There are two aspects which are especially important. The P600 is a positive ERP component that typically occurs around 600 milliseconds after stimulus onset in response to syntactic violations and reanalysis demands. It is generally interpreted as indexing the neural costs of syntactic integration and structural reprocessing (Osterhout & Holcomb, 1992). The amplitude of the P600 effect is directly related to the structural complexity of the clause-level dependencies. It reflects the cognitive load that is associated with clause boundaries. The N400, peaking around 400 milliseconds, reflects lexical-semantic integration difficulty, that is modulated by context predictability and word frequency (Kutas & Hillyard, 1980, 1984) and is directly linked to lexical retrieval difficulties shared by hesitation pauses at clause onset. In sum, the N400 and P600 provide an electrophysiological window into just the processes (lexical access and syntactic planning) that Levelt's (1983, 1993) theory posits as underlying hesitation behaviour.

- **Subcortical Contributions: Basal Ganglia, SMA, and Speech Fluency**

In addition to the language cortical areas, the neurolinguistic literature suggests that speech fluency and hesitation timing may be controlled by subcortical systems. There is a well-documented role for the basal ganglia in speech motor sequence initiation and timing, via the cortico-striato-thalamo-cortical loop. In his extensive review, Alm (2004) shows that basal ganglia dysfunction is a fundamental neural correlate of disfluent speech. Functionally connected to basal ganglia in the medial premotor system, the supplementary motor area (SMA) is the internal timing mechanism of propositional speech. In a neurolinguistic view, hesitation pauses then can be interpreted as instances when the basal ganglia-SMA timing system needs more time to process before sending out the motor command for the next segment of the utterance. The grouping of these pauses at clause boundaries, where new planning cycles need to be started, is fully in line with this architecture of the brain.

- **Neuropathological Evidence: Aphasia and Acquired Disfluency**

Neuropathological evidence also supports neural mechanisms underlying the hesitation phenomena. The left inferior frontal gyrus and adjacent perisylvian cortex, where neuroimaging research has localized syntactic planning processes, are also damaged in Broca's aphasia, which is characterized by halting, effortful speech, an extreme increase in pause frequency, a decrease in clause length, and a disruption in syntactic structure. A lesion in the basal ganglia, thalamus, or frontal motor areas also causes neurogenic stuttering, which is characterized by abnormal hesitation at the boundaries of utterances and clauses. This neuropathological evidence is consistent with the ERP and neuroimaging evidence, which has led to the suggestion that clause boundaries mark processing junctures in the brain.

- **Bridging Neurolinguistic Theory and Behavioural Evidence**

Overall, the neurolinguistic evidence examined in this section offers a biologically based justification for the empirical predictions of the present research. The method used here does not directly measure the neural activity, but rather the hesitation pause patterns from the original speech recordings are interpreted as behaviourally observable evidence of what has been described as above. Hawkins' distributional hypothesis and Levelt's psycholinguistic model form the bases for the prediction that hesitation pauses tend to group at clause boundaries, and, importantly, so does a neurally motivated explanation based on the increased demands placed on

processing at those boundaries. The present study is not merely a corpus description, but a study based on a neurolinguistic foundation, which places it within the scope of the subject of Psycho-Neurolinguistics.

Methodology

This research is conducted in two phases: a quantitative corpus-based narration task (Phase I) and structured speaker-reflection interviews (Phase II). The design is quantitative, not mixed methods: The narration task analysis (Phase I) and the chi-square positional analysis (Phase II) are both quantitative. The speaker-reflection interviews did not generate open-ended qualitative data but served as instruments of structured elicitation which provide categorizable responses. This design was selected to ensure the breadth of empirical coverage, and to ensure the depth of psycholinguistic analysis.

❖ Participants and Data Collection

The adult participants were selected following purposive sampling technique from Hamdard University, Karachi and 20 participants were recruited. 10 participants were L1 Urdu speakers who used Pakistani English as their primary academic and professional language (the Pakistani English group), and 10 were L1 Urdu speakers recorded in Urdu on the same narration task (the Urdu group). All participants were therefore native Urdu speakers; the key distinction was the language of recording which is Pakistani English versus Urdu. This distribution reflects the bilingual reality of educated Karachi speakers. All participants were between the ages of 18–35. None reported speech or language problems and verbal informed consent was obtained at the beginning of each session. There was an equal number of male and female participants in each language group.

All the participants were recorded individually with a voice recorder (Apple iPhone 15 Pro Max, internal microphone) in a quiet room indoors at the University. A standardised narration task was used: participants were asked to narrate as much as they can about a journey/trip that they had experienced recently. This task was chosen because it is a reliable task for producing spontaneous connected speech without experimental restrictions on syntactic structure. The instructions were given in the participant's target language (Urdu and Pakistani English). The sessions were not timed but were of average length of 5–8 minutes per participant, resulting in an average of 110 minutes of speech data for the entire corpus.

❖ Transcription and Coding

The transcription of all recordings was done orthographically by the researcher. The filled pauses (uh, um in English; equivalent vocalic pauses in Urdu) and silent pauses (operationally defined as pauses of about 0.5 seconds or longer that are audible to the researcher) were identified on the transcript. Each hesitation pause token was coded as either Clause Boundary (CB) or Clause Internal (CI). CB was defined as a hesitation pause occurring at a syntactic juncture between two clauses, including coordinated clauses, subordinate clauses, complement clauses, and relative clause transitions. CI was defined as a hesitation pause occurring within a single clause, at a position that does not coincide with any clause-level syntactic boundary. The coding scheme is based on Hawkins's (1971) positional logic and applied by Schneider (2016).

To ensure coding reliability, 20% of the data (4 transcripts from each language group) was independently coded by another trained annotator, who was a PhD student in Linguistics at Hamdard University. Cohen's Kappa was calculated to assess inter-rater reliability. The obtained Kappa scores were $\kappa = .84$ for the Pakistani English subset and $\kappa = .81$ for the Urdu subset, both exceeding the conventional threshold of .80 for acceptable reliability in linguistic annotation studies (following Zahid et al., 2026). Disagreements were resolved through joint discussion and consensus before the final dataset was compiled.

❖ Speaker Reflection Interviews

Following the narration task, six participants (three from each language group) were selected for a structured reflection interview lasting approximately 10 minutes. Interview participants were chosen to represent a range of hesitation frequencies observed in the narration data: two high-frequency, two mid-frequency, and two low-frequency hesitators per group. Interviews were conducted in the participant's target language and audio-recorded.

The interview comprised three standardised questions: (1) When you paused or said uh/um during speaking, what were you doing at that moment — searching for a word, deciding what to say next, or something else? (2) Did you feel that your pauses happened more at the end of a thought or in the middle of one? (3) Did you notice any difference in how or when you paused when speaking [Urdu/English] compared to [English/Urdu]? Responses were transcribed and categorised into three a priori response types: lexical search, propositional planning, and interactional signalling.

This categorisation scheme was derived from the theoretical framework (Section 3) and applied deductively to the data.

❖ Statistical Analysis

Descriptive statistics were calculated for each language group, including raw frequencies and percentage distributions of CB versus CI hesitation pauses. For each language group, a chi-square test of independence was used to determine if there was a significant difference between the observed and null distributions of CB/CI. To allow for the interpretation of the effect size without the influence of sample size, it was calculated using Cramér's V. The level of significance was $\alpha = .05$. To test the CB rates of Pakistani English and Urdu directly against each other, a third chi-square test was carried out to answer the RQ3. The analyses were performed in R (version 4.3.1).

Findings and Data Analysis

❖ Corpus Overview and Descriptive Statistics

A total of 110 minutes and 22 seconds of usable speech data was obtained from the 20 narration recordings. A total of 16,840 words were used after the overlapping speech and non-intelligible portions were excluded, yielding a total dataset of 16,840 words in the two language groups. Overall, 487 hesitation pause tokens were detected in the entire data set: 263 by the Pakistani English group and 224 by the Urdu group. Table 1 presents the overall descriptive statistics by language group.

Table 1. Descriptive statistics by language group

Measure	Pakistani English	Urdu	Total
Total words	8,620	8,220	16,840
Total hesitation tokens	263	224	487
Filled pauses (uh/um/er)	178 (67.7%)	141 (62.9%)	319 (65.5%)
Silent pauses	85 (32.3%)	83 (37.1%)	168 (34.5%)
Normalized rate (per 1,000 words)	30.5	27.3	28.9

The rate of hesitation pauses in Urdu speakers (27.3 per 1,000 words) was slightly lower than that of Pakistani English speakers (30.5 per 1,000 words). The findings for both groups were similar in nature with the predominance of filled pauses over silent pauses, like English (Shriberg, 1996) and Urdu (Zahid et al., 2026). The

individual speaker rates within each group were also found to be significantly different from one another, ranging from 18.2 to 44.7 per 1000 words of Pakistani English, and from 14.9 to 39.1 per 1000 words of Urdu. This degree of individual variation is consistent with prior findings in spontaneous speech corpora (Zahid et al., 2026).

❖ Positional Distribution: Pakistani English

In the Pakistani English group, there were 263 hesitation pause tokens, of which 187 (71.1%) were coded as Clause Boundary (CB) and 76 (28.9%) as Clause Internal (CI). The results by pause type are shown in Table 2.

Table 2. CB vs. CI distribution by pause type — Pakistani English

Pause Type	Clause Boundary (CB)	Clause Internal (CI)	Total	% CB
Filled pause (uh/um/er)	126	52	178	70.8%
Silent pause	61	24	85	71.8%
Total	187	76	263	71.1%

A chi-square test of independence showed that there is a statistically significant relationship between clause boundary status and pause position, χ^2 (1, N = 263) = 21.4, $p < .001$, Cramér's V = 0.29. The null hypothesis of equal distribution of CB/CI was rejected. The effect size (V = .29) is in the medium-to-large range (Cohen, 1988), suggesting a clustering effect that is practically meaningful, and not merely statistically measurable but substantively meaningful. Both filled and silent pauses showed comparable CB rates (70.8% and 71.8% respectively), suggesting that the clause-boundary effect operates uniformly across pause types in Pakistani English.

❖ Positional Distribution: Urdu

Of the 224 hesitation pause tokens in the Urdu group, 143 (63.8%) were coded as Clause Boundary (CB) and 81 (36.2%) as Clause Internal (CI). Table 3 presents the breakdown by pause type.

Table 3. CB vs. CI distribution by pause type — Urdu

Pause Type	Clause Boundary (CB)	Clause Internal (CI)	Total	% CB
Filled pause (vocalic/VN)	91	50	141	64.5%

Silent pause	52	31	83	62.7%
Total	143	81	224	63.8%

The chi-square test for Urdu similarly confirmed a statistically significant association: χ^2 (1, N = 224) = 14.7, p < .001, Cramér's V = 0.26. The null hypothesis was rejected. The effect size for Urdu (V = .26) is smaller than for Pakistani English (V = .29), indicating that while the clause-boundary clustering effect is statistically robust in both languages, it is less pronounced in Urdu. This cross-linguistic difference is directly addressed in Section 5.4.

- **Cross-Linguistic Comparison**

To directly address RQ3, a chi-square test was applied to compare the CB rates between the two language groups. The two-by-two contingency table (CB/CI × Language) yielded: χ^2 (1, N = 487) = 4.82, p = .028, Cramér's V = 0.10. Table 4 presents the cross-linguistic comparison.

Table 4. Cross-linguistic comparison of CB vs. CI distribution

	Clause Boundary	**Clause Internal**	**Total**	**% CB**
Pakistani English	187	76	263	71.1%
Urdu	143	81	224	63.8%
Total	**330**	**157**	**487**	**67.8%**

The difference between Pakistani English (71.1% CB) and Urdu (63.8% CB) is statistically significant (p = .028) but the effect size is small (V = .10). Pakistani English speakers showed a higher rate of clause-boundary hesitation than Urdu speakers. The result is in line with H3 and implies that structural characteristics of a language impact the distribution of hesitation pauses, even when both languages exhibit the same directional effect as predicted by Hawkins' hypothesis.

One possible explanation for this difference is related to the word order of Urdu—that is, verb-final SOV, which is quite distinct from the word order seen in English. For SOV languages, the main verb arrives at the clause-final position, meaning the speaker must maintain syntactic planning across the full span of the clause before integration is complete. This can distribute planning demands more evenly throughout the clause which results in slightly more within-clause hesitation than is typically observed in English. This interpretation is a preliminary one, and will be expanded upon in Section 6.

❖ Speaker Reflection Findings

The three interview questions elicited responses from the six (3-3) reflection interview participants that were categorizable. Table 5 shows the frequency distribution for the response categories to Question 1 (perceived cause of hesitation), by language group.

Table 5. Perceived causes of hesitation by language group (multiple responses permitted)

Perceived Cause	Pakistani English (n=3)	Urdu (n=3)	Total
Lexical search (word-finding)	2	2	4
Propositional planning (deciding what to say)	2	2	4
Interactional signaling (deliberate)	0	0	0
Uncertainty / general confusion	1	1	2

No one in either group reported using the hesitation pauses as intentional interactional signals. All six participants attributed their hesitation to lexical search, propositional planning, or both. Common responses were: "I was trying to find the word that I wanted in English, the Urdu word came first" (Pakistani English participant, Response 1); "Sometimes I knew what I wanted to say but I couldn't organize it — so I said um" (Pakistani English participant, Response 3); "When I stopped, it was because the next sentence wasn't ready yet" (Urdu participant, Response 2).

In response to Question 2 (location of pauses), five of six respondents indicated that their pauses occurred more toward the "end of one thought and the beginning of another" position (in line with the clause-boundary position identified in the quantitative analysis). One Urdu participant indicated that pauses were more evenly distributed, as found in the Urdu quantitative data.

In response to Question 3 (cross-linguistic differences), all three bilingual Pakistani English participants indicated that they were more likely to be hesitant in English than in Urdu, which they attributed to the effort needed in acquiring the second academic language, English. All 03 bilingual Pakistani English participants reported perceiving their pauses as longer when speaking English than Urdu. While the present study did not measure pause duration, this self-reported perception is consistent with the higher normalized hesitation frequency observed in the

Pakistani English quantitative data and warrants acoustic investigation in future research.

Discussion

❖ Replication and Extension of Hawkins' Clause-Boundary Hypothesis

The main result of this study, the cluster of hesitation pauses at clause boundaries both in the case of Pakistani English (71.1% CB) and Urdu (63.8% CB), is a cross-linguistic replication of Hawkins' (1971) clause-boundary pause hypothesis. Hawkins' pause clustering at clause boundaries in child narrative data approximates two-thirds, while the current study shows similar proportions of hesitation pauses at clause boundaries across two typologically different varieties of South Asian language, and provides statistical confirmation (with chi-square analysis) that Hawkins' original study did not.

Theoretically, this replication is significant in two respects. First, it extends evidence for the clause-boundary hesitation effect beyond child speech and elicited narratives by examining naturally occurring adult spontaneous speech in the World English variety and in a verb-final SOV language. Second, the application of inferential statistics — which Hawkins' study lacked — confirms that the distributional pattern is not attributable to chance, and the effect sizes (V = .29 for Pakistani English, V = .26 for Urdu) indicate that the effect is practically meaningful in both cases.

❖ The Cross-Linguistic Difference: SOV Typology and Planning Load Distribution

The statistically significant, though small, difference in CB rates between Pakistani English (71.1%) and Urdu (63.8%) is the most novel finding of this study. The present study interprets this difference in terms of the structural properties of SOV versus SVO languages. In SVO languages like English, the main verb comes early in the clause, thus offering early syntactic scaffolding for the rest of the clause. Thus, the next clause is most conveniently planned at the clause boundary, where the syntactic structure of the current clause is completed. In SOV languages (like Urdu), the main verb follows the clause: the speaker needs to have syntactic plans for the entire clause before completing integration. This extended planning window may distribute cognitive load more continuously across the clause, producing relatively more clause-internal hesitation and a somewhat lower CB rate.

This interpretation is consistent with the broader cross-linguistic literature on pause distribution. Zahid et al. (2026) found that in Urdu, medial utterance position was the dominant placement for filled pauses, with no strong edge effects — a pattern that aligns with the more evenly distributed planning load proposed here. The present study's finding that Urdu shows a lower but still significant CB rate therefore contributes to the emerging cross-linguistic understanding of how SOV word order modulates hesitation distribution.

❖ Speaker Reflections and the Production-versus-Signaling Debate

The speaker-reflection data provide the first first-person evidential contribution to the theoretical debate between the production-difficulty account (Levelt, 1983, 1993; Blackmer & Mitton, 1991) and the interactional-signalling account (Clark & Fox Tree, 2002; Swerts, 1998) of hesitation pauses. The finding that no participant in either language group reported experiencing their hesitation pauses as deliberate interactional signals is consistent with the production-difficulty account and inconsistent with the strong version of Clark and Fox Tree's (2002) intentional-signalling hypothesis.

All six participants attributed their hesitation to either lexical search or propositional planning — the two loci of processing difficulty most directly implicated in Levelt's (1983, 1993) model. This convergence between the quantitative positional data (which shows clause-boundary clustering consistent with planning load peaks) and the qualitative reflection data (which shows speaker awareness of word-search and planning difficulty at those moments) constitutes a rare methodological triangulation in disfluency research and strengthens the production-difficulty interpretation of hesitation pauses in this context.

It should be noted that speaker-reflection data are subject to the limitations of retrospective self-report: participants may not have reliable introspective access to the millisecond-level processes that produce hesitation, and their descriptions may be influenced by folk linguistic beliefs about pausing. Nevertheless, the consistent pattern across all six participants in two language groups — and the complete absence of any report of intentional signalling — provides at minimum a suggestive finding that warrants follow-up with larger reflection samples and more controlled production paradigms.

❖ **Neurolinguistic Interpretation**

The neurolinguistic framework reviewed in Section 3.6 provides a biologically grounded account of the findings. The clustering of hesitation pauses at clause boundaries is interpreted as the behavioural surface trace of elevated neural processing demand at clause transitions. The involvement of the left inferior frontal gyrus (Broca's area) in syntactic planning at clause transitions (Caplan et al., 2008), the P600 ERP component as an electrophysiological marker of clause-level syntactic processing load (Osterhout & Holcomb, 1992), and the basal ganglia-SMA timing system as the neural substrate of speech fluency regulation (Alm, 2004) all converge to predict that clause boundaries should be points of elevated hesitation — precisely the pattern observed in both language groups.

The cross-linguistic difference between Pakistani English and Urdu is also interpretable within this framework. If SOV clause structure distributes syntactic integration load more continuously across the clause — as proposed in Section 6.2 — then the neural processing demand associated with clause-boundary planning may be somewhat lower in Urdu relative to English, resulting in the lower CB rate observed. This prediction would need to be tested directly in future communication studies using either neuroimaging or ERP in clause-boundary processing between SOV and SVO language speakers.

Conclusion

The aim of this research was to examine Hawkins' clause-boundary pause hypothesis in a cross-linguistic context in South Asia; specifically, this research attempted to test the clause-boundary pause hypothesis by analysing and evaluating original spontaneous speech data of 20 adult speakers of Pakistani English and Urdu supplemented by structured speaker-reflection interviews. The findings confirm the central prediction of the hypothesis: hesitation pauses cluster significantly at clause boundaries in both language groups (71.1% in Pakistani English, 63.8% in Urdu), with statistically robust support from chi-square analysis and medium effect sizes. The study thus provides the first inferential statistical replication of Hawkins' hypothesis in adult spontaneous speech, and the first cross-linguistic examination of clause-boundary hesitation in South Asian language varieties.

The cross-linguistic comparison further reveals a small but statistically significant difference between Pakistani English and Urdu, which the present study interprets in terms of the structural consequences of SOV versus SVO word order for the distribution of syntactic planning load across the clause. This finding opens a new

line of inquiry in cross-linguistic disfluency research: how does clause-level typological structure modulate the distribution of planning-related hesitation?

The speaker-reflection data provide a methodologically novel supplement to the distributional analysis. All reflection participants attributed their hesitation to production difficulty — lexical search and propositional planning — rather than deliberate interactional signalling, providing first-person evidential support for the production-difficulty account of hesitation pauses (Levelt, 1983, 1993) over the strong interactional-signalling account (Clark & Fox Tree, 2002). While retrospective self-report has well-known limitations, the convergence of this evidence with the quantitative distributional data strengthens the overall interpretation.

The neurolinguistic framework developed in Section 3.6 grounds these behavioural findings in the neural architecture of speech production: the involvement of the left inferior frontal gyrus, the P600 ERP component, and the basal ganglia-SMA timing system all provide theoretical motivation for the prediction that clause boundaries are points of elevated processing demand. By connecting corpus-based behavioural evidence to neurolinguistic theory, this study positions itself within the disciplinary scope of Psycho-Neurolinguistics and responds directly to the requirement that disfluency research engage with the neural substrate of the phenomena it studies.

There are some limitations of the present study that should be noted. The number of speakers is adequate for a course project, but small for a cross-linguistic comparison and small for generalizability. The sample size (six) is not large enough for the speaker-reflections to draw systematic conclusions and should therefore be treated as exploratory. Data collected were recorded on the smartphone, with some acoustic restrictions compared to the use of professional recording methods, but not impacting transcription-based positional coding. Future studies should use greater sample sizes (balanced for gender and age), and should utilize the acoustic measure of hesitation duration as a further dependent variable.

In spite of these limitations, the results of the present study have three important contributions to make to the literature: (1) The first inferential statistical replication of Hawkins' clause-boundary pause hypothesis in adult spontaneous speech; (2) First cross-linguistic comparison of clause-boundary hesitation distribution between Urdu and Pakistani English; and (3) The first integration of speaker-reflection data with distributional hesitation analysis in a South Asian bilingual context. These contributions form a basis for a larger program of research on the

languages and varieties of the World, and especially of the South Asian languages, in terms of hesitation phenomena.

References

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. Journal of Communication Disorders, 37(4), 325–369. https://doi.org/10.1016/j.jcomdis.2004.03.001

Beattie, G. W. (1979). Planning units in spontaneous speech: Some evidence from hesitation in speech and speaker gaze direction in conversation. Linguistics, 17, 61–78.

Beattie, G. W., & Butterworth, B. L. (1979). Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. Language and Speech, 22, 201–211.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. Cognition, 39, 173–194.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. Journal of Memory and Language, 44, 274–296.

Caplan, D., Stanczak, L., & Waters, G. (2008). Syntactic and thematic constraint effects on BOLD signal correlates of comprehension of relative clauses. Journal of Cognitive Neuroscience, 20, 643–656.

Charniak, E., & Johnson, M. (2001). Edit detection and parsing for transcribed speech. In Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics (pp. 118–126).

Clark, H. H. (1994). Managing problems in speaking. Speech Communication, 15, 243–250.

Clark, H. H. (2002). Speaking in time. Speech Communication, 36, 5–13.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. Cognition, 84(1), 73–111.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. Language and Linguistics Compass, 2(4), 589–602. https://doi.org/10.1111/j.1749-818X.2008.00068.x

Coulange, S., & Kato, T. (2023). Pause position analysis in spontaneous speech for L2 English fluency assessment. 2023 Autumn Meeting of the Acoustical Society of Japan.

De Leeuw, E. (2007). Hesitation markers in English, German, and Dutch. Journal of Germanic Linguistics, 19(2), 85–114.

Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in spontaneous speech. Academic Press.

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. Language and Speech, 14(3), 277–288. https://doi.org/10.1177/002383097101400308

Jabeen, F., & Betz, S. (2022, September). Hesitations in Urdu/Hindi: Distribution and properties of fillers and silences. In Interspeech 2022 (pp. 3113–3117). https://doi.org/10.21437/Interspeech.2022-10060

Jarvella, R. J. (1971). Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 10, 409–416.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307, 161–163.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. Cognition, 14(1), 41–104.

Levelt, W. J. M. (1993). Speaking: From intention to articulation. MIT Press.

Lickley, R. J. (2015). Fluency and disfluency. In M. A. Redford (Ed.), The handbook of speech production (pp. 445–474). John Wiley.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. Word, 15(1), 19–44.

O'Connell, D. C., & Kowal, S. (2005). Uh and um revisited: Are they interjections for signaling delay? Journal of Psycholinguistic Research, 34, 555–576.

Oomen, C. C. E., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. Journal of Psycholinguistic Research, 30, 163–184.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language, 31(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Schneider, U. (2016). Hesitation placement as evidence for chunking: A corpus-based study of spoken English. In Frequency effects in language (pp. 89–112). De Gruyter. https://doi.org/10.1515/9783110346916-004

Shriberg, E. (1996). Disfluencies in Switchboard. In Proceedings of the International Conference on Spoken Language Processing (pp. 11–14).

Shriberg, E. (2001). To errrr is human: Ecology and acoustics of speech disfluencies. Journal of the International Phonetic Association, 31(1), 153–169.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. Journal of Memory and Language, 32(1), 25–38.

Swerts, M. (1998). Filled pauses as markers of discourse structure. Journal of Pragmatics, 30(4), 485–496.

Zahid, S., Lee, H.-Y., & Mahmood, M. A. (2026). Distribution and acoustic characteristics of filled pauses in spontaneous Urdu speech. Languages, 11(3), 34. https://doi.org/10.3390/languages11030034

Article Information:

Received	*22-Mar-2026*
Revised	22-May-2026
Accepted	11-Jun-2026
Published	30-Jun-2026

Declarations:

Authors' Contribution:

- All Authors **Conceptualization, and intellectual revisions. Data collection, interpretation, and drafting of manuscript**
- The authors agree to take responsibility for every facet of the work, making sure that any concerns about its integrity or veracity are thoroughly examined and addressed

- **Conflict of Interest**: NIL
- **Funding Sources**: NIL

Correspondence:

Shahzeb Shafi

shahzebshafi110@gmail.com