

**The Relationship between L2 Listening Proficiency and
L2 Aural Language Processing**

Paul Joyce

Department of Law, Kindai University
3-4-1 Kowakae, Higashi-shi, Osaka 577-8502, Japan
Email: pauljoyce@hotmail.com

Abstract

Listening has been widely characterised as a multi-faceted process encompassing a range of linguistic and psycholinguistic components (see Rubin, 1994). However, for learners at different levels of L2 proficiency, there is uncertainty over the relative importance of the various sub-skills. To address this issue, a number of linguistic and psycholinguistic sub-skills that are associated with L2 listening proficiency were selected and operationalised. This battery of discrete point measures, as well as two L2 proficiency tests, was administered to 443 Japanese university students. After the data had been subjected to descriptive and inferential analysis, the findings indicated that L2 listening comprehension is most closely associated with L2 syntactic knowledge, followed by the ability to recognise words in connected speech. The results also revealed that listeners at different proficiency levels process the language in decidedly different ways. Less proficient learners were discovered to be far more dependent on the linguistic and psycholinguistic sub-skills that are closest to the surface of the message. On the other hand, owing to the development of their syntactic knowledge and recognition of words in connected speech, more proficient listeners benefitted from a greater

interaction between their more closely entwined higher and lower level processing skills.

Keywords: L2 listening, L2 proficiency, L2 listening construct, L2 processing

Introduction

Listening is the most widely used of the four traditional language skills (Feyton, 1991). And it has been found to play a critical role in second language (L2) learning (Long, 1985). However, as well as being important, L2 listening has also been found to be the most difficult language skill to learn (El-dali, 2017; Graham, 2003). The reason for this could relate to uncertainty over how the L2 listening process operates. One particular area of ambiguity is the relationship between aural language processing and L2 proficiency. It has been recognised that listening comprehension is the result of a series of complex and active processes (Hansen & Jensen, 1994) including phoneme recognition, lexical segmentation, and syntactic processing sub-skills (Rubin, 1994). However, to facilitate a greater understanding of which sub-skills L2 learners most need to develop, this theoretical perspective needs to be underpinned by empirical research to determine how these sub-skills interact with the proficiency level of the L2 listener.

Language processing is typically conceptualised through two diametrically opposed theories of comprehension: the “bottom-up” and the “top-down” views (Tsui & Fullilove, 1998). The bottom-up perspective holds that the aural signal is divided into distinct levels of processing. These levels include the phoneme, the word, the clause, the sentence, and finally the intended meaning of the complete utterance (Buck, 2001). It is posited that the output from the lower levels of processing is employed as the input for the higher levels. In contrast, the “top-down” approach emphasises the use of background, world, and contextual knowledge to construct semantic expectations of the language, which the speech signal confirms or rejects (Buck, 2001). Notably,

the models only permit unidirectional processing. In the case of bottom-up listening, this means that the acoustic message is processed from the phoneme level upwards. However, should the listener's interpretation of a phoneme string be constrained by their knowledge of a word's existence, this would qualify as top-down processing. Consequently, it would fall outside the model's explanatory realm. Since it is widely concurred that a complex interplay exists between the two processing styles (e.g. Frauenfelder & Komisarjevsky-Tyler, 1987; Witkin, 1990), it is evident that L2 listening can be neither regarded as entirely bottom-up nor top-down based.

In light of the deficiencies that are evident in both the bottom-up and top-down theories, the interactive-compensatory model has been advanced (Stanovich, 1980). As the name suggests, this model proposes that higher-level semantic and lower-level linguistic processes can be employed simultaneously. Since the late 1970s, the interactive model has proven to be the dominant conceptual paradigm. Nevertheless, the model provides little guidance on the relationship between proficiency level and higher and lower processing level. And while some studies have found that comprehension breakdowns are due to shortcomings in higher-level processing skills (e.g. Hansen & Jensen, 1994; Shohamy & Inbar, 1991), others have concluded that deficiencies in lower-level processing skills are responsible (e.g. Tsui & Fullilove, 1998; Wolff, 1987).

An alternative standpoint on the process of listening comprehension is provided by cognitive psychology. From this perspective, listening is viewed as comprising a number of complex, active sub-skills (Goh, 2017). The interaction between the various component sub-skills is postulated to determine the degree of comprehension (Byrnes, 1984; Call, 1985). One such cognitive model (Anderson, 1995) consists of three stages; perception, parsing and utilization. While these stages are considered to be interactive and recursive, Anderson (1995) also notes that they are by necessity partially ordered in time. During the initial perception stage, the acoustic signal is encoded into a

meaningful form. To enable this, the listener pays close attention to the aural input, which upon contact is momentarily retained in the echoic memory (Bacon, 1992). While in the echoic memory, it is asserted that the continuous speech stream is segmented into phonemes (Anderson, 1995). At this early stage, the listener applies their knowledge of the rules of segmentation to transform the text into meaningful lexical representations. In the second stage of the listening process, the words undergo parsing. That is, the lexical items that were identified at the perception stage are combined into chunks. The chunking is primarily directed by cues to meaning, and the listener's knowledge of syntactic structures. Following the aggregation of the words into chunks, the individual speech segments are converted into a mental representation of the message. When processing speech, prosodic features of the language influence the nature of the message. For instance, the placement of sentence stress affects the semantic force of the utterance. Once a mental representation of a section of language has been derived, this chunk of information is transferred to short-term memory. Crucially, short-term memory is limited in its storage capacity, especially for L2 input (Call, 1985). Therefore, it is essential that language is parsed efficiently. Through the aggregation of chunks of understanding, the listener derives a full mental representation of the message. During the utilisation phase of the comprehension process, the listener relates their mental representation of the text to their existing knowledge. This knowledge is stored within long-term memory as scripts, and other interrelated concepts.

As has been discussed above, a wide range of sub-skill processes are employed during the three phases of aural comprehension. From reviewing the literature, linguistic and psycholinguistic sub-skills of primary importance in this process are syntactic knowledge (Alderson, 1993; Rost, 1990), vocabulary breadth (Bonk, 2000; Cheng & Matthews, 2016; Kelly, 1991; Révész & Brunfaut, 2013), the ability to recognise words in speech through phonological modification knowledge (PMK) (Brown & Hilferty, 1986; Ito, 2001), working memory capacity (Call, 1985;

Harrington & Sawyer, 1992; Vandergrift & Baker, 2018), phonological awareness (Mack, 1988; Mora, 2005), sentence stress awareness (Vanderplank, 1985) and metacognitive listening strategy use (Goh, 1998; Vandergrift, 2005). While the collective and relative importance of these listening sub-skills for Japanese learners has been previously addressed (see Joyce, 2011), the relationship between these processes and learners at different proficiency levels remains unclear. To help bridge the gap between L2 listening theory and practice, the following research question was formulated:

To what extent does the development of Japanese learners' selected linguistic knowledge and psycholinguistic processing sub-skills differ as a function of their L2 listening comprehension ability?

Methodology

Participants

The research was conducted in Japan at a university specializing in foreign language studies. All of the 443 participants were Japanese L1 speakers who were enrolled as full-time English language major undergraduates. In terms of proficiency, the participants could broadly be described as being from a false beginner to an upper intermediate level. All of the 443 students undertook all of the different research instruments. The participants were classified into groups based upon their combined score on two listening proficiency tests. Participants who scored in the top third ($n = 148$) were considered the higher proficiency group and those in the bottom third ($n = 148$) the lower proficiency group. When using the CEFR (Common European Framework of Reference for Languages) scale, the higher proficiency group were at around the B1 level, and the lower proficiency group at the A2 level. Since the selection of the participants was determined by the cooperation of EFL instructors, a convenience sample was used. Informed consent was obtained from all participants.

Materials

The section below details how the L2 listening construct and the sub-skills considered important for its success were operationalised.

L2 Listening Proficiency.

Two different L2 listening proficiency tests were employed. The first was the listening section of a university in-house general proficiency test. The second test was the listening section of the TOEFL. Although both were communicative language tests that targeted the general listening proficiency domain, they varied significantly in their format and delivery. The video-mediated in-house test was mainly based upon extended texts and had around six items per passage. And since the test questions were printed in a test booklet, the candidates had a clear purpose for listening. On the other hand, for the most part, the assessed material for the TOEFL consisted of short texts that were associated with a single item. In contrast to the first test, the TOEFL questions were presented aurally, after the listening passage had been heard. While the two tests were very different, together it was considered that they provided a well-rounded measure of general listening proficiency.

L2 Syntactic Knowledge.

The L2 Syntactic Knowledge test was based on the Listening Comprehension Test (LCT), a commercially produced aural test of grammar (ELI LCT Manual, 1986). The content of the LCT was derived from two commonly used grammar textbooks (Krohn, 1971; Lado & Fries, 1958). To ensure the purity of the construct, the grammatical forms were presented in short decontextualised statements or questions. Learners were required to select a multiple-choice option that was either similar in meaning to a statement they heard or answered a question that they heard. When reviewing the LCT, Kuehn (1993) observed, “inferences made from [LCT] scores should be limited only to recognition of basic structures in English.” To improve the trait

purity of the LCT, it was ensured that all of the lexical items contained in the sentences were either within the 1000 most frequent word families, or encompassed by the list of core vocabulary items that are taught at all Japanese junior high schools. Also, when the listening materials were rerecorded, the texts were delivered in a relatively slow and formal manner to minimise the influence of reduced forms. As is the case for most of the sub-skill measures, an example test item is contained in Appendix 1.

L2 Vocabulary Breadth.

The content of the L2 Vocabulary Breadth test was sampled from a lemmatised version of the 10 million word spoken version of the British National Corpus (BNC). In order to ensure that the meaning of the assessed vocabulary could not be inferred from context, the assessed lexical items were each aurally presented in isolation. However, due to the lack of contextual support for the vocabulary, and the resulting potential for phonological misperception, the assessed items were each presented to the listeners twice. The participants were required to select the meaning of the words they heard from five multiple choice options. To safeguard the trait purity of a test, the answer choices consisted of lexical items in the participants' first language (see Nation, 2001).

L2 Phonological Modification Knowledge (PMK).

The PMK construct was operationalised through a dictation test. Since there is currently a lack of quantitative data on the frequency of common reduced forms, the assessed material was derived from the reduced forms that researchers (Brown & Hilferty, 1989; Ur, 1984; Weinstein, 1982) consider to be the most important. For instance, there is concurrence on the reduced perceptual saliency of certain grammatical words, especially modal and auxiliary verbs, and personal pronouns:

you shouldn't have (Ur, 1984: 46)

shoul'da (should have) (Brown and Hilferty, 1989: 27)

To maximise the trait purity of the test, three steps were undertaken. Firstly, to carefully control the linguistic difficulty of the assessed material, all of the structural forms presented were derived from a foundation level textbook (Murphy, 2003). And all of the assessed vocabulary was drawn from a list taught at all Japanese junior high schools. Secondly, to reduce the influence of semantic processing and thereby decrease the overlap between this sub-skill test and the L2 proficiency measure, the sentences in the test were decontextualised. Through decontextualization it has been found that, “students can perform very well on a dictation test, and yet have very little understanding of the gist of what they have written down” (Dirven & Oakeshott-Taylor, 1985: 14). Thirdly, to ensure that the participants did not have to call upon their L2 linguistic knowledge, the assessed sentences were very short. As noted by Buck, “...when segments are very short, and they do not challenge the test-taker, writing down a few words of spoken text is little more than a simple transcription exercise. The listening skills involved are probably just word recognition” (2001: 77). When marking the tests, one point was awarded for each correctly identified assessed word. For more details on this test see Joyce (2014).

Working Memory (English and Japanese tests).

Since working memory (WM) capacity has been found to be heavily dependent on whether the target language is in the participants’ L1 or L2 (Cook, 1977), both English and Japanese random digit tasks were used. After hearing a series of digits, the participants were tasked with reproducing the sequence of numbers.

L2 Phonological Awareness.

L2 Phonological Awareness was operationalised through a minimal pair (AX) phonemic discrimination test format. The research instrument focused on six phonemic contrasts that have been deemed particularly difficult for Japanese learners to differentiate (Kenworthy, 1987). During this task, the participants

listened to 80 different word pairs. There were 60 pairs containing words that differed by one phoneme and 20 identical distracter pairs.

L2 Sentence Stress Awareness.

The L2 sentence stress awareness task involved the participants listening to a series of 30 short decontextualised sentences. As well as hearing the texts, the material was also printed in the test booklet. After listening to each of the sentences, the participants selected which of the lexical items held the main stress. For each sentence, there were five possible answer choices and the word that was selected to receive the primary stress was chosen at random. To ensure that syntactic knowledge did not become a contaminating factor, the test sentences were drawn from a foundation level L2 English grammar book (Murphy, 2003). Furthermore, all of the vocabulary in the test was drawn from either the 1000 most frequent English words or the core vocabulary syllabus taught to all Japanese junior high school students. To ensure that the test was fair, the items were independently validated by a group of native English speakers.

L2 Metacognitive Listening Strategies Usage.

L2 Metacognitive Listening Strategy Usage was evaluated through an established instrument, the Metacognitive Awareness Listening Questionnaire (MALQ) (Vandergrift, 2005). The questionnaire contains a series of statements each corresponding to a strategy associated with successful L2 listening comprehension. It was developed with reference to previous research in the field (e.g. Bacon 1992; Goh, 1998, 2000; O'Malley & Chamot, 1990; Vandergrift, 1998, 2003) and has been found to be reliable and unidimensional (Ehrich & Henderson, 2018). The participants indicated the frequency of their strategy usage through a five-point scale. The questionnaire was translated into the learners' L1. After piloting, it was determined that a revised 17 item version of the MALQ would be used in the main study.

Data Collection

The various research instruments were administered over three sessions. Typically, the first session covered the main test battery, which was completed within a 90-minute class period. To control for any possible sequencing effect, the order of the tests was counterbalanced. The second session comprised the TOEFL listening test and the Metacognitive Awareness Listening Questionnaire (MALQ). The in-house proficiency test was completed last and was part of the students' institutional course requirements. To allow a valid comparison between the various test scores, it was essential that there was minimal learning between the test administrations. Therefore, all the data collection sessions were conducted within a period of three weeks.

Data Processing

To enable a comparison of the participants across different tests, the Rasch model was used to convert the raw scores into IRT ability estimates. However, since IRT values are often negative, which are both unintuitive and unsuited to some statistical techniques, the logit figures were transformed into positive scores that were centred on a mean of 50. To generate a less method-dependent estimate of general listening comprehension ability, the participants' scores on in-house and TOEFL tests were combined. To obtain a composite score for each learner, the IRT person ability estimates on the two tests were first separately calculated, transformed, and then averaged. An established procedure for deriving a reliability estimate for the combined test was used (see Evans, 1996).

Results

Background Data

All of the 443 students undertook all of the different research instruments. The transformed IRT scores for the various research instruments can be found in Table 1. For the full dataset, the Cronbach's alpha reliability values for the test scores were

found to range between .74 and .87, while the Rasch person reliability estimates fell between .75 and .84. To determine whether there were meaningful differences in the performance of the higher and lower proficiency groups on each of the measures, independent t-tests were conducted. There was found to be a statistically significant difference between the two proficiency groups on all of the criterion and explanatory variables ($p < .05$).

Table 1. *Summary of the Test Results*

Test	Level	k	mean	SD	min.	max.
In-house & TOEFL	Higher	39 + 50	58.75	4.00	53.75	71.05
	Lower		45.26	2.71	33.95	48.65
Syntactic Knowledge	Higher	60	56.33	5.18	40.50	71.00
	Lower		44.18	3.64	31.80	53.50
Vocabulary Breath	Higher	40	64.04	11.52	29.90	90.50
	Lower		58.13	9.56	32.00	90.50
PMK	Higher	69	62.36	9.47	42.80	87.70
	Lower		49.93	7.05	29.40	64.60
WM (English)	Higher	16	55.10	14.85	15.40	84.70
	Lower		52.28	14.24	15.40	84.70
WM (Japanese)	Higher	16	44.09	14.16	14.30	76.30
	Lower		38.07	15.73	14.30	84.70
Phonological Awareness	Higher	60	61.12	9.64	43.80	94.50
	Lower		57.01	7.09	43.80	94.50
Sentence Stress	Higher	30	64.68	10.84	41.70	87.40
	Lower		55.70	10.73	31.30	79.90
Metacognitive Usage	Higher	17	58.11	5.46	42.20	73.80
	Lower		54.65	5.17	40.00	70.60

Since the participants were divided by their proficiency scores, there were substantial differences in the main criterion variable; the combined in-house and TOEFL tests scores for the two groups (Higher; $Mean = 58.75$, $SD = 4.00$; Lower; $Mean = 45.26$, $SD = 2.71$, $t(294) = 34.01$, $p < .001$, $d = 3.95$). In terms of the explanatory variables, the sub-skills that showed the greatest difference in scores between the higher and lower proficiency

groups were syntactic knowledge (Higher; *Mean* = 56.33, *SD* = 5.18; Lower; *Mean* = 44.18, *SD* = 3.64, $t(294) = 16.25$, $p < .001$, $d = 2.71$) and PMK (Higher; *Mean* = 62.36, *SD* = 9.47; Lower; *Mean* = 49.93, *SD* = 7.05, $t(294) = 12.81$, $p < .001$, $d = 1.49$). Since the full data-set was divided into three groups based on proficiency level, there was a restricted range of sampling and a consequential reduction in the score variance. As a result, there was necessarily a reduction in the covariation between the criterion and explanatory variables. However, through correcting for attenuation, it was possible to compensate for the measurement error and estimate the correlations between the true scores. Through this means, a more balanced comparison between the higher and lower proficiency participants has been provided. Nevertheless, the results should be viewed as exploratory.

The full correlation matrixes for the higher and lower proficiency sub-groups are displayed in Tables 2 and 3. The numbers in the top row refer to the different tests that are displayed in the first column. The raw Pearson correlations are below the diagonal, and the disattenuated correlations are above it.

Table 2. Higher Sub-Group Correlations ($n = 148$)

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. KEPT-TOEFL	-	.91	.10	.80	.08	.10	.08	.28	.14
2. Syntactic Know.	.60***	-	.28	.61	.16	.31	.23	.35	.14
3. Vocabulary Br.	.07	.22**	-	.18	.21	.19	.10	.09	-.06
4. Phon. Mod. Know.	.56***	.49***	.15*	-	.19	-.01	.24	.29	.18
5. Eng. Memory	.05	.13	.17*	.15*	-	.63	.14	.22	-.09
6. Jpn. Memory	.07	.24**	.15*	.00	.49***	-	.06	.08	-.05
7. Phon. Awareness	.06	.19*	.09	.21**	.12	.05	-	.25	.15
8. Sent. Stress Aw.	.19**	.27***	.08	.24**	.18*	.07	.21**	-	.04
9. Metacog. Strats.	.09	.10	-.04	.14*	-.07	-.03	.11	.04	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$ (one-tailed)

Table 3. Lower Sub-Group Correlations ($n = 148$)

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. KEPT-TOEFL	-	.92	.17	.57	.22	.52	.45	.54	.33
2. Syntactic Know.	.34***	-	.23	.54	.27	.25	.13	.31	.07
3. Vocabulary Br.	.07	.18*	-	-.02	.05	.01	.08	.04	.02
4. Phon. Mod. Know.	.22**	.39***	-.02	-	.21	.34	.34	.28	.20
5. Eng. Memory	.09	.19*	.04	.16*	-	.75	.07	.32	-.01
6. Jpn. Memory	.21**	.19*	.01	.27**	.59***	-	.06	.40	-.03
7. Phon. Awareness	.17*	.09	.07	.26**	.05	.04	-	.05	.16
8. Sent. Stress Aw.	.21**	.23**	.03	.21**	.25**	.32***	.04	-	.13
9. Metacog. Strats.	.13	.05	.01	.16*	-.01	-.02	.12	.10	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$ (one-tailed)

The covariances between the variables bear important similarities and differences. Regarding the similarities, in both cases, L2 syntactic knowledge formed the strongest relationship with the criterion variable. For the higher proficiency participants, the disattenuated correlation was .91 ($r = .60$, $p < .001$), and for the lower proficiency group it was .92 ($r = .34$, $p < .001$). Furthermore, for both sets of learners, the second strongest interaction with L2 listening proficiency involved PMK. However, for the higher ability group the corrected correlation stood at .80 ($r = .56$, $p < .001$), while for the lower proficiency learners the equivalent value was a rather more modest .57 ($r = .22$, $p < .01$). There was also found to be a substantial overlap between L2 syntactic knowledge and PMK for the higher group of .61 ($r = .49$, $p < .001$) and the lower group of .54 ($r = .39$, $p < .001$).

There were also found to be some striking differences between the two sets of results. Across the proficiency groups, the correlations for WM, phonological awareness and sentence stress awareness for the lower proficiency group were higher than for the higher proficiency students. In the case of WM (Japanese), the disattenuated correlation of .52 ($r = .21$, $p < .01$) for the lower proficiency group, greatly exceeded that of the higher group (.08, $r = .07$, n.s.). A similar pattern emerged for the phonological

awareness sub-skill. The lower proficiency group recorded a corrected correlation with L2 proficiency of .45 ($r = .17, p < .05$), while the equivalent value for the higher proficiency group was .08 ($r = .06, n.s.$). Finally, the disattenuated correlation between the criterion variable and sentence stress awareness was .54 (.21, $p < .01$) for the lower proficiency learners. This is in contrast to the value of .28 (.19, $p < .01$) for the higher proficiency listeners. As will be considered in the discussion section, it is notable that the lower proficiency learners rely more heavily upon the sub-skills that lie closest to the acoustic signal. Since these components form constituent parts of L2 listening, it is reasonable to hypothesize that knowledge of these sub-skills has an important role in rendering the L2 input comprehensible.

Discussion and Conclusion

As has been discussed, the aural sub-skills with the greatest disattenuated correlations with the criterion variable have proven to be syntactic knowledge and PMK. Thus, the findings from this study suggest that proficiency in these sub-skills is central to L2 listening comprehension. Although these linguistic variables were found to be important to both proficiency groups, higher proficiency learners displayed far greater knowledge of them. It is considered that higher proficiency listeners primarily benefit from their superior knowledge in two main ways. Firstly, they can use their understanding of reduced forms to locate the word boundaries in the speech stream. This is extremely important as individual words in continuous speech are often highly indistinct owing to extensive phonological modification. Through better identifying the lexical items, higher proficiency listeners were provided with syntactic and semantic constraints that aided in the recognition of subsequent words and enabled them to develop expectations of the forthcoming language. Secondly, as higher proficiency listeners were better able to parse the smaller linguistic units into larger entities, they are able to process the language into larger chunks than lower proficiency listeners. As a result, they were capable of developing a more

coherent representation of the speech stream from which they could apply their higher-level processing skills to enrich their understanding. Therefore, through the facilitation of their syntactic knowledge and PMK, higher proficiency learners benefit from a complex and closely entwined interplay between their higher and lower level processes. As a result, information at a whole range of levels is simultaneously accessed to carefully predict, check, and monitor the listener's developing interpretation.

Conversely, in the case of the lower proficiency learners, their relatively poor familiarity with grammar and connected speech were serious hindrances to their aural understanding. Due to insufficient knowledge of the reduced forms, the learners were limited in their ability to segment the language they heard. As a result, they had difficulty applying their grammatical and pragmatic knowledge to anticipate and parse the language that they heard. This inability to efficiently parse the language necessitated a heavy reliance upon the perceptual stage of understanding to grasp the content of the speech. Thus, as has been found in previous studies (e.g. Hansen & Jensen, 1994; Shohamy & Inbar, 1991), less proficient listeners were found to be highly dependent on the linguistic and psycholinguistic sub-skills that lay closest to the surface of the message, such as phonological awareness, and working memory (Japanese). Due to this shallow processing style, less proficient listeners were forced to base their understanding on associations between content words rather than processing across clauses (Taylor & Taylor, 1990).

In terms of the phonological awareness results, for the higher proficiency group, there was found to be a very low corrected correlation between L2 listening proficiency and the phonological awareness scores (.08, $r = .06$, n.s.). In contrast, the lower proficiency learners recorded much greater values (.45, ($r = .17$, $p < .05$)). When considering this difference, it is worth noting that since the perceptual realisation of the phoneme varies according to its immediate phonetic environment (Liberman,

1970), the acoustic signal is not a firm foundation upon which processing can be built. To compensate for the deficiencies in the clarity of the speech stream, more proficient L2 listeners apply their higher level semantic, syntactic, and lexical knowledge to help predict, segment and decode words. On the other hand, due to their shallow processing of L2 aural material, less proficient listeners have to accurately select the target item in the speech stream from multiple similarly sounding word candidates. As well as placing a strain upon their L2 phonological awareness sub-skill, this also places heavy demands upon their WM. Thus, consistent with previous research (Byrnes, 1984; Call, 1985), the L2 phonological awareness results indicate that efficient word recognition is facilitated by the interaction between sound and multiple information sources.

In regard to the purer measure of WM (WM: Japanese), the lower ability group recorded a much greater disattenuated correlation with the in-house-TOEFL measure (.52 ($r = .21$, $p < .01$)) than the higher ability listeners (.10 ($r = .07$, n.s.)). It is possible that lower ability listeners were more dependent on WM as they had a marginally smaller memory capacity. However, this is considered unlikely as both higher and lower proficiency listeners from this study displayed an average digit memory span in the normal adult range (Gernsbacher, 1980). Instead, it is posited that the difference in the correlations was due to the lower group being forced to reanalyze the speech stream while it was in their echoic memory to a greater extent than the higher listeners. That is, due to their inability to efficiently segment and parse the language into meaningful units, the lower group was more dependent on the sensory traces in memory. Furthermore, owing to their need to divert cognitive resources towards the reanalysis of the auditory material held in their WM, the less proficient listeners were likely to have been further hampered by a lack of attentional capacity to construct a global understanding of the text.

As has been discussed, due to a lack of consensus amongst researchers on the relative importance of the various aural sub-

skills, there has been limited agreement on the most advantageous approach to L2 listening proficiency development. For instance, it has been variously suggested that teachers and learners should primarily focus on lexical expansion (Kelly, 1991), phonemic awareness (Reynolds, 1998), or listening strategies (Goh, 2000). This study has clarified that comprehension is most closely associated with knowledge of grammar and phonological modification for both higher and lower proficiency L2 listeners. Therefore, it is logical that through an increased pedagogic and learning focus on these sub-skills, L2 listening ability will develop with greater efficiency. When considering the teaching of lower proficiency students, the findings could be used to advocate an emphasis on phonological awareness and working memory (Japanese). However, as has been discussed, it is considered that less proficient listeners' dependence on the lowest level sub-skills is symptomatic of their deficiency in their syntactic understanding and PMK. Therefore, when considering which aural sub-skills lower level listeners should concentrate on, it is advised that instructors retain a focus on grammar and connected speech.

Overall, the picture of listening that has emerged both reinforces and adds detail to the process of L2 listening comprehension that is portrayed in the literature (e.g. Anderson & Lynch, 1988; Jensen & Hansen, 1995). Aural comprehension is a highly complex interactive process that involves successful listeners harmoniously relating linguistic information and higher-level inferences to generate meaning. Crucially, it is postulated that this symbiotic relationship is enabled by the development of the listeners' syntactic knowledge and the ability to recognise words in connected speech. On the other hand, there is a clear implication that owing to insufficient syntactic resources and L2 phonological modification knowledge, less proficient listeners process texts in a shallower manner. That is, they are forced to place a heavy reliance on the lowest level processing skills, which tend to merely yield fragments of understanding. When this proves inadequate, less proficient listeners become reliant on higher level schematic knowledge to bridge the gap to meaning (Jensen &

Hansen, 1995; Tsui & Fullilove, 2004). However, due to the impoverished nature of the lower level guidance to the higher-level processing, the application of their prior knowledge is often little more than guesswork. Thus, in summary, the results suggest that a generalized knowledge of grammar and connected speech is essential to the facilitation of L2 listening comprehension.

The Author

Paul Joyce is an Associate Professor at Kindai University in Osaka, Japan. He is from Britain. He undertook his M.A. (TESL/TEFL) at Birmingham University (England) and conducted his Ph.D. in Applied Linguistics at Surrey University (England). His research interests include vocabulary acquisition, formative assessment, and second language listening.

References

- Alderson, J. C. (1993). The relationship between grammar and reading in an English for academic purposes test battery. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research: Selected papers from the 1990 Language Testing Research Colloquium*. Alexandria, VA: TESOL.
- Anderson, A. & Lynch, T. (1988). *Listening*. Oxford: Oxford University Press.
- Anderson, J. R. (1995). *Cognitive psychology and its implications*, (4th Edition). Freeman, New York.
- Bacon, S. M. (1992). Phases of listening to authentic input in Spanish: A descriptive study. *Foreign Language Annals*, 25, 317-334.
- Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, 14-31.
- Brown, J. & Hilferty, A. (1986). Listening for reduced forms. *TESOL Quarterly*, 20(4), 759-763.
- Brown, J. & Hilferty, A. (1989). Teaching reduced forms. In *Gendai Eigo Kyoiku* (January issue), 26-28.

- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Byrnes, H. (1984). The role of listening comprehension: A theoretical base. *Foreign Language Annals*, 17, 317-329.
- Call, M. E. (1985). Auditory short-term memory, listening comprehension, and the input hypothesis. *TESOL Quarterly*, 19(4), 765-781.
- Cheng, J. & Matthews, J. (2018). The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing*, 35(1). 3–25.
- Cook, V. J. (1977). Cognitive processes in second language research. *IRAL*, 15(1), 1-20.
- Dirven, R. & Oakeshott-Taylor, J. (1985). Listening comprehension (Part 2). *Language Teaching*, 18, 2-20.
- Ehrich, J. F., & Henderson, D. B. (2018). Rasch analysis of the Metacognitive Awareness Listening Questionnaire (MALQ). *International Journal of Listening*. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/10904018.2017.1418350>
- El-dali, M. (2017). L2 Listening Comprehension: Is it a Language Problem or Listening Problem? *English Linguistics Research*, 6(1), 14-37.
- ELI LCT Manual* (1986). English Language Institute: The University of Michigan.
- Evans, L. D. (1996). Calculating achievement composite scores for regression discrepancy models. *Learning Disability Quarterly*, 19(4), 242-249.
- Feyton, C. (1991). The power of listening ability: An overlooked dimension in language acquisition. *The Modern Language Journal*, 75(ii), 173-180.
- Frauenfelder, U. & Komisarjevsky-Tyler, L. (1987). The process of spoken word recognition: An introduction. *Cognition*, 25, 1-20.
- Gernsbacher, M. (1990). *Language comprehension as structure building*. Hillsdale: Erlbaum.

- Goh, C. C. M. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2, 124-147.
- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28, 55-75.
- Goh, C. C. (2017). Cognition, metacognition, and L2 listening. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning*, Volume 3 (pp. 214-228). Routledge.
- Graham, S. (2003). Learner strategies and advanced level listening comprehension. *Language Learning Journal*, 28, 64-69.
- Hansen, C. & Jensen, C. (1994). Evaluating lecture comprehension. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 241-268). Cambridge: Cambridge University Press.
- Harrington, M. & Sawyer, M. (1992). L2 working memory capacity and the L2 reading skill. *Studies in Second Language Acquisition*, 14, 25-38.
- Ito, Y. (2001). Effect of reduced forms on ESL learners' input-intake process. *Second Language Studies*, 20(1), 99-124.
- Jensen, C. & Hanson, C. (1995). The effect of prior knowledge on EAP listening performance. *Language Testing*, 12(1), 99-119.
- Joyce, P. (2011). Componentiality in L2 listening. In B. O'Sullivan (Ed.), *Language testing: Theory and practice*. (pp. 71-93). Palgrave MacMillan, London.
- Joyce, P. (2014). The Development and Validation of a Second Language Listening Reduced Forms Test. *Iranian Journal of Language Testing*, 4(2), 155-174.
- Kelly, P. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced foreign language learners. *IRAL*, 29(2), 135-149.
- Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.
- Krohn, R. (1971). *English Sentence Structure*. University of Michigan Press.

- Kuehn, P. (1993). *Review of the listening comprehension test*. Buros Institute of Mental Measurements Test Reviews Online.
- Lado, R. & Fries, C. C. (1958). *English Practice Patterns*. University of Michigan Press.
- Liberman, A. M. (1970). The grammars of speech and language. *Cognitive Psychology*, 1, 301-323.
- Long, M. (1985). Input and second language acquisition theory. In S, Gass & C, Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, MA: Newbury House.
- Mack, M. (1988). Sentence processing by non-native speakers of English: Evidence from the perception of natural and computer-generated anomalous L2 sentences. *Journal of Neurolinguistics*, 3, 293-316.
- Mora, J. C. (2005). Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and non-words by Spanish/Catalan bilingual learners of English. *Proceedings of the ISCA workshop in speech perception*. London, UK.
- Murphy, R. (2003). *Essential grammar in use with answers: A self-study reference and practice book for elementary students of English*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- O'Malley, J. M., & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Revesz, A., & Brunfaut, T. (2013). Text characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition*, 35(1), 31-65.
- Reynolds, B. (1998). Phonemic awareness: Is it language specific? *Literacy across Cultures* 2/2.
- Rost, M. (1990). *Listening in language learning*. New York: Longman.

- Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal*, 78(ii), 199-221.
- Shohamy, E. & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question-type. *Language Testing*, 8, 23-40.
- Stanovich, K. E. (1980). Towards an interactive compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16(1): 32-72.
- Taylor, I. & Taylor, M. M. (1990). *Psycholinguistics: Learning and using language*. Englewood Cliffs, NJ: Prentice Hall.
- Tsui, A. B. M. & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics*, 19(4), 432-451.
- Ur, P. (1984). *Teaching listening comprehension*. Cambridge: Cambridge University Press.
- Vandergrift, L. (1998). Successful and less successful listeners in French: What are the strategy differences? *The French Review*, 71, 370-395.
- Vandergrift, L. (2003). Orchestrating strategy use: Towards a model of the skilled L2 listener. *Language Learning*, 53, 461-494.
- Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness, and proficiency in L2 listening. *Applied Linguistics*, 26(1), 70-89.
- Vandergrift, L. & Baker, S. (2018). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, 74(1), 79-100.
- Vanderplank, R. (1985). Isochrony and intelligibility. In J. Tommola & K. Battarbee (Eds.). CDEF 84. *Papers from the conference of Departments of English in Finland* (pp. 301-328). Turku, Finland: Publications of the Department of English.
- Weinstein, N. (1982). *Whaddaya Say?*. Culver City, California: ESL Publications.

- Witkin, B. R. (1990). Listening theory and research: The state of the art. *International Journal of Listening*, 4, 7-32.
- Wolff, D. (1987). Some assumptions about second language text comprehension. *Studies in Second Language Acquisition*, 9, 307-326.

Appendix 1

Sample Questions

L2 Syntactic Knowledge.

Test takers heard: *Who brought Bill his lunch today.*

- Printed test item:
- (a) His sister is.
 - (b) His sister did.
 - (c) To his sister.
 - (d) His sister had.

L2 Vocabulary Breadth.

Test takers heard: *interrupt...interrupt*

Printed test item:

1. a. 折る b. さえぎる c. 囲む d. そらす e. はぐ

Translation of test item: (not provided to test takers)

- a. fold b. interrupt c. surround d. divert e. strip

L2 Phonological Modification Knowledge (PMK).

Three sentences test takers transcribed:

You would not tell him.

Do you want to do it?

He has to go with them

L2 Phonological Awareness.

Focusing on the /l/-/r/ contrast, the final test included:

Test takers heard: *plod...prod*

Printed test item: same different

L2 Sentence Stress Awareness.

Test takers heard: Do you want to play *tennis*?

Printed test item: Do you want to play tennis?

- (A) (B) (C) (D) (E)

L2 Metacognitive Listening Strategies Usage.

Sample item:

As I am listening, I predict what will happen.