

Disfluencies In English And Uzbek Spontaneous Speech: A Comparative Psycholinguistic Analysis Of Filled Pauses, Repetitions, And Self-Repairs

Shodieva Gulsanam Arshiddin qizi

National University of Uzbekistan, PhD university, Uzbekistan

Received: 30 August 2025; **Accepted:** 25 September 2025; **Published:** 28 October 2025

Abstract: Background. Disfluencies (filled pauses, repetitions, self-repairs) provide a real-time window on speech planning and monitoring. Objective. To compare how speech rate (fast vs. slow) relates to disfluency types, positional distribution, and production stages in spontaneous English and Uzbek speech.

Methods. We analyzed ≈100 minutes per language from televised interviews (The Graham Norton Show; Darakchi.uz), ≈10 speakers per language (≥18 y.o.). Disfluencies were coded by type (filled pause, repetition, self-repair), position (initial/medial/final), and stage (conceptual planning, formulation/encoding, articulation, self-monitoring). A pause threshold of ≥200 ms was used; speech rate buckets (fast/slow) were assigned from observed WPM.

Results. English: 78 filled pauses, 36 repetitions, 3 self-repairs. Uzbek: 151 filled pauses, 13 repetitions, 7 self-repairs. Fast speakers produced more errors but fewer fillers; slow speakers produced fewer errors but more fillers. Disfluencies clustered utterance-medially in both languages. Gender patterns showed small asymmetries (e.g., more exchanges/perseverations among Uzbek males).

Conclusion. Disfluency profiles co-vary with speech rate and sociolinguistic style, reflecting universal psycholinguistic mechanisms modulated by community-specific communicative practices.

Keywords: Disfluency; filled pause; repetition; self-repair; speech errors; speech rate; psycholinguistics; English; Uzbek; self-monitoring.

Introduction: Psycholinguistics primarily aims to understand the complex relationship between human cognition and language, and most analyses in this field are conducted on natural speech. In this study, psycholinguistic analysis of speech samples drawn from interviews, spontaneous conversations, and free interactions constitutes an important component of our research. The first fundamental approaches to analyzing spoken language emerged in the 1950s–1960s, when psycholinguistics took shape as an independent discipline. Investigating errors in natural speech is one of the starting points of psycholinguistic analysis. The core theoretical principles were articulated by V. Fromkin in 1970. She argued that speech errors are not random; rather, they are systematic disruptions that occur during speech

production [1]. Large corpora of errors compiled by scholars such as David Fay and Anne Cutler, as well as V. Fromkin, brought notable innovation to the field of psycholinguistics [2]. They systematically distinguished phonological, morphological, and syntactic errors in spoken language and showed how these errors map onto particular stages of speech production.

In Psychology and Language, Clark and Clark similarly contend that disfluencies and speech errors arising in natural conversation serve as a valuable window into speakers' mental processes, especially planning and lexical selection.

METHODS

There are various methods for the psycholinguistic analysis of speech errors. One is Fromkin's observation-

and-recording method, in which hundreds of spontaneous errors were collected and classified into phonological, morphological, and syntactic types. In subsequent years, this method was extended by other scholars. For example, the large error corpus compiled by D. Fay and A. Cutler was oriented toward malapropisms and disruptions in lexical selection, thereby shedding light on the mechanisms of lexical choice and retention during speech production. The strength of this observation-and-recording approach is that it captures errors occurring in natural speech outside artificial settings, thus revealing how speech mechanisms operate under everyday conditions.

Disfluencies occupy a central place in natural speech activity. They are directly linked to a speaker's speech rate, individual style, communicative context, and the complexity of ongoing psycholinguistic processes.

Disfluencies appear as interruptions in the flow of speech and are manifested through pauses, repetitions, self-repairs, and fillers (e.g., uh, um, hhh). In this section, disfluencies observed in English and Uzbek speech are analyzed comparatively. The analysis draws on the following criteria:

Pauses — their duration (≥ 200 ms), position (utterance-initial, medial, final), and function; •

Repetition — are-use of a sound, syllable, or word; •

Repairs — immediate correction by the speaker upon detecting an error; • Fillers — units such as um, uh, hhh, mmm. For analysis, video recordings in English and Uzbek were selected as data sources. The English interviews were taken from The Graham Norton Show. The Uzbek interviews were selected from the Darakchi.uz YouTube channel. Each recording was approximately 100 minutes in length. Across both languages, about ten speakers (male and female), all over the age of 18, participated, and each spoke in their native language.

The aim is to determine, for English and Uzbek speakers, the frequency and distribution of disfluencies, as well as the speech-production stages at which they tend to arise. This section also highlights in which gender groups and among faster versus slower speakers' disfluencies are more frequently observed. Such an approach enables a comparative analysis of disfluencies not only from a linguistic perspective but also from psycholinguistic and social viewpoints.

Example 1 (10:56) Host: Qanaqasiga qattan oldiz buni? Maftuna: Chunki ichkaridagiga kuchiz yetsade tashqaridagiga farqi qolmaydi tasir qilmaydi Host: Demak, sizga oila oiladan ta'sir. oilada bo'lganmi bu muammo yani demoqchiman masalan nima qilasan bu narsani degan bir... This speech sample embodies both

a disfluency and a speech error from a psycholinguistic perspective. First, the unnecessary repetition of the word "oila" ("family") constitutes a disfluency (repetition). In addition, the confusion that arises during the selection of morphological forms—oila → oiladan → oilada—is evaluated as a speech error.

This error occurred at the stage of grammatical encoding, more specifically during morphological inflection. The speaker selected the lexical item oila at the conceptual level, but showed uncertainty in generating the appropriate grammatical form, resulting in the successive use of three different case forms. From the standpoint of syntactic structure, the error occurred at the beginning of the sentence.

Psycholinguistic research explains the relatively high frequency of errors in sentence-initial position by the fact that speech planning has only just begun and a fully specified structure has not yet been formed. By part of speech, the error involves a noun. The base lexeme oila should have appeared in the nominative case, but the speaker confused it with the ablative (oiladan, 'from the family') and the locative (oilada, 'in the family'). This is recorded as a morphological substitution error. The fragment also exhibits a self-repair: the speaker halts the utterance and begins to clarify ("that is, what I mean is..."). This illustrates how disfluencies and errors often co-occur and are interrelated within the dynamics of online speech production.

Example 2 (3:22) Graham Norton: "It's a it's a warm first time"

Speaker's gender and style: man (Graham Norton), speaks fast Stage of speech: This type of speech error pertains to the articulation and self-monitoring

Stages of speech production. Clark & Clark likewise classify such repetitions as performance errors. In this instance, there is a reparandum ("It's a...") and a repair ("it's a warm first time"), indicating active self-monitoring. Error location: The error occurs at the beginning of the sentence, on the pronoun plus the auxiliary verb to be, in a declarative clause.

Error type and identification: In this segment, the error type is repetition. Repetition signals a disruption in speech planning and typically arises at the stages of lexical selection or grammatical structuring. The speaker initiates the utterance but hesitates and attempts the start again.

Example 3 (3:29) Graham Norton: "Tom, have you recovered from your last appearance?"

Tom Hiddleston: "Uh, just about. Yeah".

Speaker's gender and style: male (Tom Hiddleston); he typically speaks fluently and quickly, with humor.

Stage of speech: In this exchange, Tom answers with

slight deliberation. The phrase “just about” signals uncertainty/approximation and functions as a filled pause arising at the formulation stage.

Error location: Sentence-initial position, realized as an initial adverbial (“just about”) before the propositional content. Error type and identification: This is a filled pause, i.e., a fluency error—a natural disfluency.

Example 4 (3:32) Graham Norton: “...have you told the others what happened the last time?”

Tom Hiddleston: I...I don't think I have.

Speaker's gender and style: male (Tom Hiddleston); he typically speaks fluently and rapidly, with humor. When measured, his speaking rate is generally 160–180 words per minute, with precise observations ranging 150–190 WPM.

Stage of speech: Formulation + articulation. Pause duration: 150 milliseconds (from 0:54 to 0:69).

Error location: Sentence-initial position, occurring on the pronoun within the speaker's reply.

Error type and identification: Repetition. After the first “I”, there are no filled pauses; instead, the speaker produces a second “I” separated by a 150 ms silent pause (i.e., an unfilled gap between the two tokens of I).

Example 4 (14:30)

Samuel: “So she decided we need to figure out a way to get men talking about, you know, cancer and let them know that, you know, a lot of men in this country die from cancer.”

Speaker's gender and style: male (Samuel L. Jackson), a fast-speaking guest.

Stage of speech: The first “you know” is used during the utterance to allow time for planning before firmly specifying the topic. The second “you know” both reinforces the idea and helps maintain the continuity of the speech flow. Before transitioning to the lemmas “cancer” and “a lot of men in this country,” he employs “you know” to bridge into the forthcoming material.

Error location: Mid-clause.

Error type and identification: This is a filled pause. While many speakers use “you know” automatically every two or three sentences, in Samuel L. Jackson's speech it appears only in certain stretches; here, its function is to ease planning before continuing.

Example 5 (17:43)

Graham: Have you seen the new Star Wars? Samuel: yeah Graham: okay All right, It would have been better with you in it. Yeah, that, that's really, yeah, that's that's that's the bottom line .

Speaker's gender and style: male (Samuel L. Jackson),

a fast-speaking guest. Stage of speech: The speaker is aiming to deliver the main takeaway (“the bottom line”) but experiences a brief delay in selecting the appropriate lexical item. To avoid breaking the flow, he repeats “that's” several times. No phonetic error is observed at the articulatory level—the words are pronounced correctly. The disruption arises prior to articulation, i.e., during the planning/formulation stage. Error location: Structurally, the repetition occurs utterance-medially (mid-clause). This supports the view that it is not a random pause but the result of hesitation in conceptual planning. The example occurs in a reply within a live, audience-facing talk-show setting, where self-monitoring appears active: the speaker tracks his own output and, at the moment of difficulty, continues via repetition rather than pausing.

Error type and identification: The speaker repeats “that's” three times in succession. This is a classic repetition, serving as a time-buying strategy to complete message formulation. In this instance, repetition also fulfills a communicative role—intensifying emotional tone, heightening dramatic effect, and focusing the audience's attention.

Example 6 (24:05) Tom: And I jump in the air and he dodges out the way and I bring the spear down and it wedges itself into the floor. And then I use the use it to spin myself around and kick him in the chest, which has been interpreted by fans as um a demonstration of my hidden talent at pole dancing.

Speaker's gender and style: male (Tom Hiddleston); he typically speaks fluently and quickly, with humor.

Stage of speech: This pattern indicates active self-monitoring. From a psycholinguistic perspective, the error arises at the phonological encoding stage or during pre-articulatory preparation—that is, the speaker momentarily delays transferring the intended phoneme from the mental plan into overt articulation. Error location: Both errors occur utterance-medially (i.e., in mid-clause).

Error type and identification: First case: The form “d he dodges” constitutes a false start followed by immediate self-repair. The speaker initiates the verb dodges with a brief blockage (an incomplete d) and corrects himself right away. Second case: Although it may superficially resemble repetition, this is not a classic repetition but a false start + syntactic repair. The speaker begins with “use the...,” then instantly restructures the clause and continues as “use it to spin myself.” Here, he likely intended a frame such as “use the [object],” but during planning the syntactic structure shifted, prompting an on-the-fly correction to “use it.”

Example 7 (10:57) Mahliyo Azamatova: Chunki man birinchi kontent qilishdan oldin ko'raman. Manga ko'p

anaqa uhhh g'oya ko'proq manga keladida, ko'raman yozib qo'yaman, erim aytadilar nimadur miyangga kelsa yozib qo'y deydilar

Stage of speech: It arises at the conceptual planning stage and serves functions such as buying time, maintaining fluency, and keeping the listener engaged. Location in the utterance: Mid-clause, occurring before a noun

Error type and identification: In spontaneous speech, lexical placeholders like “anaqa” (roughly “uh/like/that thing”) function here as fillers. In this case, the sequence “anaqa” + “uhhh” semantically “covers” the micro-pause, allowing the speaker to prepare to continue. “Anaqa” carries no propositional content but plays an important psycholinguistic role as a disfluency within the clause structure.

RESULTS

When comparing disfluencies in English and Uzbek, not only their frequency but also their dynamics relative to speech rate stand out. Observations show that fast speakers tend to produce more errors overall yet use fewer fillers; conversely, slow speakers make fewer errors but rely more on filled pauses. This pattern held for both English and Uzbek respondents. For example, among Uzbek speakers, Samandar Hamroqulov (5 pauses in a 10-minute segment) and Maftuna Arabbayeva (8 pauses in a 10-minute segment) exhibit fast styles yet relatively low pause frequency. Among English speakers, Sara Pascoe was noted as one of those with the fewest pauses. These results suggest that fast talkers generally avoid “padding” the stream with pauses—prioritizing speed at the cost of more errors—whereas slower speakers exert tighter control over the flow, yielding fewer errors but making filled pauses an integral feature of their delivery.

There is a direct relationship between speech rate and disfluencies: fast speakers make more errors but use filled pauses only rarely, whereas slow speakers reduce errors but increase the frequency of pauses. This contrast reflects different strategies of speech control and distinct implementations of underlying psycholinguistic processes. Based on spontaneous English and Uzbek speech, we conducted a comparative analysis of disfluency types, their frequency, and their distribution across speech-production stages and positional criteria. The findings highlight several key points.

Overall frequencies. In the English data, we recorded 78 filled pauses, 36 repetitions, and 3 self-repairs. In the Uzbek data, we observed 151 filled pauses, 13 repetitions, and 7 self-repairs. Thus, filled pauses were markedly less frequent in English, while they appeared at a much higher rate in Uzbek. This suggests that

Uzbek speakers—given their speech rate and style—tend to “buy time” more often by pausing in order to continue the utterance. Role of speech rate. Speech rate emerged as a key factor. Fast speakers produced more errors overall yet very few pauses; slow speakers produced fewer errors but relied more on filled pauses. For instance, among Uzbek respondents, Samandar Hamroqulov produced only 5 pauses and Maftuna Arabbayeva 8 pauses in 10-minute segments, while among the English respondents Sara Pascoe stood out as one of the speakers with the fewest pauses. These results corroborate a direct link between speech rate and disfluencies.

Positional distribution. Most disfluencies in both English and Uzbek occurred mid-clause. This accords with psycholinguistic models: the middle of an utterance is where speakers are actively selecting lexical items, forming grammatical structure, and engaging in phonological encoding. By contrast, errors at the beginning or end of clauses were considerably less frequent.

Gender differences. Gender also mattered. Among Uzbek male respondents, exchanges and perseverations occurred more often than among females. In the English data, errors were more evenly distributed, though repetitions and self-repairs were slightly more common among male speakers. Interacting factors. Disfluencies in English and Uzbek are intertwined not only with the linguistic system itself but also with speech rate, gender, personal style, and social factors. Fast speakers typically minimize pausing—prioritizing speed at the cost of more errors—while slow speakers constrain errors but integrate filled pauses as a regular feature of delivery. This pattern indicates that disfluencies rest on universal psycholinguistic mechanisms, while also reflecting speech culture and communicative strategies specific to each community.

REFERENCES

1. Clark, H. H., & Clark, E. V. (1977). *Psychology and Language*. Harcourt.
2. Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111.
3. Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, 37(3), 201–242.
4. Fromkin, V. A. (Ed.). (1973). *Speech Errors as Linguistic Evidence*. Mouton.
5. Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production. *Cognition*, 92(1–2), 101–144.

6. Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press.
7. Shriberg, E. (1996). Disfluencies in Switchboard. In *Proc. ICSLP* (pp. 11–14).