PanSIG Journal Vol. 10 - 2024
ISSN 2759-9965

How to cite the article:

Hart, E. (2024). Pitch frequency's impact on L2 listening comprehension. *PanSIG Journal*, *10*(1), 168–174. https://doi.org/10.37546/JALTPanSIGJ10.1-21

Research Article

Pitch Frequency's Impact on L2 Listening Comprehension

Edwin Hart

Fukui University of Technology

Abstract

While volume and pronunciation are often considered when discussing listening comprehension, vocal pitch frequency is a salient and often overlooked factor. Especially in schools where second language (L2) learners consist of students from different cultures with different native frequency pitches, vocal pitch frequency on the part of the instructor and the materials used may play a key part in the ease or difficulty with which students will have with comprehending material (Novia et al., 2018). Participants in this study were asked to listen to two speeches of approximately equal English level (CEFR B1), followed by a short comprehension test. The first speech was recorded and shifted to a lower pitch (118 Hz), while the second speech was recorded and shifted to a higher pitch (149.2 Hz). A pilot test was run with the same speaker with no pitch shift (139.8 Hz). Test results demonstrated a significant difference in performance depending on the frequency of the speaker. Implications to language instruction to internationally diverse L2 learners will be discussed.

リスニング理解について議論する際には、音量や発音が考慮されることが多いですが、音声ピッチ周波数は顕著でありながら見落とされがちな要素です。特に、第二言語（L2）学習者が、母国語の周波数ピッチが異なるさまざまな文化の生徒で構成される学校では、講師側の音声ピッチ周波数と使用される教材が、生徒が教材を理解する際の容易さや困難さに重要な役割を果たす可能性があります（Novia et al., 2018）。この研究の参加者は、ほぼ同じ英語レベル（CEFR　B1）の2つのスピーチを聞き、その後短い理解テストを受けました。最初のスピーチは録音されて低いピッチ（118 Hz）にシフトされ、2番目のスピーチは録音されて高いピッチ（149.2 Hz）にシフトされました。同じ話者でピッチシフトなし（139.8　Hz）のパイロットテストが実行されました。テスト結果では、話者の周波数に応じてパフォーマンスに大きな違いがあることが示されました。国際的に多様なL2学習者への言語指導への影響について説明します。

Listening comprehension tests are a mainstay of second language proficiency evaluations. Pitch can have an impact on listeners' comprehension and retention (Pourfannan et al., 2022). Certain studies have even indicated that the use of female voice overs often yielded better test results by L2 learners (Novia et al., 2018). This study aims to narrow down even further the more precise causative factor in the improvement of listening scores by checking if in fact the pitch frequency of the voice being heard is more of a contributing factor to intelligibility over the gender of the voice being spoken. The results of the current study will then be reported, which sought to address the following research questions (RQs): To what degree, if any, does vocal pitch frequency affect the L2 listening comprehension accuracy of Japanese tertiary students of English as a foreign language (EFL)? To what degree, if any, does L2 listening comprehension accuracy of Japanese tertiary students of EFL align with their perception of pitch?

Literature Review

Listening comprehension tests are a commonly used form of examination in standardized English testing such as TOEFL (Test of English as a Foreign Language) and TOEIC (Test of English for International Communication). As such, there have been studies examining the efficacy of different vocal qualities of the speakers in the exams and any correlation in test results. Listening comprehension (TOEFL) tests with women speakers conducted in an Indonesian university yielded high overall test scores in those who participated (Novia et al., 2018). On average, women have a higher vocal pitch frequency than men (Fitch et al., 1970), which lends some credence to high-pitched voices having an effect on listening comprehension tests. Conversely, lower-pitched voices lent themselves to higher accuracy of memory recall in earwitness testimony (by L1 English speakers) (Mullennix et al., 2009). In the case of text memory recall, L1 speakers listening to both higher and lower pitched voices yielded better memory recall than listening to a normally pitched voice (Helfrich et al., 2011). Higher levels of comprehension and retention in online lectures to L1 speakers (regardless of the lecturer's gender) tend to favor high-pitched voices (Samoza et al., 2015). While the effect of voice pitch on comprehension is well-established, there has been little study done in the context of an L2 learning environment.

The methods and findings in this study aim to focus primarily on the effects, if any, of vocal pitch frequency on L2 English students primarily from Japan.

Pitch has a unique position amongst other suprasegmental features of speech in that it often also affects

not only the listening comprehension of the listener, but also can contribute to other aspects as to how the speech is perceived. In designing robotic assistants, studies have found that both female and male voices modulated to a similar fundamental pitch frequency (225 Hz and 225 Hz respectively) resulted in better understanding and retention of information by human listeners (Pourfannan et al., 2022) clearly indicating a relation between pitch and memory retention. High-pitched voices can in one sense seem more friendly and childlike, or conversely can even seem more aggressive, which also have an effect on text memory (Helfrich 2004). As such, the use of pitch and its effect on test takers warrants important consideration when creating materials for L2 students.

Methodology

Participants

The current study was conducted at a small, private Japanese university. The student body primarily consists of students in technical fields such as sports sciences, electrical and mechanical engineering. While the university does not offer an English major, English is a required subject for all students in all four years of their study. Students typically take two English classes a week (90 minutes each). Courses in years one and two focus on listening and conversational skills, with later years lending more practice in standardized tests such as TOEIC. The university is located in a rural area of Japan, with little-to-no spontaneous interactions with L1 English speakers.

The English ability levels of students are evaluated annually with an abridged version of the TOEIC Bridge test. The results of these tests are also considered for students' class placements. This test is solely a listening comprehension test, with no speaking or writing components. The levels of students in the university follow the CEFR (Common European Framework of Reference for Languages) model: Advanced level (C1, C2), Intermediate level (B1, B2), and Beginner level (A1, A2), with the majority of students falling within the B and A levels of ability. The researcher informed the students before the test that for the purposes of his research the students would be asked to listen to two short speeches and be quizzed afterwards. The participants were also advised by the researcher that involvement in the research was completely voluntary, and that participation (or lack thereof) would have no bearing on or affect their current class grade in any way. They were also told that they could opt out of the study at any time.

In total, 189 students were involved in the study. However, only the data from participants who took part in both the Low-Pitched and High-Pitched tests were used for the pair comparison data analysis. The participants were divided into three groups, with names derived from the form of stimulus they received: Pilot Group (Group M, n = 49), Low-Pitched (Group LP, n = 110), and High-Pitched (Group HP, n = 110). Of the 110 participants, 89 were male (88 Japanese, 1 Vietnamese) and 21 were female (19 Japanese, 1 Vietnamese, 1 Thai).

Listening and Assessment Materials

Two original passages of approximate English level and length were created for this experiment by the researcher (Appendices A and C). The original passages were adapted from general information gathered from Wikipedia and adapted to match the ability range of the students in the university (CEFR B, A). The passages were regarding dragonflies and red lantern flies, respectively. These topics were carefully considered so as to eliminate the possibility of topic familiarity interfering with test results (e.g., Alexander et al., 1994; Rubin, 1992). Using a web-based vocabulary level analyzer based on the CEFR-J word list (cvla.langedu.jp), the contents of the passages were checked to be within the CEFR-B1.2 and B1.1 levels (see Table 1). This was done to ensure that the English used would not be too difficult and be within the ability for the students tested (CEFR B, A), with over 60% of the terms in each passage actually being in the A1 range. Some words which fell outside of the B1, 2 range included proper nouns and loanwords (e.g., dragonfly, nymph, North America). Both sets of the comprehension questions were designed so as to be able to reflect specific information directly from the text which was read. An endeavor was made to have similar specifics be asked about (the colors of the insects, number of wings, their eating habits).

Table 1

Vocabulary Level Analysis of Texts Used in Comprehension Tests

CEFR	Dragonfly Text		Spotted Lantern Fly Text	
	AvrDiff	*BperA*	*AvrDiff*	*BperA*
A1	1.27	0.06	1.27	0.06
A2	1.41	0.11	1.41	0.11
B1	1.50	0.13	1.50	0.13
B2	1.67	0.19	1.67	0.19
Input	1.58	0.18	1.50	0.16
Estimated Level	B1.2	B2.1	B1.1	B1.2

Note. As per cvla.langedu.jp, AvrDiff denotes average word difficulty with A1 = 1, A2 = 2, B1 = 3, and B2 = 4. BperA denotes the ratio of B level words to A level ones.

Three variants of the stimulus used for the test were prepared. A male speaker of L1 Standard American English (the researcher), first recorded audio of himself reciting Passage A (Appendix A). The recording was extracted and converted into a wav audio file for presentation with no pitch modification (139.8 Hz) to the pilot group. The same audio was modified, and pitch shifted two semitones down (118Hz) for use for the low pitch listening group.

A second recording was made of the researcher reciting the Passage B (Appendix C). This audio was modified, and pitch shifted two semitones up (149.2 Hz). These pitch frequencies were chosen in reference to stay within established boundaries of the frequency range for an adult male (Fitch et al., 1970) without sounding too unnatural or distorted. Other studies have found that pitch shifting the samples beyond 20 Hz in either direction can make the sample become noticeably artificial, unnatural and distracting (Pourfannan et al., 2022).

The audio was recorded using a Shure SM7B dynamic microphone connected via XLR to a Tascam Portapro X8. The audio files (wav) were recorded at a sample rate of 48 kHz and then pitch shifted up or down as needed in the open-source software Audacity. After modification and export, the samples were put through Praat to more precisely measure the average pitches of the audio samples.

Data Analysis

Two original 10-item quizzes were written by the researcher to correspond to each test to evaluate the participants' level of comprehension (Appendices B and D). To ensure quick and accurate comprehension of the quiz items, this outcome measure was created in the participants' L1 (Japanese) by the researcher and checked by a colleague with a Japanese Language Proficiency Test (JLPT) N1 proficiency level. The quiz was piloted with a sample population of 49 students from the same university. Inferential statistics (t-tests) were run to determine if statistical differences existed between the groups.

After participating in both the LP and HP sessions, students were asked to take part in a post-test impressions questionnaire (Appendix E) to gauge the perception of how they performed versus their actual test results.

Procedures

At the beginning of each listening test, the participants were given time to read over the 10 comprehension questions before the audio was played. While such reading of questions before listening begins is a procedure commonly done in TOEIC examinations, reading the questions beforehand considers working memory as an independent variable in L2 listening comprehension (e.g., Fay & Buchweitz, 2014; Linck et al., 2013; Wallace & Lee, 2020). By familiarizing themselves with the questions and answers beforehand, the participants could focus on listening, thereby reducing other variables such as comprehension. Students who took part in the tests were told that they would be listening to a short speech in English which would only be played once, and to look over the questions before the audio was played. Once all the participants indicated that they were ready, each group listened to the audio sample corresponding to their test.

The Pilot Group heard the audio of the speaker with no pitch shift (139.8 Hz), Group LP heard the audio of the low pitch shifted (118 Hz) speaker, and Group HP heard the audio of the high pitch shifted (149.2 Hz) speaker. After the audio was finished playing, participants were allowed to answer the quiz questions at their own pace without the restriction of a time limit.

After the LP and HP groups finished their tests, they were asked to answer an exit questionnaire regarding their perception on their own performance in each test as well as how they perceived the ease of understanding for each of the differently pitched audio clips.

Results

A dataset was created using the website langest.jp, which operated using the statistics program R v3.2.1, from which all statistical calculations were performed. Descriptive statistics were first calculated for each group, with means and standard deviations reported (Table 2). Although there were originally 149 participants in the study, only the results from participants who took part in both the LP and HP groups were used in the data analysis.

Table 2

Descriptive Statistics of Comprehension Quiz Scores

	Pilot Group		Group LP		Group HP	
	M	*SD*	*M*	*SD*	*M*	*SD*
Quiz Score	5.50	2.03	5.06	2.16	5.54	1.83

Note. The various group participant sizes were as follows: Pilot Group n = 49. Group LP, Group HP n = 110. The scoring scale was from 0 (minimum) to 10 (maximum).

A paired *t*-test of Group LP and Group HP was conducted to determine if a statistical difference was present between the groups. A significant statistical difference was found between the comprehension scores of the low-pitched

and high-pitched groups: $p = 0.007$, $d = 0.26$ [0.07 ~ 0.45]. If effect sizes were taken into account, the results could be made even more apparent.

The exit questionnaire the participants filled out regarding their self-perceived performances on the tests seem to reflect these results, with 82% of students asked saying that they felt that the higher pitched voice test was easier to understand. Despite all audio recordings being done by the same male L1 speaker of English with the only digital audio augmentation being pitch shifting up and down by two semitones, some students erroneously associated the higher pitched voice as a female voice.

Discussion

The current study sought to investigate whether the pitch frequency of a speaker would have an impact on listening comprehension test results, with results demonstrating that tests conducted with the high-pitched (149.2 Hz) voice yielded statistically significant higher scores ($p = 0.007$, $d = 0.26$). This is information important for consideration in the development of listening comprehension assessment materials especially for instructors who design and create their own materials to suit the specific needs and levels of their students.

Limitations and Future Directions

While 107 of the 110 participants in this study were Japanese nationals, the student body of the university also includes foreign students from Vietnam, Thailand, Malaysia, Mongolia, and China. Some cultural factors related to the L1 of the participants may be a fruitful avenue of study in checking if factors relating to the L1 of participants has a notable influence on their perception of pitch in an L2 context. While variation of pitch as well as lexical styles are demonstrated as important factors in a pitch accented language such as Japanese (Sano S. and Sano C., 2019), effects of pitch on L2 learners of English (a non-tonal language) in tonal based languages (such as Chinese or Thai) would also be another course of study of interest to this researcher.

As there has been research indicating an important link between the fundamental pitch and tonal slope in learners' L1 languages and their ability to grasp certain phonetic concepts in Asian languages such as Chinese and Japanese (Guion and Pederson, 2007), a focused study on the impact of students' L1 on their L2 listening comprehension would be an interesting continuation on this theme.

Conclusion

Research on the effect of pitch and listening has shown that pitch does have an effect on memory recall (Helfrich et al., 2011; Samoza et al., 2015) and even influences notions regarding the speaker (Aung et al., 2020). However, in the context of lower-level learners of L2 English, pitch also had a statistically significant effect on the results of listening comprehension tests. In the current study, 110 students of L2 English listened to two passages of comparable English level (CEFR B1), with one pitch-shifted down (118 Hz) and one pitch-shifted up (149.2 Hz). Students had higher results on listening comprehension tests, as well as perceived it to be easier to understand. These results indicate that the pitch of a voice used in testing materials bears serious consideration not only for ease of testing for the participants, but for their results as well. Free, open source software such as Praat and Audacity can be used by test creators and makers of educational materials to carefully measure and fine tune voices to suit the ability levels of their students. Such care to voice and pitch modulation can also have implications in fields such as creating AI text to voice models to create voice types conducive to better information retention in the targeted listeners.

References

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). The influence of topic knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learning and Individual Differences, 6*(4), 379–397. https://doi.org/10.1016/1041-6080(94)90001-9

Aung, T. & Puts, D. (2020). Voice pitch: a window into the communication of social power. *Current Opinion in Psychology, 33*, 154–161. https://doi.org/10.1016/j.copsyc.2019.07.028

Fay, A., & Buchweitz, A. (2014). Listening comprehension and individual differences in working memory capacity in beginning L2 learners. *Letrônica, 7*(1), 113–129. https://doi.org/10.15448/1984-4301.2014.1.16839

Fitch, J.L. & Holbrook, A. (1970). Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology – Head and Neck Surgery, 92*(4), 379-382, Table 2 (p. 381). https://doi.org/10.1001/archotol.1970.04310040067O1

Guion, S. G. & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O-S Bohn and M.J. Munro (Eds.) *Language Experience in Second Language Speech Learning: In honor of James E. Flege* (pp. 57–77) John Benjamins. https://doi.org/10.1075/lllt.17.09gui

Helfrich, H., & Weidenbecher, P. (2011). Impact of voice pitch on text memory. *Swiss Journal of Psychology, 70*(2), 85–93. https://doi.org/10.1024/1421-0185/a000042

Helfrich, H. (2004). Paralinguistic behaviors and culture. *Encyclopedia of Applied Psychology, 2*, 797–813. https://doi.org/10.1016/b0-12-657410-3/00035-0

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2013). Working memory and Second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2

Mullennix, J. W., Stern, S. E., Grounds, B., Kalas, R., Flaherty, M., Kowalok, S., May, E., & Tessmer, B. (2009). Earwitness memory: Distortions for voice pitch and speaking rate. *Applied Cognitive Psychology, 24*(4), 513–526. https://doi.org/10.1002/acp.1566

Novia, N., Bahri, S., & Inayah, N. (2018). The influence of speakers' voice in TOEFL listening test. *Research in English and Education (READ), 3*(3), 184–192.

Pourfannan, H., Mahzoon, H., Yoshikawa, Y., & Ishiguro, H. (2022). Towards a simultaneously speaking bilingual robot: Primary study on the effect of gender and pitch of the robot's voice. *PLOS ONE, 17*(12). https://doi.org/10.1371/journal.pone.0278852

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education, 33*(4), 511–531. https://doi.org/10.1007/bf00973770

Samoza, P. R., Sugay, J. F., Arellano, E., & Custodio, B. (2015). An evaluation of the effect of various voice qualities on memory retention. *Procedia Manufacturing, 3*, 1503–1510. https://doi.org/10.1016/j.promfg.2015.07.399

Sano, S., & Sano, C. (2019). The role of extra-linguistic factors in pitch range variation: A corpus study of spoken Japanese. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.) *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 290–294) Australasian Speech Science and Technology Association Inc.

Wallace, M. P., & Lee, K. (2020). Examining second language listening, vocabulary, and executive functioning. *Frontiers in Psychology, 11*(1122), 1–14. https://doi.org/10.3389/fpsyg.2020.01122

Appendix A

Script for the Listening Passage A

Sometimes in the summer, you can see large insects flying around. Many of these large flying insects are dragonflies.

Dragonflies are flying insects which have four large wings and long slender bodies. When at rest, dragonflies hold their wings horizontally.

Adult dragonflies often have brilliant colors. Their overall color is often a combination of yellow, red, brown and black. Freshly emerged adults are often a pale white color, but get their normal color after a few days. Once fully developed, the wings of dragonflies are a clear color with black.

Young dragonflies, called nymphs, are usually a blend of brown, green and grey.

Dragonflies are excellent hunters, and often catch their food in the air. Adult dragonflies eat other insects such as mosquitoes, butterflies, and sometimes even other dragonflies.

Dragonfly nymphs are also good hunters, and eat animals smaller than themselves. For example, worms, other baby insects (called larvae), and sometimes small frogs and fish.

In addition to having excellent flying and hunting ability, dragonflies also have advanced eyesight. Dragonflies see about three hundred images in a second, and can also see in three hundred sixty degrees. About eighty percent of the dragonfly's brain is used for its eyes.

Appendix B

Comprehension Quiz For Listening Passage A (English translation)

1. What is this speech about?
   a) dragons
   b) dragonflies
   c) earthworms
   d) flies
2. What kind of bodies do dragonflies have?
   a) two wings
   b) long and thin
   c) short and round
   d) thin and round
3. What color are matured dragonflies?
   a) clear
   b) gray
   c) pale white
   d) yellow
4. What is a "nymph"?
   a) a young dragonfly
   b) a small insect
   c) a mature dragonfly
   d) a type of color
5. What color are nymphs?
   a) yellow and red
   b) pale white
   c) brown and gray
   d) clear
6. Which do mature dragonflies typically eat?
   a) mosquitoes
   b) mice
   c) frogs
   d) fish
7. Which do nymphs typically eat?
   a) earthworms
   b) large frogs and fish
   c) praying mantises
   d) butterflies
8. Dragonflies have advanced ____?
   a) biting power
   b) flying ability
   c) degrees
   d) eyesight
9. How many images does a dragonfly's brain see in one second?
   a) 4
   b) 80
   c) 300
   d) 360
10. What is 80% of a dragonfly's brain used for?
    a) flying
    b) hunting
    c) eyes
    d) brain
11. What is your native language?
    a) Japanese f) Malay
    b) Korean g) Indonesian
    c) Chinese h) Bangladeshi
    d) Vietnamese i) English
    e) Thai j) Other
12. If "Other", please specify: _____

Appendix C

Script for the Listening Passage B

In North America, the spotted lantern fly has become a problem. This insect is native to Asia, but arrived in North America in 2014. The spotted lanternfly has 4 wings - two large grey wings with small black spots, and two small red wings with small black spots. This fly is about 2.5 cm long and 1.2 cm wide. It has a small, black head, and red eyes. Its stomach is often a yellow color.

Although they have 4 wings, spotted lantern flies are not good fliers. Instead, they jump from plant to plant to feed. Spotted lantern flies do not have teeth or a stinger, so they cannot bite or hurt humans. However, they can harm plants and trees. For example, spotted lanternflies eat the sap, or juice, of plants and trees. Sometimes many spotted lanternflies eat at the same time, so this can damage or kill many plants and trees. This is especially a problem

for businesses that grow fruits. Because spotted lanternflies also eat the juice of grape vines, these flies can have a very large impact on wine business and local economies.

The spotted lantern fly has no natural predators in North America, so people were told one thing - if you see a spotted lanternfly, kill it!

Appendix D

Comprehension Quiz For Listening Passage B (English translation)

1. Where are spotted lanternflies originally from?
   a) North America
   b) Asia
   c) South America
   d) Europe
2. How many wings do spotted lanternflies have?
   a) two wings
   b) four wings
   c) they do not have wings
   d) six wings
3. What year did spotted lanternflies arrive in North America?
   a) 2014
   b) 1914
   c) 2040
   d) 2004
4. Spotted lanternflies are about _____.
   a) 1.2 cm
   b) 2.5 cm
   c) 2.5 mm
   d) 4 cm
5. What color are the wings of spotted lanternflies?
   a) yellow
   b) red with black spots
   c) gray with black spots
   d) clear
6. How are spotted lanternflies harmful?
   a) they harm local wildlife
   b) they have poison which is dangerous to humans
   c) they emit a foul odor
   d) they are harmful to local agriculture
7. Which do spotted lanternflies typically eat?
   a) earthworms
   b) other insects
   c) pollen
   d) tree sap
8. Spotted lanternflies are not good at _____.
   a) biting
   b) hunting
   c) seeing
   d) flying
9. In North America, what are people told to do if they see a spotted lanternfly?
   a) kill it
   b) capture it
   c) take a picture
   d) report its location to the authorities
10. According to the passage, spotted lanternflies are _____.
    a) pests
    b) poisonous
    c) beneficial organisms
    d) an endangered species
11. What is your native language?
    a) Japanese
    b) Korean
    c) Chinese
    d) Vietnamese
    e) Thai
    f) Malay
    g) Indonesian
    h) Bangladeshi
    i) English
    j) Other
12. If "Other", please specify: _____

Appendix E

Post-testing Impressions Questionnaire (English translation)

1. Between samples A and B, which do you feel was easier to understand/comprehend?
   a) A was much easier to understand
   b) A was a little easier to understand
   c) they were the same
   d) B was a little easier to understand
   e) B was much easier to understand
2. Please explain your choice from the above.