

The Impact of Input Frequency on Chinese EFL Learners, Acceptability Judgments of English Passive Constructions

Yuxin Zhang^{1,*}

¹ Faculty of Arts, University of Melbourne, Melbourne, Australia

* Corresponding Author Email: yzhang2786@student.unimelb.edu.au

Abstract. Based on the Usage-Based Approach, this study employs the Acceptability Judgment Task (AJT) to explore the influence of input frequency on Chinese learners' acquisition of English passive structures. By researching 30 Chinese English learners' judgment on the "naturalness" of 20 target passive sentences, this study analysed that the high-frequency passive structures are generally high, with the score distribution concentrated in the 6-8 point range, indicating stable performance; while the scores for low-frequency structures fluctuate significantly, with some sentences having significantly lower scores, mostly concentrated in the 5-6 point range. The box plot and ranking chart further reveal that the median and interquartile range of the high-frequency group are higher than those of the low-frequency group, and the low-frequency group has more extreme low values. Based on the results, this research argues that frequency would greatly impact on second-language learners' acquisition of certain grammatical structures, especially English passive voice. This finding may also provide further support for the input frequency of diverse grammar for textbook compilation and language teaching.

Keywords: Passive Voice, EFL, Usage-Based Approach, Acceptability Judgment Task.

1. Introduction

Reviewing the Usage-Based Approach (UBA), which states that language acquisition is not driven by innate grammatical rules but is gradually formed through cognitive processes of language usage, one can more intuitively understand the process of second language acquisition. In other words, the more often learners encounter specific language structures in real contexts, the easier these structures will be internalised and acquired [1].

One of the biggest challenges for second language learners to internalise or acquire English, especially for native speakers of Chinese, is passive voice, given its grammatically complex feature. One of the main reasons, based on the UBA, can be that Chinese lacks a corresponding structure. With less exposure, Chinese learners may struggle with the progress of language transfer. Moreover, even if Chinese learners have mastered the passive voice, they may still avoid using this structure in actual language output [2, 3].

Currently, there is a considerable amount of research on the acquisition of passive voice in English. Yet one noticeable gap is that most focus on the output and practice, whilst few on how input impacts its acquisition. Hence, this research aims to study from an alternate perspective and examine Chinese learners' preferences through the Acceptability Judgment Task (AJT). The research results will provide empirical support for the role of input frequency effects in the acquisition of second language structures and offer references for teaching arrangements.

2. Literature Review

2.1. Usage-based Language Approach and Frequency Effect

In the process of language acquisition, learners gradually accumulate and integrate language patterns to form their mental lexicon [4, 5, 6]. Such repeated input and frequent exposure to certain language norms reinforce their learning outcome [7]. Whereas the mental lexicon is formed, the language learner's comprehension strengthens and output becomes more fluent. Hakimov and Backus further emphasised the role of frequency in language processing, as higher exposure frequency will

lead to stronger mental representations, regardless of the complexity of the structure [8]. Ellis et al. went further to propose that language "rules" are not innate but are patterns that learners gradually infer through long-term exposure to language input [1]. Therefore, input frequency becomes a key driving force in language acquisition, shaping the processing mode of the grammatical structure and the judgment of acceptability.

2.2. English and Chinese Passive Voice

If a comparative analysis is conducted on the corpus, it can be observed that there are differences in the usage frequency of passive structures between English and Chinese. Xiao's research indicates that in the British National Corpus, the frequency of English passive structures is approximately 1026 times per 100,000 words, while in the Lancaster Chinese Corpus, the frequency of Chinese passive structures is approximately 110 times [9]. Given this nearly tenfold difference, it can be inferred that influenced by the native language environment, Chinese English learners encounter fewer examples of passive structures in real contexts, which may hinder their acquisition of the passive voice.

Furthermore, differences in structure and typology also warrant further exploration. In English, the passive voice usually follows a fixed pattern, where the linking verb "be" is combined with the past participle. However, in the Chinese language, passive voice also greatly relies on lexical or contextual clues rather than fixed grammatical forms of just bèi. This typological difference makes the English passive voice lack an equivalent form in the native language, making it difficult to learn [2, 3]. Studies have shown that Chinese learners often choose active or agent-omitted structures to replace English passive sentences in non-adversity contexts, showing a clear tendency of "structural avoidance".

2.3. Empirical Studies of the Frequency Effect: Judgment Tasks and Production Tasks

2.3.1 Acceptability Judgment Task (AJT)

The Acceptability Judgment Task is widely used to explore the relationship between construction frequency and semantic judgment. Bybee and Eddington studied the collocation relationship between the verb "become" and adjectives in Spanish [1]. They found that high-frequency combinations were considered more natural and acceptable than other low-frequency combinations. Duan, Zhao, and Lum used AJT to explore the position of prepositions in English wh-relative clauses (pied-piping and stranding) [10]. The results showed that learners' acceptance of preposition-front structures was influenced by their proficiency. Their judgment performance was also affected by the phrase frequency and collocation strength.

2.3.2 Production Tasks (SCT and Gap-fill)

Production tasks are also frequently used in related research, such as sentence completion tasks (SCT) and gap-fill tasks. Domazetoska and Zhao employed SCT to examine the use of as-verbal structures in English [11], requiring learners to complete full sentences based on verb cues and contextual prompts to assess their structural production ability. They also used cloze tasks to test learners' vocabulary retrieval ability within specific structural frameworks and found that frequent structures were more likely to be produced correctly. Duan et al. also used SCT to observe learners' tendencies for preposition placement when dealing with wh- clauses [10], and the results showed that learners were highly sensitive to the frequency of collocations during the production process, especially high-level learners tended to produce more natural and frequent structures.

3. Methodology

This study employed the Acceptability Judgment Task (AJT) as the main experimental method to investigate how the frequency of English input affects the acceptability judgments of passive structures by Chinese learners. By systematically controlling the combination features and task-related variables, the study aimed to isolate the role of frequency in syntactic processing. The goal

was to reveal the internal mechanism by which input frequency shapes bilingual syntactic comprehension.

3.1. Participants

30 undergraduate students aged between 20 and 23 are included in the experiment. All of which are native Chinese speakers from a certain Chinese university, with more than 6 years of English education and have obtained an intermediate level of English literacy (CEFR B1-A2 or equivalent). All the participants were notified the process of the test and understood that this research would be completely anonymous. Written contents were also signed.

3.2. Experimental Materials

The research utilised an AJT questionnaire, consisting of 30 English sentences based on corpus collocation frequency. The sentences were structured as follows:

1. 20 target items: All were grammatically correct passive voice sentences, following the structure “be” (was/were) + past participle, with no “by + agent” constructions. Sentence length was controlled between 10 and 14 words.
2. 10 high-frequency collocations: Their passive forms appeared more than 4,000 times in the COCA corpus (raw frequency).
3. 10 low-frequency collocations: Their passive forms appeared fewer than 2,000 times in the COCA corpus.
4. 10 filler sentences: These were non-passive, active sentences with natural semantics and varied structures, designed to distract participants and minimise guesswork about the experiment’s purpose.

All collocations were sourced from the COCA corpus and categorised based on their actual co-occurrence frequency. To control for lexical familiarity, the team referenced CEFR vocabulary levels to exclude potential interference from vocabulary difficulty. The sentences were rewritten by the researcher from the original collocations to ensure structural consistency and semantic naturalness, and underwent multiple rounds of review.

3.3. Experimental Design

The experiment was implemented through an online questionnaire, published using the Qualtrics platform. Each participant is required to complete 30 single-sentence judgment tasks. Each task consists of an English sentence, and the score is given based on its “naturalness”. A Likert scale of 1-10 was used, with 1 being “very unnatural” and 10 being “very natural”. The order of all sentences was randomised to avoid the concentration of high or low frequencies. Two practice questions were used to familiarise subjects with the scoring process before the formal scoring (including one dysfunctional sentence to guide the scoring criteria). All the participants were asked to finish the tests in one sitting, with no chance to go back and make any further modifications to previous answers as well. Each individual’s test timing is expected to be within 5-8 minutes and the tests were finished on participants’ personal devices.

3.4. Data Analysis

This study conducted a statistical analysis on the AJT scoring data to investigate how input frequency affects the acquisition of English passive voice structures.

To achieve that, the questionnaire data were cleaned, and all incomplete, illogical or improperly formatted responses were excluded. Also after excluding 2 practice questions and 10 fill-in sentences (non-passive structures), 20 target passive sentences were left for analysis. Each target sentence adopted a uniform structure (“was/were + past participle”), with controlled length, and excluded the structure “by + actor” to reduce the effect of subject indication.

At the descriptive statistics level, the average scores of each participant in the high-frequency group (n=10) and the low-frequency group (n=10), as well as the average score and standard deviation

of each sentence, were calculated. Box plots were drawn to show the distribution of scores of the two groups, highlighting the median, interquartile range and degree of dispersion, to visually present the differences between the high-frequency and low-frequency structures.

To further explore the relationship between frequency and acceptability, bar charts were drawn by grouping by frequency and sorting by score to display the average score of each passive sentence. At the same time, histograms were made to present the rounded average score distribution of the two groups (interval range: 4-8), clearly contrasting the trend of acceptability.

Considering possible confounding variables, such as individual differences in the scoring scale and semantic preferences, the study suggests that future attempts could be made to use other methods (such as reaction time analysis, production tasks or cross-group comparisons) to enhance the explanatory power of the frequency variable.

4. Results

As can be seen from Figure 1, the average acceptability scores of the 20 passive sentences are presented. The frequent sentences are marked in yellow, while the infrequent sentences are marked in orange. The top 6 sentences with the highest scores are all of the frequent passive structures. Although some infrequent sentences (such as "All reports were finished before the deadline") also achieved relatively high single-sentence scores, the distribution of infrequent words is generally more scattered, and in the lower-score ranges, the scores are usually lower. While the mean difference in ratings between high- and low-frequency input groups does not reach conventional levels of statistical significance, the consistent directional trend across items provides robust qualitative evidence for a frequency effect, suggesting that input frequency meaningfully influences learners' metalinguistic judgments of grammatical acceptability.



Figure 1. Distribution of passive sentence acceptability ratings (ordered by frequency grouping)

The box-and-line plot in Figure 2 below is used to compare the overall acceptability score distributions of frequent passive sentences and infrequent passive sentences. The horizontal axis represents the two frequency groups, while the vertical axis shows the scores on a scale of 1 to 10.

The median of the high-frequency group is slightly higher, and the interquartile range (IQR) position is also higher. This indicates that the overall scores of these sentences tend to be more favourable. Although the dispersion of scores in both groups is similar, the extreme values in the low-frequency group are lower, meaning that the lowest-scoring sentences in the low-frequency group

have a lower score than those in the high-frequency group. This pattern further confirms the view that the higher the input frequency, the more positive the acceptance evaluation.

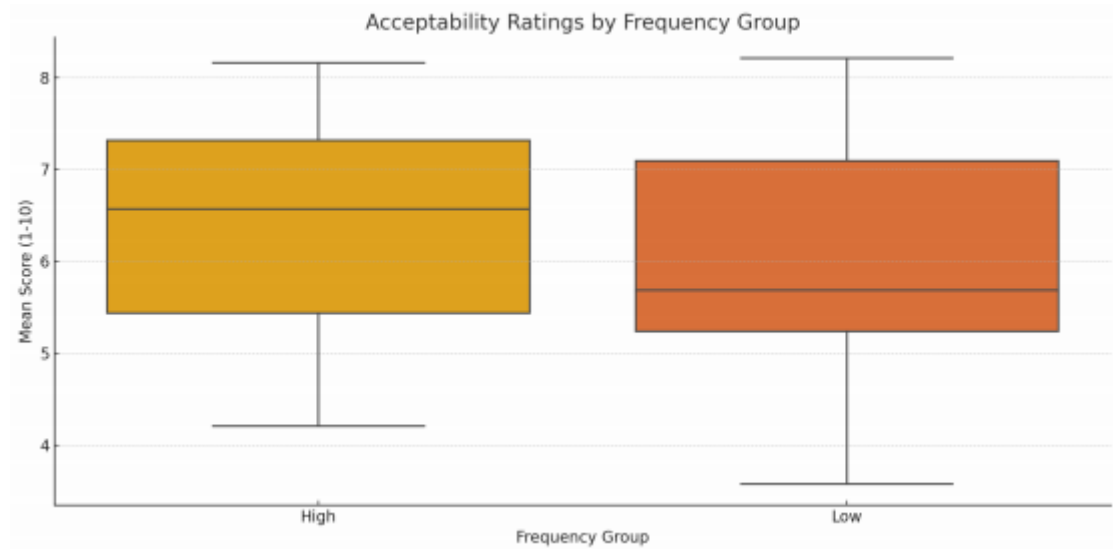


Figure 2. Boxplot of high–frequency vs. low–frequency passive sentence scores (Boxplot)

Figure 3 presents the average acceptability scores for all target passive sentences, ranked in descending order. High-frequency sentences are plotted in blue; low-frequency sentences, in orange.

The top-ranked items (i.e., those with the highest acceptability scores) consist predominantly of high-frequency sentences, whereas the bottom-ranked items (i.e., those with the lowest scores) are largely composed of low-frequency sentences. This clear gradient reflects the relative stability brought about by high-frequency input, while low-frequency input shows greater volatility.

The graph further facilitates the identification of individual sentences exhibiting markedly low acceptability, such as “The project was begun...”, enabling targeted discussion of their semantic or formal properties.

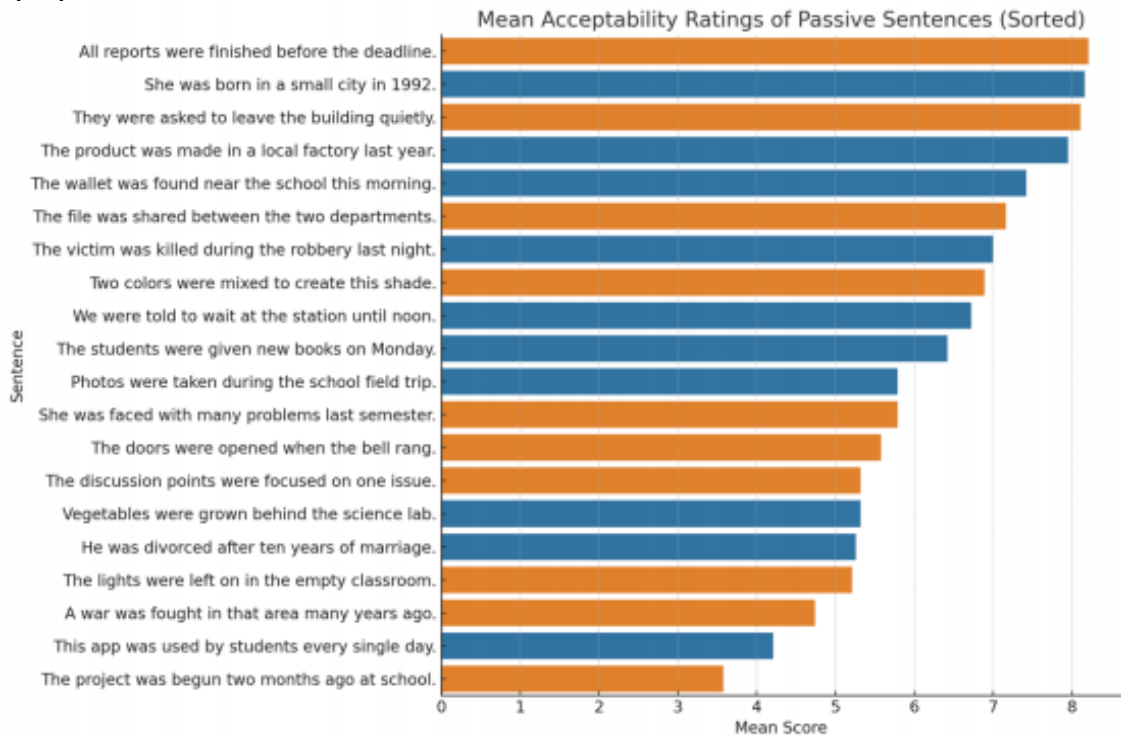


Figure 3. Ranking chart of average passive sentence scores (all sentences in descending order)

The distribution of passive sentence acceptability ratings is shown as a histogram in Figure 4. The y-axis shows how many phrases fall into each scoring bin, while the x-axis shows the rounded mean acceptability ratings, which range from 4 to 8. To identify high-frequency (blue) and low-frequency (orange) sentence clusters, bars are colour-coded.

High-frequency phrases are primarily centred in the 6–8 range, with a little positive tilt toward higher scores, according to the data. Low-frequency sentences, on the other hand, are distributed more evenly and widely over the 5–7 range, with a tiny percentage falling into the 4–5 range. The low-frequency group has no significant concentration in the highest-scoring intervals (i.e., above 6.5), despite the fact that both groups include precisely one phrase evaluated at 4.

This distribution supports a little but noticeable trend: learners prefer to see structures they encounter more frequently in input as more acceptable. The pattern indicates that exposure frequency influences acceptability assessments, even if the impact is not very strong.

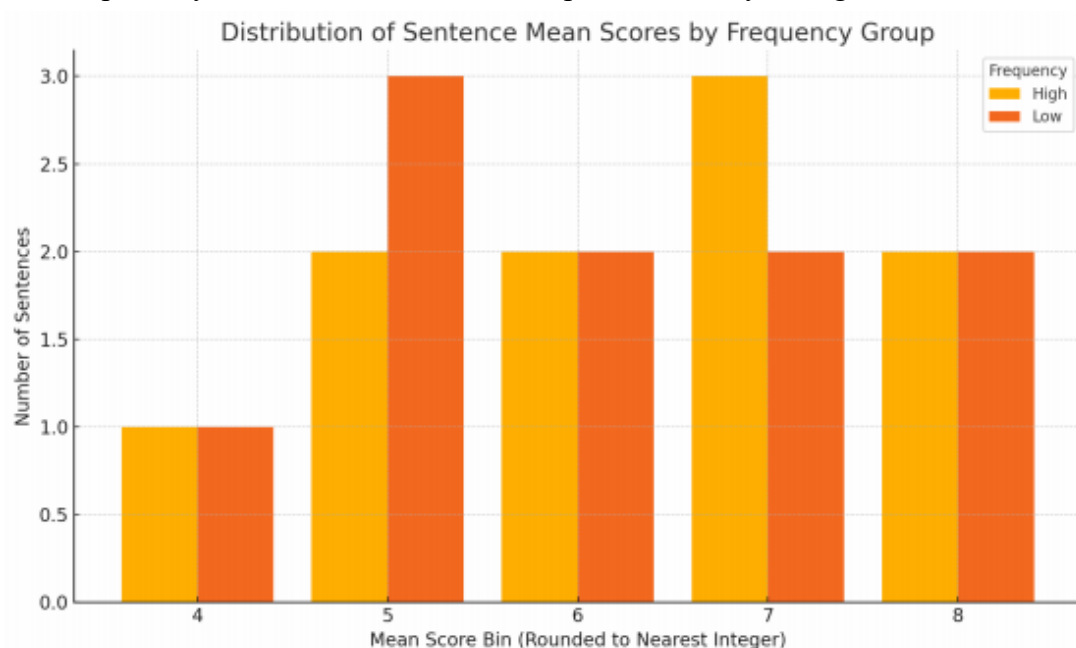


Figure 4. Distribution of the number of sentences in different score intervals (high vs. low frequency)

5. Conclusion

This study provides certain preliminary evidence for the input frequency effect proposed by the usage theory in second language acquisition. Although the sample size of the analysis was small, which limited the effectiveness of statistical tests, the observed data patterns still revealed some meaningful findings.

The descriptive results show that the acceptability scores for high-frequency passive structures (such as "was born", "was made") are generally higher, and the score ranges are more concentrated and more stable. In contrast, low-frequency structures (such as "began", "faced") have greater variability, and the scores of some sentences are significantly lower. This is consistent with the usage-based theory which suggests that higher input frequency strengthens mental representations, making the structures easier to be judged grammatically and evaluated for naturalness.

The further analysis in this paper reveals that the frequent sentences mainly fall within the 6-8 score range, while the less frequent sentences are mostly concentrated in the 5-6 score range, and some even lower. This trend indicates that frequency does influence people's judgment of the acceptability of specific structures. Additionally, some less frequent structures (such as "had divorced") are grammatically correct but may be negatively affected due to their relatively less semantic association. This suggests an interaction between input frequency and semantic familiarity.

Another interesting result from the data is that the individual difference is significant, especially for low-frequency structures. The reasons may relate to previous learning experience or teaching input. It shows that the frequency effect may differ accordingly. In order to further explore the underlying reasons and also the process of frequency impact on second language acquisition, it is recommended that future research can include more samples and include production activities or measure reaction times.

In conclusion, this paper offers certain perspectives on input frequency and grammar learning, also provides evidence-based data as a reference for future frequency-based teaching strategies and textbook compilation.

References

- [1] Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). Constructions and usage-based approaches to language acquisition. *Language Learning*, 66(1), 23–44. <https://doi.org/10.1111/lang.12150>
- [2] Li, S. (2014). An investigation into the influence of Chinese topic-prominent features on Chinese EFL learners' acquisition of passive voice. <https://doi.org/10.17863/CAM.1494>
- [3] Wang, Y., & Pongpaiboj, N. (2021). Avoidance of the English passive construction by L1 Chinese learners. *LEARN Journal: Language Education and Acquisition Research Network*, 14(1), 294–316. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/248692>
- [4] Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1–24. <https://doi.org/10.1093/applin/ami038>
- [5] Ellis, N. C. (2012). What can we count in language, and what counts in language acquisition, cognition, and use? In S. Th. Gries & D. Divjak (Eds.), *Frequency effects in language learning and processing* (pp. 7–34). De Gruyter. <https://doi.org/10.1515/9783110274059.7>
- [6] Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 1–24). John Benjamins. <https://doi.org/10.1075/sibil.48>
- [7] Bybee, J., & Eddington, D. (2006). A usage-based approach to Spanish verbs of "becoming." *Language*, 82(2), 323–355. <https://www.jstor.org/stable/4490159>
- [8] Hakimov, N., & Backus, A. (2021). Usage-based contact linguistics: Effects of frequency and similarity in language contact. *Journal of Language Contact*, 13(3), 459–481. <https://doi.org/10.1163/19552629-13030009>
- [9] Xiao, R. (2015). Contrastive corpus linguistic studies of English and Chinese: The cases of passive constructions and classifiers. *Journal of Chinese Linguistics Monograph Series*, 25, 82–114. <https://www.airitilibrary.com/Article/Detail?DocID=P20181128001-201512-201811280012-201811280012-82-114>
- [10] Duan, H., Zhao, H., & Lum, J. (2025). Native and nonnative speakers' preferences for preposition pied-piping versus stranding in English wh-relative clauses. *Language Learning*. <https://doi.org/10.1111/lang.12712>
- [11] Domazetoska, I., & Zhao, H. (2025). Distributional cues in construction acquisition: A comparative study of native and nonnative English speakers using the as-predicative construction. *Language Learning*. <https://doi.org/10.1111/lang.12710>