

Original Paper

L2 Proficiency and L2 Developmental Stages:

A Learner Corpus Analysis

Yumiko Yamaguchi^{1*}

¹ International Education Center, Tokai University, Japan

Received: October 23, 2019 Accepted: November 5, 2019 Online Published: November 27, 2019

doi:10.22158/selt.v7n4p516

URL: <http://dx.doi.org/10.22158/selt.v7n4p516>

Abstract

This paper presents part of the results of a learner corpus study of English oral and written production by a large number of Japanese native speakers. Each participant was asked to perform two tasks, namely spoken and written narratives, using a picture book titled “Frog, where are you?” (Mayer, 1969) containing 24 wordless pictures. For the analyses in the current study, the data from 80 learners, focusing on audio-recorded and transcribed spoken narratives, was used. The Japanese learners’ speech production was examined based on Processability Theory (PT; Pienemann, 1998, 2005; Bettoni & Di Biase, 2015) as well as on the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001). Results show that there is a correlation between second language (L2) proficiency levels and L2 developmental stages in a learner corpus of L2 spoken English. On the other hand, the dispersion is found to increase at higher stages as shown in previous studies (e.g., Granfeldt & Ågren, 2013; Hagenfeld, 2017).

Keywords

L2 proficiency, L2 development, CEFR, processability theory, learner corpus

1. Introduction

The aim of the present paper is to investigate empirically the relationship between second language (L2) proficiency as measured by the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001) and L2 developmental stages hypothesized by Processability Theory (PT), a theory of second language acquisition (SLA). The CEFR offers a comprehensive description about “what language learners have to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively” (2001, p.1). It characterizes language learners’ communicative proficiency at six levels, including A1, A2, B1, B2, C1 and C2. In the CEFR ratings, learners at the A1/A2 levels can be regarded as Basic Users, those at the B1/B2 levels can be

considered Independent Users, and those at the C1/C2 levels are generally thought of as Proficient Users. The CEFR has been widely used as a common reference tool for SLA, and found useful in syllabus construction, curriculum coordination, and in the preparation of teaching materials and examinations. However, few empirical studies with L2 learner data have been done on the CEFR levels (e.g., Hulstijn, 2011). In particular, little is known about how the CEFR levels of L2 proficiency are related to the L2 developmental stages as predicted in various SLA theories (e.g., Granfeldt & Ågren, 2013; Hagenfeld, 2017; Yamaguchi, 2019).

In one major SLA theory, Processability Theory (PT; Pienemann, 1998; Pienemann, Di Biase, & Kawaguchi, 2005), a universal hierarchy of L2 development is assumed to exist. Based on Levelt's (1989) Speech Model and Lexical Functional Grammar (LFG; e.g., Bresnan, 2001), PT hypothesizes as to the learners' developmental stages regarding the acquisition of grammatical structures, including morphology and syntax. In 2005, PT was expanded to include hypotheses regarding the acquisition of syntactic structures (Pienemann, Di Biase, & Kawaguchi, 2015) in accordance with the development of LFG.

One of the hypotheses in this PT extension (Pienemann, Di Biase, & Kawaguchi, 2015), the Lexical Mapping Hypothesis (The LMH; Pienemann, Di Biase, & Kawaguchi, 2015), focuses on the development of argument mapping between thematic roles and grammatical functions in sentence construction. Table 1 presents the developmental stages of English syntactic structures hypothesized in the LMH.

Table 1. Developmental Stages for English Syntax Based on the Lexical Mapping Hypothesis (after Pienemann, Di Biase, & Kawaguchi, 2005)

STAGE	STRUCTURE	EXAMPLE
4. NON-DEFAULT MAPPING	passive	<i>The ball was kicked by Lisa</i>
	causative	<i>Dad made Bob wash the car</i>
3. DEFAULT MAPPING + ADDITIONAL ARGUMENT	Ditransitive	<i>Janet gave Ben a present</i>
	Canonical sentence + Oblique argument	<i>Tim put the book on the desk</i>
2. DEFAULT MAPPING	Canonical word order e.g., agent-event-patient	<i>Lisa kicked a ball</i>
1. LEMMA ACCESS	single words	<i>Look</i>
	formulas	<i>Thank you</i>

In the LMH, L2 learners are assumed to start constructing sentences using “default mapping” when they become able to produce utterances of more than one word. In “default mapping”, the highest available role in the thematic hierarchy, the Agent, is mapped onto the Subject (SUBJ) grammatical

function. The sample sentence (1) shows typical “default mapping” with a transitive verb “kick”, which requires two arguments.

(1) *Lisa kicked a ball*

In the sentence (1), the most prominent role, the Agent “Lisa” is mapped onto the SUBJ and the less prominent role, the Patient “the ball” is mapped onto the Object (OBJ). Many scholars, such as Pinker (1984) and Slobin (1985), also claim that beginning learners map the most prominent thematic role onto the SUBJ.

It is assumed that learners gradually learn how to direct the listener’s attention to a particular thematic role lower in the hierarchy by promoting it to the SUBJ and de-focusing the highest role by mapping it onto a grammatical function other than the SUBJ, or by suppressing it. A typical case of non-default mapping is the Passive. In the sample sentence (2), the Patient “the ball” is mapped onto the most prominent grammatical function, SUBJ, while the highest thematic role, the Agent, is suppressed, and appears as Adjunct, “by Lisa”. L2 learners are predicted to become able to produce “non-default mapping” which requires a much higher processing cost only after “default mapping” is in place.

(2) *The ball was kicked by Lisa*

2. Previous Studies

The developmental stages for morphology and syntax hypothesized in PT have been tested in various L2 contexts (e.g., Bettoni & Di Biase, 2015, Keßler, Lenzing, & Liebner, 2016; Lenzing, Nicholas, & Roos, 2019; Pienemann, 1998; Pienemann, Di Biase, & Kawaguchi, 2005). On the other hand, the LMH in PT has been tested mainly in Japanese L2 (Kawaguchi, 2005, 2007, 2008, 2009a, 2009b, 2016). Regarding English L2 contexts, the development of the passive construction has been the focus in both longitudinal (e.g., Di Biase, Kawaguchi, & Yamaguchi, 2015) and cross-sectional study (e.g., Keatinge & Keßler, 2009; Wang, 2009) research.

For instance, Di Biase, Kawaguchi and Yamaguchi (2015) tested the LMH using data from Yamaguchi’s (2013) two-year longitudinal study of a Japanese primary school child learning English in Australia. They found that the child first acquired default mapping, and gradually added one more argument after default mapping, and finally started producing non-default mapping, such as passives. While Di Biase, Kawaguchi, and Yamaguchi (2015) showed support for the LMH, their study finished just after the child started producing non-default mapping.

While the relationship between L2 proficiency measured by the CEFR and L2 developmental stages hypothesized in SLA theory has not yet been investigated extensively using learner data (e.g., Hulstijn, 2011), some recent research has reported a positive connection between the CEFR levels and the developmental stages predicted in PT (Granfeldt & Ågren, 2013; Hagenfeld, 2017; Yamaguchi, 2019).

Granfeldt and Ågren (2013) examined written data produced by 38 Swedish speakers learning French as a third language (L3). In their study, the development of morphosyntax in L3 French was analyzed based on PT and the communicative proficiency was measured by two CEFR raters. While they found

a strong connection between the CEFR rating and the PT analysis, a dispersion at more advanced stages was shown to exist. Nevertheless, they claim that learners' communicative proficiency up to the CEFR B1 level and morphosyntactic development up to PT's stage 3 seem to develop more or less at the same rate.

Hagenfeld (2017), who examined the CEFR-based rating of 21 novice and 10 expert raters on 14 original and 8 edited files of the same speech samples, also found a positive correlation between the CEFR levels and PT stages, again, only at lower levels. Yet, these studies show some connection between L2 proficiency and L2 developmental stages for morphosyntactic structures.

More recently, Yamaguchi (2019) examined the relationship between the CEFR levels and PT stages using a learner corpus of English speaking and writing by 60 Japanese speakers. Yamaguchi (2019) examined the Japanese learners' use of English syntax based on the LMH in PT extension (Pienemann, Di Biase, & Kawaguchi, 2005) and compared L2 developmental stages found in PT analysis with L2 proficiency level measured in the CEFR ratings. The results indicate that there seems to be a connection between L2 proficiency levels and L2 developmental stages for English syntactic structures. However, since a majority of the learners at the highest PT stages for English syntax (i.e., stage 4) were rated as Basic Users (i.e., A2 level) in the CEFR rating, it was suggested that English L2 learners at the highest PT stage were not necessarily at higher CEFR levels.

In order to explore the possible relationship between L2 proficiency and L2 development in more details, the present study addresses the following research questions.

1. To what extent are L2 proficiency levels as measured by the CEFR rating and L2 developmental stages as found in PT analysis related in a corpus of L2 spoken English by Japanese native speakers?
2. Does the dispersion increase at more advanced stages as found in previous PT research?

3. Methodology

In this study, 80 participants' speaking data were randomly chosen from the Japanese Learner Corpus of English Narratives (JaLCEN) constructed by Yamaguchi and Usami (2017a, 2017b). JaLCEN consists of 946 files of English Spoken and Written narratives by 473 Japanese L1 speakers, aged 18-30. For the data collection, each participant was asked to perform both spoken and written narratives using a picture book called "*Frog, where are you?*" (Mayer, 1969). This picture book contains 24 wordless pictures and has been extensively used in a number of language acquisition studies. Half of the learners (i.e., 40 learners) were asked to start with spoken narratives and the other half to start with written narratives to minimize the ordering effects. Their speech production was audio-recorded and transcribed, and their writing was recorded with pen and paper by the learners.

As for the data analysis, the participants' PT stages for English L2 syntax are examined based on the LMH in PT. Their CEFR levels are measured by two experienced raters. While most previous studies examined L2 development based on accuracy, PT has applied the emergence criterion. According to PT,

“emergence can be understood as the point in time at which certain skills have, in principle, been attained or at which certain operation can, in principle, be carried out” (Pienemann, 1998, p. 138). PT claims that using a grammatical structure at a high level of accuracy, even 80% to 90%, does not guarantee that the learner will be able to continue producing that structure at the same or higher level of accuracy in the future. The present study applies the emergence criterion in PT to determine if a target syntactic structure is acquired.

In PT analysis, default mapping, as in (3) and (4), was coded as Stage 2 structures.

(3) #10 *bees chased Tim*

(4) #54 *the boy and dog found the frog*

Then, default mapping with additional argument, as in (5) and (6), was coded as Stage 3 structures.

(5) #2 *they named their new frog friend Froggie*

(6) #8 *the dog put his head into the jar*

If a learner was able to produce sentences constructed by non-default mapping, as in (7), he/she was considered to have reached Stage 4 for English syntax.

(7) #34 *the dog was chased by many bees*

4. Results and Discussion

4.1 CEFR Levels

Table 2 summarizes the results of the CEFR rating for L2 proficiency levels found in English spoken narratives by 80 Japanese L1 speakers. According to the table, two learners were rated as A1 level, while 60 learners were regarded as A2 level. Thus, a majority of the participants in this study, namely 77.5%, are regarded as Basic Users. On the other hand, 15 learners were rated as B1 level, while 3 learners were rated as B2 level. It means that only 22.5% of the participants are regarded as Independent users.

Table 2. CEFR Levels Found in English Spoken Narratives by 80 Japanese L1 Speakers

CEFR level	A1	A2	B1	B2
n = 80	2 (2.5%)	60 (75%)	15 (18.75%)	3 (3.75%)

4.2 PT Stages

Table 3 presents the results of the PT analysis for developmental stages for English L2 syntax found in spoken narratives by 80 Japanese L1 speakers. All the 80 participants produced various sentences with default mapping. 8 learners were found to use only default mapping and they can be regarded to be at stage 2. On the other hand, 31 learners produced stage 3 structures, that is, default mapping plus one more argument, while 41 learners were found to use non-default mapping. Thus, more than half of the participants, 51.25%, can be considered to be at stage 4 for English L2 syntax. Since all the participants

at stage 4 also produced structures belonging to stage 3, it can be argued that the development of argument mapping found in this study shows implicational patterns.

Table 3. PT Stages Found in English Spoken Narratives by 80 Japanese L1 Speakers

PT stage	2	3	4
n = 80	8 (10.0%)	31 (38.75%)	41 (51.25%)

4.3 Relationship between L2 Proficiency and L2 Developmental Stages

Table 4 summarizes the correlation between L2 proficiency as measured by the CEFR raters and the L2 developmental stages as found in the PT analysis of the spoken narratives of 80 Japanese L1 speakers. As mentioned above, the CEFR ratings for the 80 participants in this study range from A1 to B2, and the results of the PT analysis covered the four stages hypothesized in PT.

According to the table, only PT stage 2 learners are rated at the CEFR A1 level. That is, no PT stage 3 or 4 learners were rated at the CEFR A1 level. PT stage 2 learners were rated either as CEFR A1 or CEFR A2 level, so they were regarded as Basic Users according to the CEFR analysis. On the other hand, only PT stage 4 learners are rated at the CEFR B2 level. It should also be noted that only PT stage 3 or 4 learners were rated at the CEFR B1 or B2 level, classifying them as Independent Users.

Table 4. Correlation of CEFR Levels and PT Stages Found in English Spoken Narratives by 80 Japanese L1 Speakers

B2	0	0	3
B1	0	5	10
A2	6	26	28
A1	2	0	0
CEFR			
	PT	2	3
		4	

Figure 1 shows the results of the analysis of the correlation between L2 proficiency and L2 development found in the CEFR ratings and PT analysis on English spoken narratives by 80 Japanese L1 speakers. In this scatterplot, the size of each circle varies according to the number of participants. The scatterplot has demonstrated that there is a linear correlation between CEFR L2 proficiency and PT L2 development of English syntax.

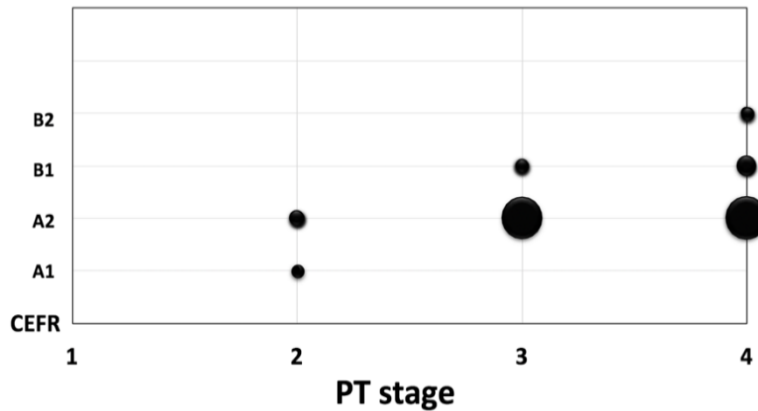


Figure 1. Correlation of L2 Proficiency and L2 Developmental Stages Found in English Spoken Narratives by 80 Japanese L1 Speakers

In order to answer the first research question, the strength of the association between the CEFR levels and PT stages was statistically assessed using a Spearman's rank order correlation analysis. The results indicate that the correlation between the levels shown in the CEFR ratings and developmental stages found in PT analysis is statistically significant at the 0.01 level ($r_{s[80]} = .319, p < 0.01$). Thus, it can be claimed that this study shows a correlation between the L2 proficiency as measured by the CEFR raters and the L2 development shown in the PT analysis. However, the statistical analysis in this study did not show as strong an association between L2 proficiency and L2 development as found in other studies such as Granfeldt and Ågren (2013) which indicated a significantly stronger correlation between the average rated CEFR level and the analyzed PT stage ($r_{s[62]} = .86, p < 0.001$) in a corpus of written L3 French. It should also be noted that while the present study analyzed English L2 speech production, Granfeldt and Ågren (2013) examined L3 French written production. Thus, further research is needed to investigate whether the discrepancy in these findings was derived from the differences in elicitation tasks, target languages, and so on.

Concerning the second research question, the results of this present study confirm that dispersion between these two methods of analysis increases at the higher stages of proficiency as found in previous studies such as Granfeldt and Ågren (2013) and Hagenfeld (2017). In particular, more than half of the PT stage 4 learners were measured as Basic Users by CEFR raters. For instance, Learner #41 produced non-default mapping twice (as in samples 8 and 9, below) and can be considered to be at stage 4 according to the emergence criterion of PT; however, the CEFR ratings for Learner #41 range from A1 to A2 (as shown in Table 5).

(8) #41 *the boy was bite, bite-d by mouse and the dog swing honeycomb, swing the tree*

(9) #41 *suddenly bird appeared the hole. the boy was dropped on the tree. honey. the dog was attacked by honey*

Table 5. The CEFR Ratings for Learner #41

Global Scale	Overall Spoken General Production	General Linguistic Range	Grammatical Accuracy	Vocabulary Range	Vocabulary Control
A2	A2	A2	A1	A2	A1

In fact, this learner often made grammatical mistakes, including wrong verb forms as in sample (8) and missing prepositions as in sample (9) “suddenly bird appeared the hole”. Also, the learner used wrong vocabulary as in sample (9), “honey” for “bees”. This suggests that the learners who often make grammatical and/or vocabulary mistakes can be rated as a Basic User in the CEFR rating system, although they produce some advanced grammatical structures. Further research is required to investigate the possible causes for the dispersion at higher stages in more detail.

5. Conclusion

The results of the analysis of the English spoken narratives by 80 Japanese L1 speakers show that there is a linear connection between their L2 proficiency level as measured by the CEFR and their L2 developmental stages as analyzed by PT. Also, as shown in previous studies (e.g., Granfeldt & Ågren, 2013), the dispersion between these two methods is found to increase at more advanced stages, in particular, PT stage 4. In addition, some learners showed the higher L2 development than L2 proficiency or vice versa. Since this study examined only speech data by 80 Japanese learners of English focusing on the development of English syntactic structures as identified in PT analysis, future research should be conducted to explore the relationship between CEFR L2 proficiency and PT developmental stages with a greater variety of learner data. In particular, the connection between L2 proficiency and L2 development was not shown to be as strong as that found in the previous research (Granfeldt & Ågren, 2013), more studies are needed to investigate whether the differences found in these studies were caused by task differences, target languages, and so on.

Acknowledgments

This work was in part supported by JSPS KAKENHI Grant Number 18K00754. The author is grateful to Hiroko Usami, Katharina Hagenfeld, Kiki Nakamura, Satomi Kawaguchi and Bruno Di Biase for their kind assistance with data collection and analysis. My gratitude also goes to Simeon Flowers for his proofreading.

References

- Bettoni, C., & Di Biase, B. (Eds.). (2015). *Grammatical development in second languages: Exploring the boundaries of Processability Theory*. European Second Language Association.
- Bresnan, J. (2001). *Lexical-functional syntax*. Malden, MA: Blackwell Publishers.

- Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Di Biase, B., Kawaguchi, S., & Yamaguchi, Y. (2015). The development of English as a second language. In C. Bettoni, & B. Di Biase (Eds.), *Grammatical development in second languages: Exploring the boundaries of Processability Theory* (pp. 85-115). European Second Language Association.
- Granfeldt, J., & Ågren, M. (2013). Stages of processability and levels of proficiency in the Common European Framework of Reference for Languages: The case of L3 French. In A. Mattsson, & c. Norrby (Eds.), *Language acquisition and use in multilingual contexts: Theory and practice* (pp. 28-38). Travaux de l'Institut de linguistique de Lund 52.
- Hagenfeld, K. (2017). *PT and the CEFR—Apples and oranges?! A theoretical and an empirical investigation of interfaces between PT and the CEFR*. Paper presented in PALA 2017. PH Ludwigsburg, Germany.
- Hulstijn, J. H. (2011) Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second language assessment. *Language Assessment Quarterly*, 8, 229-249. <https://doi.org/10.1080/15434303.2011.565844>
- Kawaguchi, S. (2005). Argument structure and syntactic development in Japanese as a second language. In M. Pienemann (Ed.), *Cross-Linguistic Aspects of Processability Theory* (pp. 253-299). Amsterdam, Benjamins.
- Kawaguchi, S. (2007). Lexical Mapping Theory and Processability Theory: A case study in Japanese. In F. Mansouri (Ed.), *Second Language Acquisition Research: Theory-Construction and Testing* (pp. 39-90). Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. (2008). Language typology and Processability Theory. In J.-U. Keßler (Ed.), *Processing Approaches to Second Language Development and Second Language learning* (pp. 89-112). Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. (2009a). Acquiring causative constructions in Japanese as a second language. *The Journal of Japanese Studies*, 29(2), 273-291. <https://doi.org/10.1080/10371390903066657>
- Kawaguchi, S. (2009b). Acquisition of non-canonical order in Japanese as a second language: The case of causative structure. In D. Keatinge, & J.-U. Keßler (Eds.), *Research in Second Language Acquisition: Empirical Evidence Across Languages* (pp. 213-239). Newcastle, UK, Cambridge Scholars Publishing.
- Kawaguchi, S. (2016). Question constructions, argument mapping, and vocabulary development in English L2 by Japanese speakers: A cross-sectional study. In J.-U. Keßler, A. Lenzing, & M. Liebner (Eds.), *Developing and Assessing Second Language Grammars across Languages*. Amsterdam, Benjamins.
- Keatinge, D., & Keßler J.-U. (2009). The acquisition of the passive voice in L2 English: Perception and production. In D. Keatinge, & J.-U. Keßler (Eds.), *Research in Second Language Acquisition:*

- Empirical Evidence Across Languages* (pp. 67-92). Newcastle, UK, Cambridge Scholars Publishing.
- Keßler, J.-U., Lenzing, A., & Liebner M. (Eds.). (2016). *Developing, modelling and assessing second languages*. Amsterdam: John Benjamins.
- Lenzing, A., Nicholas, H., & Roos, J. (Eds.). (2019). *Widening Contexts for Processability Theory*. Amsterdam: John Benjamins.
- Levelt, W. J. M. (1989). *Speaking*. Cambridge, Massachusetts: The MIT Press.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Books for Young Readers.
- Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins.
- Pienemann, M. (2005). *Cross-Linguistic Aspects of Processability Theory*. Amsterdam: John Benjamins.
- Pienemann, M., Di Biase, B., & Kawaguchi, S. (2005). Extending Processability Theory. In M. Pienemann (Ed.), *Cross-Linguistic Aspects of Processability Theory* (pp. 199-251). Amsterdam: John Benjamins.
- Wang, K. (2009). Acquiring the passive voice: Online production of the English Passive construction by Mandarin speakers. In D. Keatinge, & J.-U. Keßler (Eds.), *Research in Second Language Acquisition: Empirical Evidence Across Languages* (pp. 93-117). Newcastle, UK, Cambridge Scholars Publishing.
- Yamaguchi, Y. (2013). *Child second language learning: A study of English as a second language acquisition*. Saarbrücken, Germany: LAP Lambert Academic Publishing.
- Yamaguchi, Y. (2019). Developmental stages and the CEFR levels in foreign language learners' speaking and writing. *Studies in English Language Teaching*, 7(1), 1-13. <https://doi.org/10.22158/selt.v7n1p1>
- Yamaguchi, Y., & Usami, H. (2017a). Japanese learners' use of English grammar in speaking and writing: A Processability Approach. *Proceedings of Pacific Second Language Research Forum, 2016*, 237-242.
- Yamaguchi, Y., & Usami, H. (2017b). Plural marking in spoken and written narratives: A corpus-based study of Japanese learners of English. *International Journal of Applied Linguistics & English Literature*, 6(5), 224-234. <http://dx.doi.org/10.7575/aiac.ijalel.v.6n.5p.224>