

LANGUAGE FAMILIARITY IN DIGITAL COMMUNICATION: A USAGE-BASED STUDY OF SOCIAL MEDIA SLANG COMPREHENSION AMONG L2 ENGLISH LEARNERS

Syeda Atifa Batool atifa.batool@nu.edu.pk	Lecturer, Department of Sciences and Humanities, FAST National University of Computer and Emerging Sciences, Karachi, Pakistan.
Kamran Ali dr.kamran@hamdard.edu.pk	Associate Professor, Department of English, Faculty of Social Science and Humanities, Hamdard University, Karachi, Pakistan.
Tahira Abdul Hakeem tahira@nu.edu.pk	Lecturer, Department of Sciences and Humanities, FAST National University of Computer and Emerging Sciences, Karachi, Pakistan.

Abstract

This study investigated the processing of digital lexical items by the L2 English learners in social media language (abbreviations, neologisms, and phrases) under the lens of usage-based theory and entrenchment hypothesis. The two variables in this study, a) familiarity with a word and b) how accurately the L2 learners comprehend its meaning, operationalize the entrenchment hypothesis in this study. Ninety-seven participants rated 45 digital lexical items on a five-point Likert scale on familiarity, and subsequently, 78 participants completed the comprehension accuracy task on the 22 items that were retained out of the initial 45 items, categorized as high-familiarity ($n = 14$) or low-familiarity ($n = 8$). A paired-samples t-test revealed significantly higher accuracy for high-familiarity items ($M = 84.3\%$) than the low-familiarity items ($M = 77.6\%$; $t(77) = 3.99$, $p < .001$, $d = 0.45$). Point-biserial correlations confirmed that within-learner familiarity positively predicted item-level comprehension accuracy ($r = 0.27$, $p < .001$), supporting both hypotheses. The findings of this research validate entrenchment as a mechanism underlying L2 comprehension of digital registers.

Keywords: *Digital Registers, Entrenchment, Familiarity, Comprehension Accuracy, L2 Learners, Usage-Based Theory.*

Corresponding Author: Syeda Atifa Batool (Lecturer, Department of Sciences and Humanities, FAST National University of Computer and Social Sciences, Karachi, Pakistan)

Email: atifa.batool@nu.edu.pk

1. Introduction

1.1. Background

The wide range of digital platforms have been a turning point for communication practices over the past few decades. Digital platforms such as Facebook, Instagram, X, TikTok, YouTube and Snapchat have introduced registers of language that are constantly gaining popularity among the users of different age group. In recent times, these digital communication channels have become one of the greatest sources of exposure to English language in authentic context for Pakistani L2 English learners of all age group, specifically youngsters.

For learners of English, the sources of language input can be categorized into a) formal channels such as academic or structured language teaching programs (structured sources) or b) informal channels which includes social media and other digital media platforms (unstructured sources). The language input from the latter comprises mostly of non-standard lexical forms with a unique system of expressions, developed by the members of these digital social communities to fulfill a wide range of communicative purposes, across a variety of contexts. Interestingly, these newly developed forms of expression are not peripheral anymore and are increasingly becoming a part of the everyday linguistic environment. More specifically, young adult L2 learners are frequently exposed to informal digital communication as L2 input, yet their comprehension may not be uniform across all digital lexical items. It can be assumed that, High-familiarity words in digital communication are correctly interpreted at greater rates, reflecting efficient lexical access through repeated exposure. Or low-familiarity words in digital communication are interpreted less accurately, reflecting the processing cost associated with unfamiliar, non-canonical forms.

In linguistic research, language processing has widely been studied for relatively stable lexical forms. However, language learners today are exposed to a highly unconventional, non-standard, and context-dependent language input. The digital lexicon has taken the central stage across the globe via the digital platforms, and it does not necessarily conform to the standard vocabulary of English language. This raises an important concern related not only to ‘how’ the L2 learners understand these forms, but also, how they process them.

This confirmatory study employs the Usage-based theory to understand whether the processing of social media slang (digital lexical forms) by young adult L2 learner, such as idk, ngl and FOMO, is efficient due to familiarity with the lexical item, or whether it delays comprehension as they deviate from the canonical forms.

1.2. Research Problem

Researches in Psycholinguistics have demonstrated that the factors of familiarity and frequency of exposure may lead to lexical access and improved comprehension (Ellis,

2002; Sedivy, 2020). However, most of these researches focus on standard lexicon and formulaic sequential expressions in context of formal language (Conklin & Schmitt, 2008; Siyanova-Chanturia et al., 2011). Research on non-standard, digitally mediated abbreviated forms, context-dependent, and increasingly being accepted in the L2 input due to wide exposure to social media and digital platforms, is scarce.

Drawing on Usage-based theory and the entrenchment hypothesis, this research examines whether familiarity with digital lexical items can predict the comprehension accuracy, at different levels of familiarity (high vs. low) and also at the participant level for individual digital lexical items. This dual focus paves the way for empirical evidence, not only in the context of studying the processing of familiar lexical items extended to non-standard digital registers in L2 contexts, but also understanding the consistency of familiarity-comprehension mechanism at the item level, providing a more direct approach to test entrenchment. In accordance to the explanation above, two hypotheses are proposed:

H1: L2 learners of English will demonstrate significantly higher comprehension accuracy for high-familiarity digital lexical items than for low-familiarity digital lexical items.

H2: Self-reported familiarity with individual digital lexical items will significantly and positively predict comprehension accuracy for those items among L2 learners of English.

This research aims:

1. To categorize the digital lexical items, specifically social media slang words, into high-familiarity and low-familiarity conditions based on the L2 learners' self-reported familiarity ratings.
2. To first measure and then compare comprehension accuracy between high and low familiarity conditions, among the L2 learners of English.
3. To examine whether individual familiarity with social media slang words can predict the comprehension accuracy at item level.

Researches in Psycholinguistic suggest that familiarity plays a central role in word recognition and language processing. According to Sedivy (2020, pp.571-588) in *Language in Mind*, word recognition depends on the rapid access to stored representations, and familiar forms are more efficiently processed compared to unfamiliar ones.

Sedivy (2020) also argue that lexical knowledge is developed by the exposure to form-meaning pairs. It implies that if learners are exposed to digital lexical items such as idk repeatedly, these forms may become integrated into their mental lexicon.

However, for digitally mediated L2 input, these assumptions have not yet been tested. Therefore, the current study aims to apply the existing psycholinguistic principle

of word recognition and processing to a form of language that is highly relevant but under-researched.

2. Literature Review

Research on formulaic language has consistently shown that frequently occurring sequences are processed more quickly than novel ones. For example, Conklin and Schmitt (2008) demonstrated that formulaic sequences are recognized faster than non-formulaic strings, suggesting that they are stored as chunks rather than processed word-by-word. Similarly, Siyanova-Chanturia et al. (2011) found that repeated exposure leads to more efficient processing of multiword expressions.

In L2 contexts, Ellis (2002) highlights the role of frequency in shaping acquisition and processing. Learners rely on repeated input to build representations, which in turn facilitates comprehension.

However, most of this work has focused on standard lexical items and multiword expressions in formal language. There is very little research examining the topic of digital lexical items and their processing in L2 learners, or their impact on processing difficulty.

The current study addresses this gap by examining whether frequency effects extend to non-standard, digitally mediated forms.

2.1. Digital Communication and Emergent Lexical Forms

Digital communication has transformed how the language is used in recent times, leading to newly emerging forms, acronyms, abbreviations, neologisms, emojis, spellings, and expressions. Social media along with instant messaging apps, and discussion forums has provided users the opportunity to produce a need-based language system that gives priority to communicative efficiency, speed, creativity, identity expression, and social group affiliation. Crystal (2001) states that users innovate language forms to suit the needs of digital age interactions in online environments, thus, the digital communication has created a unique linguistic environment that meets their need. These unique linguistic forms include shortened spellings, substituted phonetics, and changed grammatical markers, with a combination of multimodal expressions of text and symbols.

According to Herring (2013), online communication has unique pragmatic and sociolinguistic features. Digital communication is essentially different from traditional communication because it combines the features of both spoken and written language. Users have adapted their language to the platform they are communicating through, audience expectations, and communication goals. Such adjustments have led to the formation of new digital lexical items that clearly deviate from standard language conventions.

In the context of education, digital communication is increasingly associated with language learning and exposure. Kohnke (2023) have argued that social media environments can expose learners to authentic and dynamic language while such forms of

language are rarely encountered in traditional classroom materials. Consequently, in real world interactions, the L2 learners are faced with the task of interpreting the non-standard digital expressions in addition to standard academic English. This rather new dimension to language learning has expanded the scope of language processing research by introducing digital lexical forms as an increasingly important area of linguistic inquiry.

2.2. Usage-Based Theory and Entrenchment

Usage-Based Theory paves the foundation for studying how frequent exposure to certain forms of language can shape linguistic knowledge. This theory proposes that language is learning is directly related to its use and repeated exposure. The more a learner interacts with the target language, the better their mental representations become over time. This framework explains how frequency of exposure is central to lexical processing, comprehension, and consequently language acquisition.

The researches of Bybee (2010) and Tomasello (2003) highlight how frequent use solidifies linguistic systems. During language processing, frequently encountered forms become easier to access and establish stronger cognitive representations. Therefore, the strengthening of form-meaning linkages brought about by frequent exposure is referred to as entrenchment. Because they take less cognitive effort to retrieve, more frequent lexical elements are processed more quickly and naturally. This viewpoint holds that through repeated exposure to particular constructions and lexical elements; learners progressively build abstract language knowledge. As a result, prior exposure and usage frequency are reflected in the familiarity to language form.

Another aspect frequently discussed in SLA research is the effect of frequency language processing. Ellis (2002) demonstrated that lexical forms that learners frequently encountered are processed more efficiently than the infrequent forms. This highlights the fact that repeated exposure to language forms can automatize and strengthen the mental representations of the lexicon.

In addition, Siyanova-Chanturia et al. (2011) have discussed the interconnected nature of lexical processing. They emphasized how frequent exposure can organize lexical representations. In their work, they have supported the view that familiarity and exposure together can contribute directly to processing efficiency and comprehension accuracy.

2.3. Frequency Effects in Second Language Learners

Frequency plays a significant role in second language processing because L2 learners have less entrenched lexical representations compared to the native speakers. Exposure to L2 vocabulary and lexical forms may develop better comprehension, gradually. 2 comprehensions.

Schmitt (2000) in their work discusses the correlation between frequency of exposure to language forms with repeated contextual usage, and comprehension. Lexical

items that frequently appear in specific context are likely to be better recognized and understood by learners.

2.4. Familiarity, Digital Lexical Processing, and Contextual Comprehension

According to Psycholinguistics, learners can comprehend those linguistic forms better that they have greater exposure to. According to Sedivy (2020), lexical items that are familiar to the learners will be processed more effectively due to stronger memory retention. Similarly, Dilley and McAuley (2008) demonstrated how contextual signals aid in the interpretation of foreign language forms. These concepts are particularly relevant for digital communication, as students regularly come across internet slang, acronyms, and non-standard spellings. Grieve et al. (2018) pointed out that social media promotes lexical innovation and originality in their study.

The above discussion highlights importance of understanding Pakistani learners' exposure to digital lexical items and their comprehension in L2 context. This study addresses this gap under the lens of usage-based framework and entrenchment.

3. Theoretical and Conceptual Framework

3.1. Word recognition and lexical access

Word recognition is defined as the cognitive process of identifying a linguistic form and then retrieving the meaning from the long-term memory associated with it. As explained in Sedivy (2020), this process is linked with familiarity to the words, where language users access the frequent and well-established forms more quickly than unfamiliar ones. In this study, digital lexical items such as idk, FOMO, and Pookie will be treated as familiar lexical items if learners have sufficient exposure. Other researches (Harrington & Dennis, 2002; Sedivy, 2020) also argue that recognition is accurate and quicker for linguistic forms which the language learners are frequently exposed to.

For the second factor, frequency and usage-based learning, it is argued that lexical knowledge develops through repeated encounters with form–meaning pairings. Over time, learners build representations that allow for more efficient processing (Ellis, 2002). This aligns with usage-based perspectives, where frequency strengthens mental representations. This principle applied to recurrent form that a learner processes in a meaningful communicative context (Bybee 2010). If abbreviated digital forms are encountered frequently enough to generate stable representations, the frequency-facilitation relationship documented for standard words should be observable in comprehension accuracy for these forms as well.

High-frequency forms are expected to be more familiar and highly comprehensible, resulting greater accuracy (correct identification of meaning), and therefore less effortful to process, while low-frequency or novel forms may require more effort. The cognitive effort can be self reported by the participants.

4. Methodology

4.1. Research Design

This research has adopted a within-subjects quasi-experimental design, having administered two sessions to the participants on the same day, in different session approx. 4 hours apart. The familiarity rating task (FRT) and comprehension accuracy task (CAT) employed in two sessions were validated by two experts for the content and construct, and piloted on 30 participants.

In the first session, the participants completed the familiarity task, and in the second session, they attempted the comprehension accuracy task. The first session introduced a list of 45 social media digital slang words, rated by the participants using a bi-dimensional Likert scale. The familiarity

ratings were aggregated to categorize the words into high familiarity and low familiarity conditions. Then, in the second session, the same participants completed a comprehension task that included the categorized high and low familiarity words. The comprehension task from the second session resulted in accuracy scores on HF and LF items serving as the dependent variable.

4.2. Participants

The participants were initially 97 Pakistani L2 learners, aged between 17-22 years ($M=20.2$, $SD=1.8$), from a private sector university in Karachi. The participant eligibility criteria were as follows:

- Years of learning English formally (school/college/university): minimum 13 years of formal education in English as Medium of Instruction
- Exposure to digital media: active use of social media for 5 or more years
- Active social media accounts: Facebook, Instagram, X/Twitter, Snapchat, or YouTube
- Daily engagement with social media: minimum 3 hours per day across the social media platforms
- Primary language on social media: English as the dominant language for communication

After the completion of both sessions, the participants with incomplete data were excluded, with a final pool of 78 participants out of the 100.

4.3. Materials

Session 1: Familiarity Rating Task (FRT)

Stimulus Pool: Initially, forty-five words commonly found on social media and internet platforms (TikTok, Instagram, X, Snapchat, YouTube) were taken to make a list of slang words. The items on the list were morphologically diverse with a balanced inclusion of (a) abbreviations such as IDK, NGL, TFW; $n=15$), (b) neologisms (Ghosting, Lowkey, Rizz;

n=15), and (c) multiword phrases or idioms (No cap, it's giving, Caught in 4K, Touch grass; n=15).

4.3.1. Two-Dimensional Rating Scale

Participants rated the items, on a 5-point Likert-scale, in two dimensions:

Dimension A (Exposure Frequency): "how often the participant has seen or encountered the word in digital communication/social media" [1 = Never; 5 = See it constantly]

Dimension B (Meaning Confidence): "how confident the participant is that they know what the word means" [1 = No idea; 5 = Fully certain]

Dimension A is the primary independent variable, operationalizing the token frequency of participants language use. The familiarity rating from Dimension A were used to item categorization in to high familiarity and low familiarity. Whereas, Dimension B, was included for a supplementary analysis to investigate whether the exposure frequency and participant's confidence of knowing the meaning co-vary (Schmitt, 2000). This methodological choice was made to validate the data for entrenchment construct, which proposes that a frequent exposure should be associated with higher level of confidence in knowing the meaning of the word.

Item Categorization Procedure: The mean of the familiarity ratings from session 1 was computed for the 45 slang words. Subsequently the words were categorized as high and low familiarity based on the mean, with a pre-determined threshold of $A \geq 4.0$ for High-Familiarity (HF) and $A \leq 2.0$ for low familiarity (LF). Words with means between $A 2.1-3.9$ were to be excluded. However, after the initial data processing, a post-hoc fallback threshold of $A \geq 3.5$ for High-Familiarity (HF) and $A \leq 2.5$ for Low-Familiarity (LF) was followed, while words with mean between $2.6-3.4$ were excluded.

Applying the fallback threshold ($\geq 3.5 / \leq 2.5$) yielded 14 HF items with ($M_A = 3.96$, $SD = 0.32$) and 8 LF items ($M_A = 2.07$, $SD = 0.42$), which resulted in asymmetric distribution of familiarity data in the sample.

Session 2: Comprehension Accuracy Task (CAT)

Target Items: After data processing from the first session, a total of Twenty-two social media slang words were retained for the second session. The comprehension accuracy task (CAT) included 14 high frequency and 8 low frequency words. In the CAT, each word was presented in a sentence providing full semantic context of the word:

Example: "She walked into the room and just ate and left no crumbs— everyone was watching." Question: "What does ate and left no crumbs mean in this sentence?"

Options: (A) consumed food rapidly (B) spoke quietly (C) performed exceptionally well (D) ate hungrily and left no food on the plate

Similarly, four multiple choice options were provided for each word. All options were plausible as semantic distractor; coherent within the context of the carrier sentence but incorrect. This was meant to prevent the carrier sentences in aiding the contextual inference and to measure genuine knowledge of the meaning of target word.

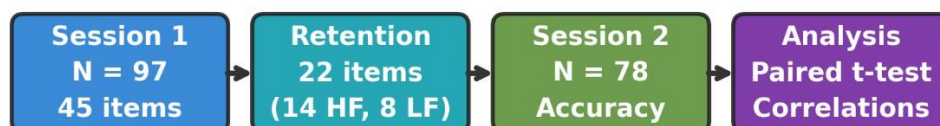


Fig 1-Participant progression and item retention through the two-session study. In Session 1, 97 participants rated 45 slang items on a 5-point familiarity scale; 22 items met inclusion criteria (mean ≥ 4.0 for high-familiarity or ≤ 2.0 for low-familiarity, yielding 14 HF items and 8 LF items). In Session 2, 78 participants (19 lost to attrition) completed a multiple-choice comprehension accuracy task on the 22 retained items. All 78 participants with complete data were included in the primary analysis.

Accuracy Scoring: The responses to CAT were used to calculate the comprehension accuracy scores, separately for HF and LF conditions:

$$\text{HF Accuracy (\%)} = (\text{Correct HF responses} / 14 \text{ HF items}) \times 100$$

$$\text{LF Accuracy (\%)} = (\text{Correct LF responses} / 8 \text{ LF items}) \times 100$$

As a result, the percentages of correct HF responses and the Percentages of LF responses formed the paired data for statistical testing.

4.4. Data Preparation and Cleaning

4.4.1. Familiarity Rating Data (Session 1)

Raw familiarity rating (1–5 Likert scale) means for each participant was computed. Only the words falling in the range of the fallback threshold were retained (2.6–3.4), and all others that fell in the mid-range were excluded from CAT and subsequent analyses, to preserve the desired categories of HF and LF, the primary experimental conditions. Afterwards, Dimension B (Meaning Confidence) ratings were extracted separately for the supplementary correlation analysis with Dimension A.

4.4.2. Comprehension Accuracy Data (Session 2)

The comprehension accuracy data was collected as binary responses (correct = 1, incorrect = 0) using the final pool of 22 slang words. The accuracy was aggregated for each participant per condition, yielding two accuracy percentages (HF% and LF%) per participant.

4.5. Data Matching

The data was carefully evaluated for both the sessions and only the complete responses were kept while incomplete data, or mismatched data from participants in both sessions was excluded. The final data that remained was N = 78.

5. Analysis

To test the hypothesis derived from Usage-based theory: that the more familiar L2 learners are with the non-standard social media slang words, the better the comprehension accuracy on those words, we have employed a multi-level analytical strategy which combines both aggregated and item-level analysis for maximum inferential power as well as theoretical precision.

5.1. Paired-Samples t-Test: Aggregate Accuracy Comparison

A paired t test was employed to compare the means of comprehension accuracy across two conditions: high familiarity and low familiarity, within the same participants. Since each of the 78 participants provided accuracy scores for each of the high familiarity ($n=14$) and low familiarity ($n=8$) conditions, the data is inherently paired. The within-subject correlation is properly accounted for using a paired t-test as it increases the within-subject statistical power by controlling the individual differences in baseline comprehension ability. The normality assumptions were met by the dependent variable (accuracy, ranging 0–100%) at the aggregate level ($N = 78$, sufficient for the robustness of Central Limit Theorem), and homogeneity of variance is acceptable (SD ratio = 1.35, which is well below the 2.0 threshold of concern).

For each participant, the percentage of correct responses on HF items and LF items was calculated yielding two values per participant. The paired t-test evaluates whether mean HF accuracy significantly differs from mean LF accuracy. The test statistic provides the mean difference between conditions and the standard deviation of the difference scores. At the degrees of freedom = $N - 1 = 77$, with two-tailed significance testing at $\alpha = .05$ was employed. For effect size, Cohen's d was calculated as $d=0.45$, which provided a scale-free measure of practical significance independent of sample size.

5.2. Point-Biserial Correlation: Entrenchment Hypothesis Test

The point-biserial correlation is a special case of Pearson's correlation, and it supports the cases when the variables types differ, with one variable being continuous and other being dichotomous. Since in our experiment, familiarity ratings (IV) is continuous and comprehension accuracy is dichotomous (correct responses = 1, incorrect responses = 0), we used point biserial. In order to test the entrenchment hypothesis, within learner relationship was quantified to understand how the familiarity of each item to a given learner reflects in how that learner answered the item correctly. The analysis on item-level, besides the aggregated t-test results, provide mechanistic evidence: does each learner show stronger accuracy on the item rated as more familiar? To study this association, point-biserial correlation is the standard measure. Hence, we computed the correlation for each participant, across the 22 items, the social media slang words. We then proceeded to aggregate the 71 valid correlations (where, data from 7 participants was excluded from the

initial 78, due to zero variance: perfect comprehension accuracy on all the items), and tested whether the mean entrenchment is significantly different from zero.

5.3. Stratified Entrenchment Analysis: Condition-Specific Effects

To test condition-specific effects, we employed Paired-samples t-test for stratified item on Fisher z-transformed point-biserial correlations. The point-biserial correlation of the aggregated means test overall entrenchment, and not on the individual participant level. Therefore, to understand the mechanism of how words already familiar to learners (HF) versus words less familiar or completely new to the learners (LF), employing the stratified data approach directly tests a theoretical nuance: does the magnitude of entrenchment depend on baseline familiarity? Hence, computing correlations separately for each condition (r of 14 HF words, and r of 8 LF words) per participant allowed a direct within-participant comparison. Using this approach enabled us to understand that if entrenchment is uniformly stronger in one condition, the paired difference will be statistically significant; if entrenchment is equivalent, there will be no significant difference. Consequently, for each participant, we computed two separate point-biserial correlation: (1) familiarity (HF items only, $n = 14$) $\times$ accuracy (HF items only), and (2) familiarity (LF items only, $n = 8$) $\times$ accuracy (LF items only), while excluding the data values for Participants with zero variance in either familiarity or accuracy for either condition (21 from HF analysis due to perfect accuracy; 21 from LF analysis for various reasons; 44 with valid data for paired comparison). Subsequently, the r values for both conditions Fisher z-transformed. A paired t-test (dependent samples) compared z value of HF versus z value of LF within the 44 matched pairs. This analysis maximized the statistical power for identifying the effects of each condition while preserving the within-subject structure.

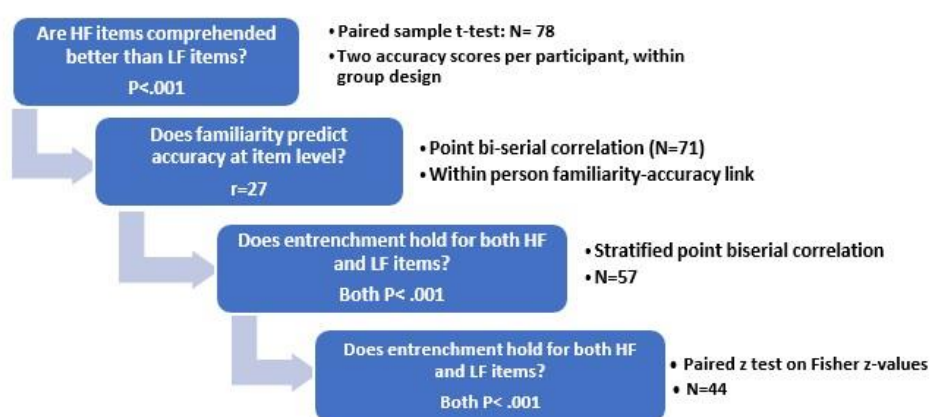


Fig 2-Statistical Workflow

6. Findings

6.1. Descriptive Statistics

In the second session (CAT) comprehension accuracy was substantially exceeded in high familiarity words ($M = 84.26\%$, $SD = 16.26\%$) compared to the low familiarity words ($M = 77.56\%$, $SD = 21.92\%$), across the 78 participants. The difference of 6.7 in percentage-point, combined with low variance of HF items, shows that high-familiarity words were more evenly comprehended as correct which points to a ceiling effect pattern typical of easier items. The Item-level descriptive statistics (means, SDs) for all 22 retained items is given in the table below.

Condition	N Items	Mean Familiarity	SD Familiarity	Mean Accuracy %	SD Accuracy %
HF	14	4.04	0.49	84.3	8.6
LF	8	1.84	0.38	77.6	9.9
Overall	22	3.22	1.26	81.6	9.3

Table 1: Descriptive States of familiarity ratings and comprehension by item condition.

Accuracy is expressed as percentage of correct responses

Paired-Samples t-Test: HF vs. LF Accuracy

Primary hypothesis (H1): A paired-samples t-test compared mean accuracy across the two item types. Results are presented in Figure 3.

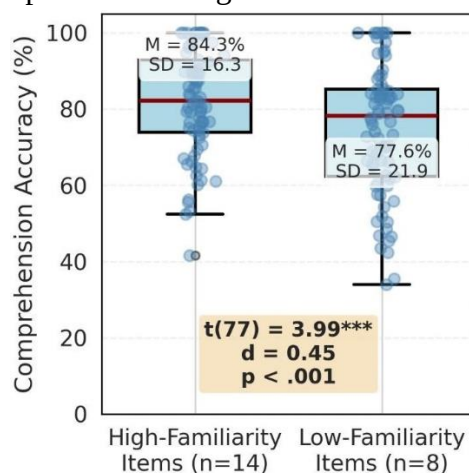


*Fig 3-Paired-samples t-test comparing comprehension accuracy for high-familiarity ($M = 84.3\%$, $SD = 16.3\%$) versus low-familiarity items ($M = 77.6\%$, $SD = 21.9\%$) across 78 L2 English learners. Individual participant data are overlaid as points. Error bars represent box plot quartiles and whiskers. *** $p < .001$, $d = 0.45$.*

The t-test yielded $t(77) = 3.99$, $p < .001$ (two-tailed), with 95% confidence interval around the mean difference of [3.35, 10.04] percentage points. Cohen's $d = 0.45$ (95% CI [0.22, 0.68]), that indicates a small-to-medium effect size. The effect may not be large, yet not negligible. This suggests while high familiarity may reliably improve the

comprehension, the magnitude of this improvement is subtle—learners are reasonably competent on low-familiarity items ($M = 77.6\%$), which indicates the absence of floor effect.

It supports the hypothesis; high-familiarity is associated with higher comprehension accuracy rather than the low-familiarity items. The positive effect size and a significant difference are consistent with the Usage-Based theory concept of entrenchment: exposure to a communicative form (operationalized via familiarity) enhances both familiarity and comprehension, even among L2 speakers.

6.1.1. Overall Point-Biserial Entrenchment

The computation of Point-biserial correlations for each participant, aggregated on the 22 total items, yielded a valid correlation on only 71 participants out of 78 (the remaining 7 were excluded due to perfect or zero accuracy on all items, which produces undefined correlations), as shown in figure 4.

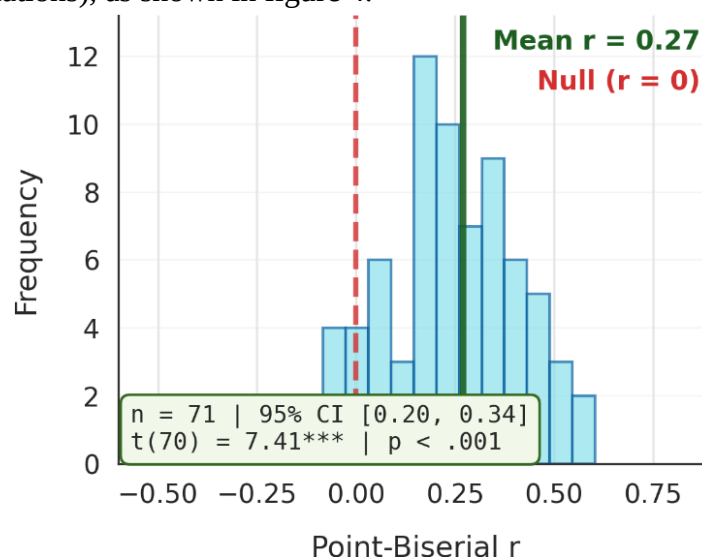


Fig 4-Distribution of point-biserial correlations between familiarity ratings and comprehension accuracy across 71 L2 learners (7 participants with zero variance excluded). The red dashed line indicates $r = 0$ (null correlation); the green line marks the sample mean ($r = 0.27$). Mean correlation with 95% CI [0.20, 0.34]. One-sample t -test: $t(70) = 7.41$, $p < .001$.

Mean Fisher $z = 0.2786$, corresponding to back-transformed mean $r = 0.27$ (95% CI [0.20, 0.34]). A one-sample t -test evaluating whether this correlation differs from zero yielded $t(70) = 7.41$, $p < .001$. This indicates that on an average learners show statistically significant relationship between their familiarity and comprehension accuracy on the social media slang words. This provides a substantial support to Usage based entrenchment hypothesis. Putting it simply, the items that were rated as more familiar, within the

individuals, were likely to be answered correctly. In the correlational context, mean $r = 0.27$ is not trivial as t indicates a more consistent, and meaningful association.

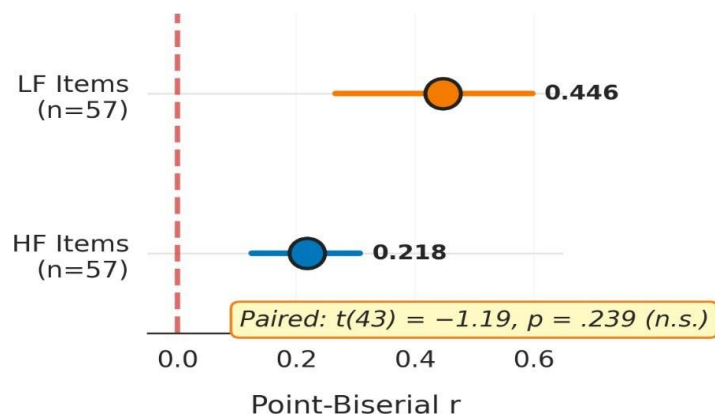


Fig 5- Stratified Entrenchment: Forest plot comparing point-biserial correlations computed separately for high-familiarity items (blue dot, $r = 0.218$, 95% CI [0.125, 0.307], $n = 57$) and low-familiarity items (orange dot, $r = 0.446$, 95% CI [0.266, 0.597], $n = 57$). Horizontal lines represent 95% confidence intervals. The red dashed line indicates $r = 0$. Paired t -test comparison ($n = 44$ matched pairs): $t(43) = -1.19, p = .239$ (not significant).

6.1.2. Paired Comparison

Finally, for the 44 participants found with valid correlation for both the HF and LF conditions, a paired sample t -test comparing familiarity and accuracy correlation, showed no significant difference in entrenchment between the high-familiarity items ($r = 0.28$) and the low-familiarity items ($r = 0.43$): $t(43) = -1.19, p = .239$. Even though, numerically low-familiarity entrenchment is 1.5× stronger, high within-subject variability hinders reliable detection of the effect of condition. This leads to the interpretation that entrenchment is uniform regardless of the difficulty level of words.

6.2. Summary of statistical workflow logic

A sequential battery of tests has been used to explore the alignment of results with the usage based theory.

1. A paired-samples t -test compared the accuracy of comprehension between the high and low conditions.
2. The point-biserial correlation was employed to quantify within-person entrenchment.
3. The stratified correlations were used to test whether entrenchment operates equally in both conditions.
4. A paired t -test was used on Fisher z -transformed correlations which compared the magnitude of entrenchment.

This hierarchical approach has systematically moved from measuring aggregate effects to providing mechanistic evidence for studying condition-specific patterns.

7. Discussion

The study tested the hypothesis derived from the Usage based theory on non-standard language items taken from the social media slang. The first hypothesis states that L2 learners can demonstrate higher level of comprehension accuracy for high-familiarity items compared to the low-familiarity items. The second hypothesis states that self-reported familiarity on individual items would predict the comprehension accuracy positively at the item level. Both hypotheses were supported. We discuss the findings, their theoretical implications, and their limitations in turn.

7.1. Confirmation of Hypotheses

Under the lens of Usage-based entrenchment hypothesis which states that greater exposure to a linguistic form has stronger cognitive representation ultimately enhancing comprehension is supported in both aggregate and item-level analyses. The difference in paired t-test ($HF > LF$, $d = 0.45$) provides substantial evidence of alignment with the usage-based theory (Langacker, 1987; Bybee, 2006). The point biserial correlation ($r = 0.27$) shows that the within L2 learner variation in familiarity is connected to the variation in comprehension. This establishes a critical link which the aggregate comparisons alone cannot, clearly showing the within-learner relationship i signature of entrenchment: cognitive representations are formed by the patterns of exposure to linguistic forms.

The stratified analysis further clarifies the entrenchment mechanism. Both HF items ($r = 0.22$) and LF items ($r = 0.45$) show significant entrenchment. The higher value of LF entrenchment, though statistically unreliable, is intriguing. The higher r value of LF items compared to HF items ($r_{LF} = 0.45$ vs. $r_{HF} = 0.22$) is interesting as it should theoretically be stronger for familiar, frequently-encountered words.

One explanation could be the presence of ceiling effects in HF items. Because the participants performed very well on 14 HF items, reducing variation in comprehension scores, their scores are near the maximum level. Correlations may appear to be weaker because there is less variation in the scores of HF and LF items. Contrary to this, the LF items show a greater variation in comprehension accuracy scores, allowing the familiarity scores of LF items to have stronger observable relationship between familiarity and comprehension accuracy.

A second explanation relates to how the items are processed cognitively. The LF items might have made the participants to consciously reflect on their previous exposure and familiarity when attempting the comprehension accuracy task. Whereas, the HF items might have been processed automatically, without a conscious effort to think and guess, due to stronger mental representation. This interpretation relates more closely to the theories of automatic and controlled processing in language comprehension.

Yet another explanation can come from the item selection criteria that was used in the study. HF items were selected because they were reported with already had high familiarity ratings, which reduced variability within the category. LF items could still have more meaningful differences in exposure and prior knowledge, consequently resulting in a stronger observed correlation.

However, the overall interpretation supports both the hypotheses. The results suggest that familiarity with the items has an influence on comprehension in both high and low familiarity conditions, while the numerical differences may be explained by the statistical and measurement factors.

8. Practical and Theoretical Implications

Pedagogical: The entrenchment effect is robust ($r = 0.27$, $p < .001$) and hence provides space for the exposure-based pedagogy. It can be argued that by increasing exposure to authentic, modern day non-standard communicative forms, or corpus-based exercises, L2 learners' comprehension for English language in authentic context can improve. The finding that entrenchment operates for LF items as well as HF items reveals that even a very brief encounter with new slang items may provide scaffolding for subsequent learning. Pedagogically, this implies that scaffolded, multiple exposures to digital communication norms can boost L2 learner competence beyond formal grammatical instruction.

Usage-Based Theory: The study has extended usage-based accounts of L2 learning to non-standard register of English language. Most research on entrenchment focuses on canonical morphosyntax or standard formal vocabulary. The current study shows that the principle of entrenchment can be extended to studying slang, which is exponentially taking over the GenZ and Gen Alpha communication, arguably making it a more peripheral, pragmatically-situated register. Furthermore, the findings also strengthen the idea that usage-based theory is a suitable framework to examine L2 mental lexicon, independent of register or linguistic domain.

9. Limitations and Scope

This study includes participants from L2 English context in Pakistan with moderate to high proficiency in English, hence the results may not be generalized to lower or higher language proficiency levels, or. Future research should examine whether entrenchment magnitude varies across different levels of proficiency.

The social media slang words were restricted to GenZ slang. Whether entrenchment patterns hold for other registers (e.g., technical jargon, regional dialects) can be explored in a sequential study since usage-based theory can widely be applied to explore entrenchment of different registers.

Participants self-reported their familiarity on a 5-point scale. Another approach would be to use more direct measures—such as exposure frequency via linguistic corpora

or neuroimaging indices to accurately trace the representation strength, this would strengthen inferences. Self-report familiarity may reflect a combination of other factors, such as perceived importance, emotional associations, beyond mere exposure.

10. Conclusions and Future Directions

This study has examined how familiarity with non-standard digital communication predicts L2 learner comprehension, for aggregate comparisons and within-learner item-level analysis. It has substantially supported the hypothesis driven from Usage-based theory. The comparison of magnitude across HF and LF conditions suggests that the mechanism of familiarity-comprehension not limited on baseline exposure or difficulty, which is a theoretically valuable finding. Future studies can: a) examine entrenchment for a broader range of registers, 2) incorporate neuroimaging or eye-tracking technology to precisely measure the cognitive mechanisms behind the familiarity-comprehension link, and 3) examine the individual learner differences in the magnitude of entrenchment to predict the overall L2 acquisition trajectories. These new dimensions of research will further deepen the understanding of how usage patterns impact L2 lexical and pragmatic knowledge.

References

- Brysbaert, M., Mander, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, 51(2), 467–479. <https://doi.org/10.3758/s13428-018-1077-9>
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language? *Applied Linguistics*, 29(1), 72–89.
- Crystal, David. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.
- Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, 59(3), 294–311. <https://doi.org/10.1016/j.jml.2008.06.006>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Grieve, J., Nini, A., & Guo, D. (2018). Mapping lexical innovation on American social media. *Journal of English Linguistics*, 46(4), 293–319. <https://doi.org/10.1177/0075424218793191>
- Herring, S. C., Stein, D., & Virtanen, T. (Eds.). (2013). *Pragmatics of computer-mediated communication* (2nd ed.). Walter de Gruyter.
- Howes, D. (1954). On the interpretation of word frequency as a variable affecting speed of recognition. *Journal of Experimental Psychology*, 48(2), 106–112. <https://doi.org/10.1037/h0059478>

- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). The effect of social media use on language learning: A meta-analysis. *Language Learning Journal*, 51(1), 1–18. <https://doi.org/10.1080/09571736.2023.2177894>
- Sedivy, J. (2020). *Language in mind: An introduction to psycholinguistics* (2nd ed.). Oxford University Press.
- Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase “time and again”: The role of frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 776–784.