

**USABILITY TESTING OF JAPANESE CAPTIONS
SEGMENTATION SYSTEM TO SCAFFOLD BEGINNERS TO
COMPREHEND JAPANESE VIDEOS**

Ya-Fei Yang

National University of Tainan/Chung Hwa University of Medical Technology
No.89, Wenhwa 1st St., Rende Shiang, Tainan County 717, Taiwan
chck.chck@msa.hinet.net

Ching-Ju Chao

National University of Tainan/TungFang Design Institute
No.110, Dongfang Rd., Hunei Dist., Kaohsiung City 82941, Taiwan
chingju@mail.tf.edu.tw

Yu-Ling Lin

National University of Tainan
33, Sec. 2, Shu-Lin St., Tainan 70005, Taiwan
whitepig564@gmail.com

Chih-Kai Chang

National University of Tainan
33, Sec. 2, Shu-Lin St., Tainan 70005, Taiwan
chihkai@mail.nutn.edu.tw

ABSTRACT

A major learning difficulty of Japanese foreign language (JFL) learners is the complex composition of two syllabaries, hiragana and katakana, and kanji characters adopted from logographic Chinese ones. As the number of Japanese language learners increases, computer-assisted Japanese language education gradually gains more attention. This study aimed to adopt a Japanese word segmentation system to help JFL learners overcome literacy problems. This study adopted MeCab, a Japanese morphological analyzer and part-of-speech (POS) tagger, to segment Japanese texts into separate morphemes by adding spaces and to attach POS tags to each morpheme for beginners. The participants were asked to participate in three experimental activities involving watching two Japanese videos with general or segmented Japanese captions and complete the Nielsen's Attributes of Usability (NAU) survey and the After Scenario Questionnaire (ASQ) to evaluate the usability of the learning activities. The results of the system evaluation showed that the videos with the segmented captions could increase the participants' learning motivation and willingness to adopt the word segmentation system to learn Japanese.

Keywords: Japanese Language Learning, Japanese Caption Processing, Japanese Word Segmentation System, Japanese Wordnet, Usability

INTRODUCTION

In 2012, 64,600 Taiwanese examinees took the Japanese Language Proficiency Test (JLPT) (approximately 10% of the total examinees). Although China and Korea are the top two ranked countries with 239,369 (not including Hong Kong and Macau) and 90,293 examinees, respectively, who took the JLPT, they have greater populations compared to Taiwan (retrieved from <http://www.jlpt.jp/e/statistics/archive.html>). However, the proportion of the Taiwanese population is more engaged in learning Japanese compared to other countries. The Japanese Language Proficiency Test (JLPT) started in 1984 as a test to assess and certify the Japanese language proficiency of those whose native language is not Japanese. In the first year, the JLPT was conducted in 15 countries, and approximately 7,000 examinees took the test. Since then, the JLPT has become the largest Japanese language test in the world, administered to approximately 610,000 examinees from 62

countries and areas worldwide in 2011.

The number of Japanese learners is increasing and would exceed the population of other foreign language learners, except English learners in Taiwan. As the number of Japanese language learners increases, Japanese language education gradually gains more attention in Taiwan. Many enterprises adopt the results of the JLPT as a basis for employment, promotion, and pay rise. Therefore, Japanese language teachers must understand students' learning motivation and needs to effectively enhance students' learning effectiveness and willingness to learn Japanese language. Those second language researchers indicated that listening comprehension is often treated as a *Cinderella* skill in language learning (Flowerdew & Miller, 2005).

Second language listening comprehension has been neglected in the field of applied linguistics for several years. However, listening comprehension is vital in the development of second language competence, and its importance in language learning has only been recognized recently (Lin, 2010). Wu (2008) used the SILL questionnaire (Strategy, Inventory for Language Learning) to survey JFL students. The survey results showed that some learners prefer watching video without Japanese captions. Moreover, those learners indicated that their motivation was promoted when they improve Japanese listening comprehension by watching blockbusters in their leisure time. While watching blockbusters, they use the meta-skills of listening comprehension (such as guess and visual clues so on) instead of comprehend Japanese captions word by word. Helping students apply listening comprehension skills in activities can improve their listening comprehension and willingness to learn Japanese.

In recent years, a significant change in the field of natural language processing has concerned the transition from the artificial production system to automatic text classification (Cardie & Mooney, 1999). Japanese National Institute of Information and Communications Technology (NICT) started to develop the Japanese WordNet in 2006 and publicly released its first version on February 2009. The current version of Japanese WordNet (v.1.1) consists of 57,238 concepts (synsets), 93,834 words, 158,058 senses (synset-word pairs), 135,692 definitions, and 48,276 examples (Isahara, Bond, Uchimoto, Utiyama, & Kanzaki, 2008). WordNet is a publicly available semantic dictionary of English widely applied in the research field of natural language processing. It distinguishes nouns, verbs, adjectives, adverbs, and a group of synonymous words, such as Hyponym, Hypernym, Antonym, and Synsets. In addition, WordNet, a fairly complete electronic lexical database, is considered to be the most important resource available to researchers in

different areas (Fellbaum, 2010; Miller, 1995). This study proposed an automatic approach that uses the ontological framework of WordNet to retrieve nouns from the Japanese captions of the selected videos as supplementary learning materials.

RESEARCH METHODOLOGY

Watching videos with captions apparently assists individuals in understanding the video content and improves their listening comprehension (Winke, Gass, & Sydorenko, 2010). Therefore, this study aimed to investigate the use of different captions and attempted to provide the supplementary learning materials for the pre-listening activity to further enhance Japanese language learners' listening comprehension. Two Japanese videos, *From Up On Poppy Hill* and *Spirited Away*, were selected as video watching activities every other week. The experiment lasted for two months.

Before designing and developing the Japanese word segmentation system, the researchers considered issues of how to (1) retrieve the Japanese captions, (2) correctly segment Japanese captions, (3) select important Japanese vocabulary for language beginners to improve their listening comprehension, and (4) automatically yield learning materials of the pre-listening activity from the system. Japanese captions could be retrieved from Japanese captions websites. In addition, using Google2SRT can also download XML captions (Closed Captions) from YouTube videos and convert them to SRT format recognized by most video players. This study adopted MeCab to segment Japanese texts into separate morphemes by adding spaces and used WordNet to retrieve nouns from the Japanese captions as learning materials of the pre-listening activity. The Nielsen's Attributes of Usability (NAU) and the After Scenario Questionnaire (ASQ) were adopted as tools to evaluate the usability of the learning abilities. The NAU uses 5 items to evaluate the attributes of the system regarding learnability, efficiency, memorability, accuracy, and subjective satisfaction (Hornbæk & Law, 2007; Nielsen & Levy, 1994; Nielsen & Hackos, 1993). The ASQ uses 3 items to evaluate the overall satisfaction of the learning activities (Lewis, 1991; Lewis, 1995). After completing the three experimental activities, the participants were asked to fill out both questionnaires. The learners responded to items on a 7-point Likert-scale, with a score of "1" indicating strongly disagree and a score of "7" indicating strongly agree.

This study developed the Japanese word segmentation system to investigate whether the system is feasible to improve the participants' listening comprehension and enhance their learning effectiveness and interests. The research hypotheses are listed as follows:

Hypothesis 1: Providing general Japanese captions are helpful to improve Japanese beginners' listening comprehension.

Hypothesis 2: Providing the segmented Japanese captions can improve beginners' satisfaction.

Hypothesis 3: Providing the segmented Japanese captions and learning materials in the pre-listening activity can improve the students' learning effectiveness, memory maintenance, and overall satisfaction with regard to the system.

RESEARCH DESIGN

The modern Japanese writing system consists of several thousand Kanji, adopted from Chinese characters, and two phonetic scripts (kana), hiragana and katakana. Japanese language learners must overcome the difficulty of appropriately segmenting Japanese texts and then learn Japanese nouns. Therefore, the main focus of the system is to appropriately segment the Japanese captions and retrieve nouns from the captions as the learning materials. This study implemented similarity measures to distinguish the nouns retrieved from Japanese WordNet. MeCab is a new morphological analyzer and part-of-speech (POS) tagger that segments Japanese captions (Kudo, 2005). Other common morphological analyzers are ChaSen, JUMAN, and MeCab. Mao (2012) stated that the accuracy of Japanese segmentation by MeCab is better than that by ChaSen and JUMAN. This study used Python, a programming language, to integrate Japanese WordNet with MeCab. Figure 1 displays the framework of the Japanese segmentation system.

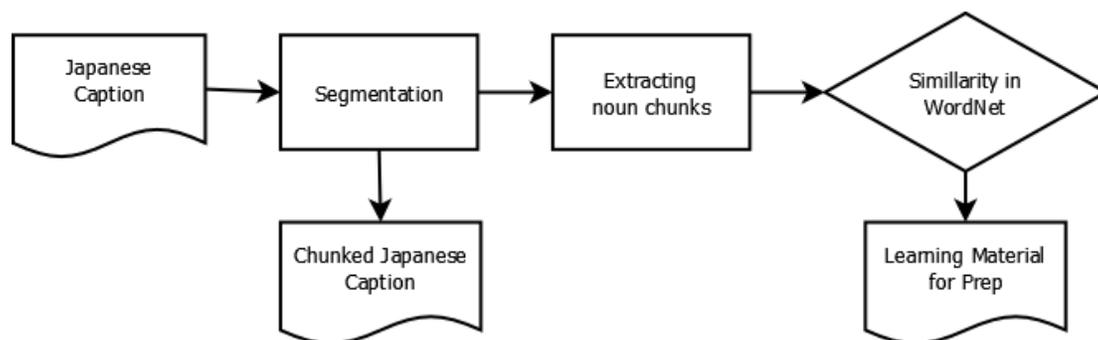


Figure 1 The Framework of the Japanese Segmentation System

This study retrieved nouns from the Japanese captions of the two videos. However, more than 700 frequently used nouns were retrieved from the Japanese captions of each 30-minute video; those retrieved nouns need to be further classified and organized as the learning materials for the pre-listening activity. This study did not filter nouns based on the high frequency words because names, locations, objects, and issues are not the focus of the videos, and they appear repetitively in the captions. This study implemented similarity measures to distinguish the retrieved nouns in Japanese WordNet and obtain a similar effect of concept clustering. Many similarity measures exist in Japanese WordNet, including (1) Wu-Palmer Similarity, (2) Resnik Similarity, (3) Path Distance Similarity, (4) Lin Similarity, (5) Leacock Chodorow Similarity, and (6) Jiang-Conrath Similarity. This study adopted the Wu-Palmer Similarity measure to sort the retrieved nouns into groups of high similarity (0.7-0.9), moderate similarity (0.6-0.7), and additional important nouns provided by an instructor as the learning materials for the pre-listening activity.

Figure 2 displays the experimental procedures. Oxford (1990) stated that visual images effectively facilitate learners to recall oral messages and store them in long-term memory. Learners can easily remember what they hear and read of target-language information along with sounds and visual images. Two Japanese animation movies were selected as supplementary learning materials for listening comprehension with an aim to enhance students' Japanese learning experience during the entire semester. This study recruited 30 participants from the Department of Tourism at a university. They were asked to watch two videos divided into four parts. The participants were asked to watch the first part of the first video with the Chinese subtitles, which can help them comprehend the video content. However, the instructor found that most participants focused only on the Chinese subtitles and ignored the Japanese pronunciation from the class discussion. Thus, this study conducted the following three experimental activities and investigated whether the participants could combine what they hear with what they see when watching the videos to enhance their Japanese listening comprehension.

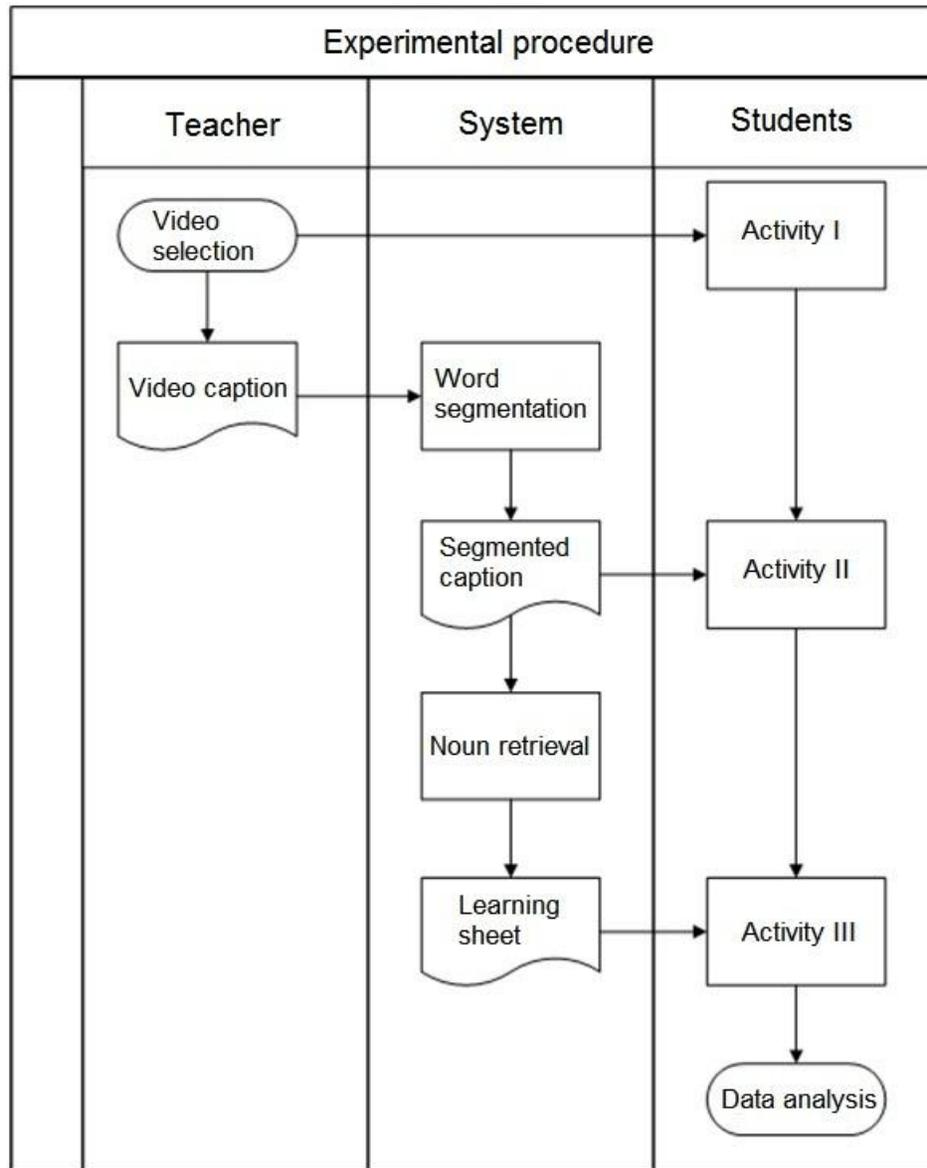


Figure 2 The Flowchart of the Experiment

In the activity I, the participants were asked to watch the second part of the first video, *From Up On Poppy Hill*, with general Japanese captions and a combination of Kanji and kana scripts. The purpose of the activity was to investigate whether the Japanese texts and sounds would lead to dual cognitive load, which could be measured using the questionnaires and observations. In the activity II, the participants were asked to watch the first part of the second video, *Spirited Away*, with segmented hiragana

captions the aim of which was to reduce the participants' cognitive load, allowing them to stay focused on listening. In the activity III, the participants were asked to learn the retrieved nouns, which were classified into highly or moderately similar ones according to the *corpus based on word similarity* in the Japanese WordNet as well as other related nouns provided by the instructor as the learning materials in the pre-listening activity. Activity III preceded watching the second part of the second video, *Spirited Away*. Therefore, the participants could understand the video content better because of the pre-listening activity.



Figure 3 Snapshot of Japanese Captions before (Right) and after (Left) Segmentation

Figure 3 shows the differences between two segmented Japanese captions of the video, *Spirited Away*. Figure 3 (right) shows the general Japanese captions, a combination of Kanji and kana scripts, without segmentation. Figure 3 (left) shows the Japanese hiragana captions, a string of morphemes annotated with parts of speech (POSS) using morphological analysis. The sentence in the snapshot of Figure 3 (left), まえのほうがいいもん, means that I prefer the previous one. After segmentation by MeCab, the sentence becomes まえのほうがいいもん, which means better before. Noteworthy, the words within parentheses in the beginning of the sentence, (千尋), indicate the female speaking character. Thus, this study removed the speaking character indicators from the original Japanese captions before using MeCab to segment the Japanese captions. Moreover, the Kanji characters (e.g., 前 and 方) are difficult for novice to pronounce. Consequently, all Kanji characters were transformed into hiragana in the segmented Japanese captions.

RESULTS

This study adopted the Nielsen's Attributes of Usability (NAU) and the After Scenario Questionnaire (ASQ) as tools to evaluate the usability of the learning abilities. The results of the questionnaires revealed that the learners have positive feedback on learnability, memorability, efficiency, accuracy, subjective satisfaction of the integration of the segmented captions in the experimental activities, which facilitate the participants to learn Japanese. The results indicated that the videos with the segmented captions can effectively increase the participants' learning motivation and that most learners are willing to adopt the segmented captions of the system to learn Japanese, which is worthy of further promotion and use.

Table 1 Effects of Providing the General Japanese Captions

NAU	N.	Mean	SE	SD	Var
Learnability: Students are easier to learn the video content	30	4.53	.213	1.167	1.361
Efficiency: The system enhances students' learning effectiveness	30	4.53	.213	1.167	1.361
Memorability: Students are easier to remember the video content	30	4.63	.200	1.098	1.206
Accuracy: Students accurately comprehend the learning content	30	4.53	.208	1.137	1.292
Subjective satisfaction: Students are satisfied with the learning activities	30	5.30	.245	1.343	1.803
ASQ	N.	Mean	SE	SD	Var
Overall, I am satisfied with the ease of completing the tasks in this scenario.	30	4.73	.249	1.363	1.857
Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario.	30	4.63	.200	1.098	1.206
Overall, I am satisfied with the support information when completing the tasks.	30	4.43	.243	1.331	1.771

Effects of Providing General Japanese Captions

The participants were required to watch 90-minute video, *From Up On Poppy Hill*, with general Japanese captions. The descriptive statistics showed that the mean value on learnability, efficiency, and accuracy of the system is 4.53 and the standard deviation is between 0.208 and 0.213, indicating that the learners can accurately comprehend the

video content. Thus, the experimental activity I can enhance learning effectiveness. Table 1 shows that the mean value is above 4 on learnability items, sufficient time to complete tasks, and ease of completing tasks, indicating that the learners agree that the learning activity is helpful to memorize the video content. Moreover, the length of the learning activity is sufficient to complete tasks without causing the cognitive burden on learning, allowing the participants to complete tasks easily. The mean value of the overall subjective satisfaction reached 5.30, showing that the learners are satisfied with the learning activity. However, with regard to the learning context, additional supplementary information may be included to assist the learners to complete tasks.

Effects of Providing Segmented Japanese Captions

The individual interviews were conducted with eight learners, which is a small sample size. In experimental activity II, the participants were asked to watch the first part of the Japanese video, *Spirited Away*, for 68 minutes, with hiragana captions containing no Kanji character in the captions. The results of the interviews indicated that the interviewees could comprehend the video content of the *From Up On Poppy Hill* mostly in the experimental activity I; however, messages sent through multiple channels as well as complex structures, codings, and visual messages of the system distracting the participants' attention from comprehending the video content. In addition, the interviewees also stated that the hiragana captions can assist them with pronunciation of the Japanese captions but are not helpful to comprehend the video content. Japanese hiragana is a phonetic script that does not have any meaning; however, each Japanese Kanji is an ideographic character that has its own meaning. Since Chinese characters are similar to Kanji ones, Taiwanese students who learn Japanese as a foreign language have a better understanding of reading Japanese captions but neglect the pronunciation of those Chinese characters in captions while watching Japanese videos.

Effects of Providing Learning Materials for the Pre-Listening Activity

Overall, 24 participants were asked to watch the video *Spirited Away* for 58 minutes in the experimental activity III. The pre-listening activity required the participants to learn Japanese vocabulary retrieved from the captions, *Spirited Away*, and classify them into highly and moderately similar nouns according to the *word similarity* in the Japanese WordNet. In addition, the instructor also provided related vocabulary to help participants recognize the meanings of the retrieved Japanese nouns and comprehend the video

content to prevent the participants from message interference and distraction. Therefore, the participants can stay focused on the listening comprehension activity. Table 2 lists the generated nouns, which served as the learning materials of the pre-listening activity before watching the video.

Table 2 Learning Materials for Pre-Listening Activity.

Semantic highly similar		Semantic moderately similar		Instructor Highlighted and provided	
Target	Related	Target	Related	Target	Related
にんげん	きやくじん	にんげん	おじさん	千(せん)	千尋(ちひろ)
にんげん	いところ	にんげん	こむすめ	白(はく)	竜(りゅう)
にんげん	ちび	にんげん	やつ	ゆばーば	ぜにーば
にんげん	ばか	にんげん	しんまい	釜爺(かまじい)	カオナシ
こむすめ	ねえさん	にんげん	ねえさん	坊(ぼう)	リン
これ、それ、あれ	ここ、そこ、あそこ	きやくじん	いところ	青蛙(かえる)	頭(かしら)
もと	ところ	きやくじん	ちび		
もの	むつつ	きやくじん	ばか		
ただ	ひとり	そこ	もと		
やつ	ばか	そこ	すみ		
		そこ	ところ		

Table 3 shows the mean value of 4.83 on learnability and efficiency, with standard deviation between 0.246 and 0.253, indicating that the participants had a better understanding of the video content in this experimental activity compared to the experimental activity II. The participants stated that the pre-listening activity could help them memorize Japanese vocabulary. Thus, they had sufficient time to accomplish the activity. The results showed that the participants could comprehend the keywords of the video better in this activity than they could in the experimental activity II, which is consistent with the utilization stage of Anderson's (1988) three-phase comprehension model, indicating that the participants can relate the mental representation of the vocabulary to their prior knowledge.

Table 3 Effects of Providing Pre-Listening Activity Materials

NAU	N.	Mean	SE	SD	Var
Learnability: Students can learn the video content easier.	24	4.83	.253	1.239	1.536
Efficiency: The system enhances students' learning effectiveness.	24	4.83	.246	1.204	1.449
Memorability: Students can remember the video content easier.	24	4.92	.275	1.349	1.819
Accuracy: Students accurately comprehend the learning content.	24	4.71	.266	1.301	1.694
Subjective satisfaction: Students are satisfied with the learning activities.	24	5.29	.285	1.398	1.955
ASQ	N.	Mean	SE	SD	Var
Overall, I am satisfied with the ease of completing the tasks in this scenario.	24	4.67	.305	1.494	2.232
Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario.	24	4.29	.304	1.488	2.216
Overall, I am satisfied with the support information when completing the tasks.	24	4.62	.355	1.740	3.027

CONCLUSION

The study developed a Japanese word segmentation system. This study aimed to investigate whether the segmented Japanese captions effectively improve learners' listening comprehension. Furthermore, it aimed to understand whether the supplementary learning materials of the pre-listening activity, which contain the retrieved nouns from the Japanese captions based on the WordNet's ontological structure, help learners comprehend the learning content. The experiment lasted for two months. The participants completed three experimental activities, which involved watching two Japanese videos, one with general Japanese captions and other with segmented Japanese captions. Concerning activity with segmented captions, they participated in the pre-listening activity before watching a video every other week. The study adopted the Nielsen's Attributes of Usability (NAU) and the After Scenario Questionnaire (ASQ) as tools to evaluate the usability of learning activities. The results revealed that learners provided

positive feedback on learnability, memorability, accuracy, and subjective satisfaction with the integration of the segmented captions in learning activities. The experimental results showed that the videos with the segmented captions can increase the participants' learning motivation, and most learners were willing to adopt the Japanese word segmentation system to learn Japanese. Thus, it would be worth to further promote and use this system.

ACKNOWLEDGMENT

The research reported in this paper has been supported by the National Science Council in Taiwan under the research project number NSC 100-2631-S-001-001, NSC 100-2628-S-024-001-MY3, and NSC 101-2511-S-024-007-MY2. Corresponding author: Chih-Kai Chang.

REFERENCES

- Anderson, A., & Lynch, T. (1988). *Listening*. New York: Oxford University Press.
- Cardie, C., & Mooney, R.J. (1999). Guest editors' introduction: Machine learning and natural language. *Machine Learning*, 34(1), 5-9.
- Fellbaum, C. (2010). WordNet. In R. Poli, M. Healy, and A. Kameas (Eds.), *Theory and applications of ontology: Computer applications* (pp. 231-243). Netherlands: Springer. http://dx.doi.org/10.1007/978-90-481-8847-5_10.
- Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. New York, USA: Cambridge University Press.
- Hornbæk, K., & Law, E.L.C. (2007). Meta-analysis of correlations among usability measures. In B. Begole, S. Payne, E. Churchill, R.S. Amant, D. Gilmore, and M.B. Rosson (Eds.), *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 617-626). New York: ACM. <http://dx.doi.org/10.1145/1240624.1240722>.
- Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., & Kanzaki, K. (2008). Development of the Japanese WordNet. In European Language Resources Association (Ed.), *Proceedings of the Sixth International Conference on Language Resources and Evaluation* (pp. 2420-2423). Paris, France: ELRA.
- Kudo, T. (2005). MeCab: Yet another part-of-speech and morphological analyzer. Retrieved January 6, 2013, from <http://mecab.sourceforge.net/>.
- Lewis, J.R. (1991). Psychometric evaluation of an after-scenario questionnaire for

- computer usability studies: The ASQ. *ACM SIGCHI Bulletin*, 23(1), 78-81. <http://dx.doi.org/10.1145/122672.122692>.
- Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78. <http://dx.doi.org/10.1080/10447319509526110>.
- Lin, Y.S. (2010). A constructivist study on Taiwanese college EFL learners' listening strategies. *Chin-Yi Journal of the Humanities and Social Sciences*, 2, 153-182.
- Mao, W.W. (2012). A research on the Japanese open-source automatic POS taggers. *Computer-Assisted Foreign Language Education*, 145, 10-14. (in Chinese)
- Miller, G.A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39-41. <http://dx.doi.org/10.1145/219717.219748>.
- Nielsen, J., & Hackos, J.T. (1993). *Usability engineering (Vol. 125184069)*. Boston: Academic Press. <http://dx.doi.org/10.1016/B978-0-08-052029-2.50007-3>.
- Nielsen, J., & Levy, J. (1994). Measuring usability: Preference vs. performance. *Communications of the ACM*, 37(4), 66-75. <http://dx.doi.org/10.1145/175276.175282>.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York : Newbury House.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65-86.
- Wu, J.H. (2008). Investigation of advanced Japanese listening comprehension learning tools and strategies. *Journal of Japanese Literature & Language in Taiwan*, 23, 229-253.