Orthoepic Competence Descriptors in Japanese Language Education: CEFR Levels B1 to C2

ITO Hideaki
University of Tsukuba, Japan
ito.hideaki.gb@u.tsukuba.ac.jp

Abstract

The Council of Europe's 2001 Common European Framework of Reference for Languages (CEFR) has shown rapid global adoption, and now includes Japanese language education though it primarily aimed at alphabetically transcribed languages. It basically acknowledges that orthoepic competence relates to comprehension of characters yet does not indicate descriptors. Descriptors examining A1 and A2 levels, using altered techniques, have already been set. In this paper, I re-examine descriptors for levels B1 to C2, which have not yet been attempted, and combine them with the results for levels A1 and A2 to present descriptors for levels A1 to C2 in overall.

Keywords: orthoepic competence; letter knowledge; CEFR; co-occurrence relations

Povzetek

Skupni evropski referenčni okvir Sveta za jezike iz leta 2001 je bil precej hitro sprejet globalno, tudi na prodročju poučevanja japonskega jezika, čeprav je bil primarno namenjen latiničnim jezikom. Okvir predvideva, da se ortopska kompetenca nanaša na razumevanje znakov, vendar ne navaja konkretnih deskriptorjev. Deskriptorji, ki preučujejo ravni A1 in A2 z uporabo spremenjenih tehnik, so že nastavljeni. V tem prispevku preučim deskriptorje na ravni med B1 in C2, ki še niso bili preskušeni, ter ugotovitve združim s tistimi na ravneh A1 in A2. Na koncu v celoti predstavim deskriptorje za vse ravni od A1 do C2.

Ključne besede: ortopska kompetenca; poznavanje črk; SEJO (CEFR); sočasni odnosi

1   Introduction

The Common European Framework of Reference for Languages (CEFR hereafter), announced in 2001 by the Council of Europe, has spread quickly throughout the world's language education community. Today, it is in the process of becoming a global language standard. As a European framework, the CEFR has particularly been

influential within European language education. Regardless of the stage of education, various European educational organizations now utilize CEFR criteria to determine course levels and degree of language ability a student should achieve upon graduating. Furthermore, the CEFR is getting accepted within Asia (including Japan) as a language education standard (Cheng, 2017). In case of Japanese language education, the expansion of its adoption has centered on the JF Standard for Japanese-Language Education (developed by the Japan Foundation) and draws upon CEFR benchmarks, which are now becoming commonplace. One example is the use of six common reference levels to describe language proficiency (Majima, 2018; Ito, 2019a). Nevertheless, the works of Meyer (2010) and North (2014) raise the problem of character comprehension when using the CEFR in Japanese language education.[1] North argues:

> In the context of current pedagogy for Japanese and Chinese it is not possible for a learner at A2 or B1 or B2 to read the types of text that appear in CEFR descriptors for the levels concerned, simply because they do not know enough signs. (p. 45)

North (2014) proceeds to suggest that:

> …using the CEFR for such languages implies either *profiling* proficiency, admitting that such learners are a higher level for listening and speaking than they are for reading and writing – which the CEFR scales will facilitate describing – or alternatively, developing completely new descriptors for reading and writing. (p. 45)

That is to say, in North's view, in particular cases there is a need to formulate entirely new descriptors for reading and writing. Länsisalmi (2012) and Shigemori Bučar, Ryu, Moritoki Škof, & Hmeljak Sangawa (2014) also point out the character/kanji issue as one of the difficulties in using the CEFR framework in Japanese teaching. Furthermore, when conducting task-based testing within Japanese language education, insufficient writing ability can cause problems with respect to completing and evaluating tasks (Kumano, Ito, & Hachisuka, 2013). If the CEFR is to be used within Japanese language education, then it is necessary to resolve issues regarding the CEFR and orthoepic competence. According to Ito (2017), the CEFR acknowledges the existence of orthoepic competence but does not indicate relevant descriptors. As a means of engaging with this problem, Ito drafted some orthoepic competency descriptors that are graded according to proficiency. However, the initial methodology for creating this draft proposal relied upon a subjective evaluation of characteristic words; Ito (2019a) thus re-examined the draft plan for the A1 and A2 orthoepic

---

[1] With respect to the Chinese Language Proficiency Test (HSK) being redesigned in 2010 on the basis of the CEFR, Meyer (2010) raises an issue regarding Chinese language education. He argues that the required number of words is too low. North (2014) subsequently expands on this point to advance a similar view in terms of Japanese language education.

competency descriptors. In this paper, I present the results of my own attempt at re-exploring levels B1 to C2, which Ito (2019a) did not investigate. Furthermore, by combining these findings with those of Ito (2019a), I show how I have been able to form orthoepic competence descriptors for levels A1 to C2.

The structure of this paper is as follows. Section 2 outlines the six stages of the CEFR's common reference levels and defines orthoepic competence. Section 3 further clarifies the research goals after reviewing prior studies related to orthoepic competence in the CEFR. Subsequently, Section 4 explains the method of analysis and its results, while Section 5 proposes some considerations drawn from those findings. Finally, Section 6 summarizes this paper and points out topics for future research.

2   Common Reference Levels and Orthoepic Competence

2.1   The Six Stages of Common Reference Levels

The Council of Europe (2001) presents six common reference levels (Figure 1), stating that: "It seems that an outline framework of six broad levels gives an adequate coverage of the learning space relevant to European language learners for these purposes" (p. 23). These common reference levels are arranged in tiers from A1 (low language proficiency) to C2 (high language proficiency). A1 or A2 indicates a "Basic User"; B1 or B2 represents an "Independent User"; and C1 and C2 signal a "Proficient User" (Ito, 2019a). However, depending on the context of its application, the CEFR also recognizes smaller divisions of B1 such as B1.1 and B1.2. As Ito (2019b) points out, there are changes in the breadth of ability expected and the abilities that are emphasized for each level. From the above, we can see that A1, A2, B1, B2, and C1, C2 do not actually express sharply delineated stages. Rather, it is more appropriate to view them as rendering the level-like nature of language proficiency into a more easily understandable form.
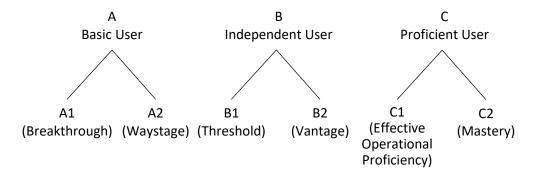

Figure 1: The Common Reference Levels (Council of Europe, 2001, p. 23)

2.2   Orthoepic Competence

The CEFR defines communicative language skills as "those which empower a person to act using specifically linguistic means" (Council of Europe, 2001, p. 9). It divides these abilities into three subgroups: (1) linguistic, (2) sociolinguistic, and (3) pragmatic. The first subgroup is composed of general linguistic range, vocabulary range, vocabulary control, grammatical competence, phonological competence, orthographic competence, semantic competence, and orthoepic competence.

With respect to orthoepic competence (the focus of this paper), the CEFR states that: "users required to read aloud a prepared text, or to use in speech words first encountered in their written form, need to be able to produce an accurate pronunciation from the written form" (Council of Europe, 2001, p. 117). Bellassen and Zhang (2008) define this skill slightly differently, as "the ability of the language user to accurately read aloud and pronounce a text or speech in a loud voice" (语言使用者在大声朗读文章或演讲稿时面对书写形式而体现出的朗读和发音能力) (p. 68).

After confirming the CEFR and the observations of Bellassen and Zhang (2008) with regard to orthoepic competence, Ito (2019a) advances the definition by describing it as "the skill of reading text or characters…aloud." Furthermore, if we read Ito (2017), we find that the author's understanding of "reading" as orthoepic competence is "the ability to pronounce vocabulary or characters within a text, and to understand the meaning of words or characters within context, and with respect to the function of how they are written" (p. 56). Ito (2019a) recognizes this sense, writing that "if we take into account the existence of kanji within Japanese, then it makes sense to consider not only pronunciation but also the meaning of words" (p. 75). However, Ito (2019a) considers that "being able to read" implies the existence of an ambiguous continuity in the space between pronunciation and the understanding of a word's meaning. On this basis, Ito (2019a) proposes a more detailed definition, asserting that "being able to read" is a "state" or "condition" whereby "an individual can pronounce characters or vocabulary, and is at least partially capable of understanding the meaning of characters or words within context, as well as the function of how they are written" (p. 76). Ito (2019a) states that this definition is not in conflict with the former one (2017). Therefore, in this paper, I use Ito's (2019a) definition of orthoepic competence as "being capable of reading."

3   Previous Studies and Research Goals

A large number of studies on the CEFR have been written for a range of languages. Of these, many have focused on the ability to accomplish tasks, which is a typical CEFR skill (Ito, 2019a). However, previous investigations that have examined the relationship between the characters of non-alphabetic languages and CEFR orthoepic competence

are, to the best of my knowledge, limited to Bellassen and Zhang (2008), Ito (2017), and Ito (2019a).

Bellassen and Zhang (2008) investigated the application of the CEFR within Chinese language education in France. They point out that, as with Japanese, there is hardly any relationship between the pronunciation of Chinese characters and the way in which they are written. For this reason, it is nearly impossible to read an unknown Chinese character, furthermore, a learner may even be led astray by their mistaken interpretation for a similarly looking characters. Bellassen and Zhang conclude that characters present a problem in terms of applying the CEFR within Chinese language education. They propose a method for measuring the ability to write Chinese characters, which they refer to as a "literacy threshold" (识字门槛). A literacy threshold selects Chinese characters based on the frequency with which they are written, their prevalence in everyday conversation, the incidence of their use in conversation, and the number of words that can be created when they are joined with other Chinese characters. The selected Chinese characters are then divided into different language proficiency groupings with the ability to recognize and write them, forming the basis for evaluating which level of language proficiency a person has reached (Table 1). The work of Bellassen and Zhang (2008) holds some significance with respect to the problem of the relationship between Chinese character and the CEFR. However, apart from noting that frequency of appearance informs the reason for their selection, authors do not offer any details on their selection of Chinese characters, which raises doubt on whether the set literacy threshold is appropriate. Ito (2019a) points out that the basis for proficiency evaluation is not a concrete descriptor but rather an individual kanji, and thus the design of the literacy threshold does not align with the thinking of the CEFR, according to which language users are understood as "social agents" that strengthen and improve their abilities by carrying out tasks within certain set environments.

Table 1: Literacy Threshold Levels and the Number of Chinese characters
according to Bellassen & Zhang, 2008, p. 69 (author's translation)

Level	Number of Chinese characters (approximately)
C2	Over 3,000
C1	2,200
B2	1,500
B1	800
A2	500
A1	250

Furthermore, Ito (2017) suggests that the CEFR should not provide a concrete descriptor for orthoepic competence because the CEFR is not something complete and static, but is rather distinguished by an orientation toward the possibility of continual expansion and refinement. Ito (2017) indicates the Council of Europe's own language in this regard by pointing to the statement that: "the framework should be open and flexible, so that it can be applied, with such adaptations as prove necessary, to particular situations" (Council of Europe, 2001, p. 7). Ito (2017) further asserts that the creation of a descriptor for orthoepic competence is an urgent task, such that it is necessary to provide a concrete descriptor now, even if an outcome is in a form that leaves room for future debate. Using this understanding, the author advances his own proposal for an orthoepic competence descriptor for Japanese.

The work of Ito (2017) is significant for having provided a concrete orthoepic competence descriptor. However, it would be difficult to claim that the proposed descriptor is highly objective. The method that the author used to draft this descriptor is problematic because the method relies upon a subjective selection of characteristic words of the abovementioned six abilities (ranked by language proficiency), which have the following linguistic competence descriptors (excluding orthoepic and semantic competence): general linguistic range, vocabulary range, vocabulary control, grammatical competence, phonological competence, and orthographic competence.

Ito (2019a) states that "it is necessary to research [orthoepic competence] further in order to ensure that it becomes a more objective measure" (p. 78). Ito used a highly objective approach to choose words for each level of language proficiency, which he (2017) originally selected arbitrarily by employing the text mining technique, along reconsidering the orthoepic competence descriptors for levels A1 and A2. Table 2 shows the results of this updated list of descriptors. Of the orthoepic descriptors originally given by Ito (2017), only the A1 and A2 descriptors have been replaced with his revisions (2019a).

Therefore, the purpose of this paper is to extend the reconsideration conducted at Ito (2019a) from the B1 level to the C2 level and to present new orthoepic competence descriptors in Japanese language education extracted in a highly objective.

Table 2: Orthoepic Competence Descriptors that Combine Ito (2017) and Ito (2019a)

Level	Orthoepic Competence Descriptors
C2	The learner can consistently read accurately and is even deeply familiar with difficult kanji outside of his/her field of expertise.
C1	The learner makes few mistakes and can read almost all characters without referring to a dictionary, including vocabulary related to his/her field of expertise.

Level	Orthoepic Competence Descriptors
B2	The learner misreads hardly any characters due to transference (interference) from his/her native language. While there are characters he/she cannot read due to a lack of orthoepic knowledge, he/she is familiar with enough characters to ensure that he/she does not struggle with communication in daily life.
B1	The learner makes mistakes in terms of mispronunciation and transference (interference) from his/her native language, but he/she is able to read the material with which he/she comes into contact in daily life.
A2	There are many cases where the learner may need to re-read a section of text or reads incorrectly; however, if he/she has the necessary basic vocabulary, then he/she is able to read material encountered in daily situations.
A1	The learner is able to read a section of text if he/she has studied the material and has a basic, concrete, and limited repertoire (words and expressions, etc.) that relates to his/her personal information.

4   Method and Results

4.1   Analytical Method

Similar to Ito (2019a) I used the linguistic competence descriptors indicated in the Japanese language version of the CEFR, including general linguistic range, vocabulary range, vocabulary control, grammatical competence, phonological competence, and orthographic competence. (There are 9 such descriptors for levels B1 and B2, 7 for C1, and 5 for C2.) I carried out the analysis using the text mining free software KH Coder 3.[2] On the first step, I identified words that most often emerged at various levels. For this purpose, I separated descriptors for different levels (from B1 to C2) and looked at the frequency rate of words that appeared at each level (Section 4.2). Next, I implemented the co-occurrence network analysis to determine co-occurrence relationships between words for each respective level (Section 4.3). Finally, I made use of the same Japanese language version of the CEFR as Ito (2019a), so that I could combine my findings with his, thus creating orthoepic competence descriptors for levels A1 to C2.

---

[2] KH Coder 3 is a text mining software created by Higuchi Koichi. It can carry out various forms of statistical analysis with respect to text-form data. As of July 2019, it has been used in over 2,800 research programs. For details on KH Coder 3, refer to Higuchi (2014).

4.2   Analysis of Word Frequency

In order to determine the frequency rate of the words used in the descriptors for levels B1 to C2, as mentioned previously, I conducted frequency analysis using KH Coder 3 for each proficiency level. Table 3 displays the highest occurring words from ranks 1 to 10 for each respective level. For B1, the first five words (in their order of highest frequency) are: "to state" (述べる), "vocabulary" (語彙), "to hold" (持つ), "state, condition" (状況), and "accurate" (正確). For B2, they are: "vocabulary", "high" (高い), "general" (一般), "to make a mistake" (間違う), and "to use" (駆使). For C1, we have: "expression" (表現), "vocabulary", "to say" (言う), "language" (言語), and "mistake" (誤り). Finally, for C2, the words are: "language", "vocabulary", "extremely" (非常), "expression", and "broad" (幅広い). However, since none of these words has a significantly high rate of occurrence, it is difficult to determine the characteristics for each level based on word frequency alone. Therefore, as we can see in the subsequent section, I used co-occurrence network analysis, which focuses on the words displayed in Table 3.

Table 3: Top Ten High Frequency Rankings for Words in Levels B1 to C2

	Level B1	Fr.	Level B2	Fr.	Level C1	Fr.	Level C2	Fr.
1	述べる (to state)	6	語彙 (vocabulary)	3	表現 (expression)	4	言語 (language)	2
2	語彙 (vocabulary)	4	高い (high)	3	語彙 (vocabulary)	2	語彙 (vocabulary)	2
3	持つ (to hold)	4	一般 (general)	2	言う (to say)	2	非常 (extremely)	2
4	状況 (state, condition)	4	間違う (to make a mistake)	2	言語 (language)	2	表現 (expression)	2
5	正確 (accurate)	4	駆使 (to use)	2	誤り (mistake)	2	幅広い (broad)	2
6	内容 (contents)	4	見る (to see)	2	些細 (unimportant)	2	あいまい (vague)	1
7	言語 (language)	2	言う (to say)	2	使用 (to use)	2	コノテーション (connotation)	1
8	考え (thought)	2	構造 (structure)	2	正確 (accurate)	2	モニター (monitor)	1
9	使う (use)	2	持つ (to hold)	2	イントネーション (intonation)	1	レパートリー (repertoire)	1
10	問題 (problem)	2	自分 (self)	2	ニュアンス (nuance)	1	意識 (consciousness)	1

4.3   Co-occurrence Network Analysis

Co-occurrence network analysis explores the relationship or strength of co-occurrence between words, which is then expressed as a network diagram. In a network diagram, each word is depicted as a circle, with the size of the circle representing the number of times the word appears. Furthermore, the existence (or lack) of lines connecting circles (words) together, as well as the thickness of those lines, embodies relationships between words and the strength of their co-occurrence. Co-occurrence network analysis is the strongest technique within the field of text mining (Ushizawa, 2018).

This section describes extracting the respective characteristic words of all levels from B1 to C2 using the co-occurrence network analysis. However, Ito (2019a) points out that:

> In order to accurately ascertain the characteristics of each level, it is desirable to focus on relationships and co-occurrence between high frequency words only. However, if we are only looking at a small number of words when conducting this analysis of high frequency words, then we may end up with an analysis of a region defined by extremely limited relationships and co-occurrence. (pp. 71–72)

According to the definition of Ito, the network diagrams consist of those constellations in which a minimum of three words had thick lines, expressing a strong relationship or co-occurrence. I followed this model. For example, in the case of B1, I only carried out co-occurrence network analysis for words with a frequency rate of six.

In cases where the network diagram did not fit this condition, I expanded the range to include words with a frequency rate that was 1 point lower. According to this approach, I determined the network diagrams used as the definition of analysis. With respect to the results of the co-occurrence network analysis for levels B1 to C2, each network diagram from B1 to C1 ended up containing words with a frequency rate as low as 2. On the other hand, level C2 only has a small number of descriptors (five), which was too few to create a co-occurrence network diagram. For this reason, I limited co-occurrence network analysis to levels B1 to C1. Regarding level C2, Ito (2019b) examined visual reception in the CEFR and commented on C2 level competence. I therefore referred to Ito (2019b) for this area while attempting to make my own slight revisions. In the next section, I present my findings from having utilized the word groupings attained through co-occurrence network diagrams (in terms of characteristic words for each level), and of having re-considered orthoepic competence descriptors for each level.

5   Considerations

In this section, I outline the findings of my re-consideration of orthoepic competence descriptors for levels B1 to C2, based on the outcomes of the co-occurrence network analysis and the results of Ito (2019b).

5.1   Orthoepic Competence Descriptors for Level B1

With respect to level B1, Ito (2017) proposes the following orthoepic competence descriptor: "The learner makes mistakes in terms of mispronunciation and transference (interference) from his/her native language, but he/she is able to read the material with which he/she comes into contact in daily life." Indeed, just as Ito (2017) states that the learner "is able to read the material with which he/she comes into contact in daily life," we can infer from the group in Figure 2 – comprised of "vocabulary" (語彙), "daily" (日常) and "topic" (話題) – that individuals at level B1 can "read *vocabulary* [associated with] *daily topics*".
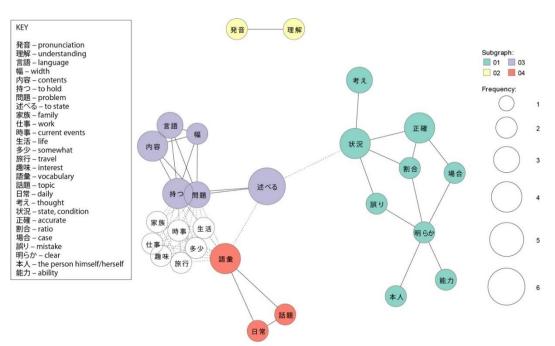

Figure 2: B1 Co-occurrence Network

In the case of the group centered on "to state" (述べる), "to state" does have a relationship with oral competence. However, "to hold" (持つ) has a relationship of co-occurrence with "to state," and "to hold" has a relationship of co-occurrence with "contents" (内容), "width" (幅), and "language" (言語). We can hence surmise that at this level, the individual can "state" or "speak" (述べる) about a "broad" (幅広い) range of "topics" (内容). Furthermore, if we look to the group centered on "state,

condition" (状況) and "accurate" (正確), we find a relationship between "state, condition," "ratio" (割合), and "accurate." "State, condition," "ratio," and "mistake" (誤り) also have a co-occurrence relationship, while "ratio" and "mistake" have a co-occurrence relationship with "clear" (明らかな). Based on this analysis, I propose changing the B1 orthoepic competence descriptor to the following: "While there are cases where [learners] may make obvious (明らかな) mistakes (間違い), they are able to relatively accurately (割合正確に) read characters related to a broad range of material (幅広い内容), including everyday topics (日常的な話題)" (Table 4). The resulting change of the section "there are cases where [learners] may make obvious mistakes" (明らかな間違いを犯す場合もあるが) represents a revision toward a slightly higher degree of abstraction when compared to that of Ito (2017): "The learner makes mistakes in terms of mispronunciation and transference (interference) from his/her native language" (発音の間違えや母語の転移（干渉）などの読み間違えもあるが). If we consider this issue from the perspective of the CEFR's concept of "plurilingualism" – according to which all linguistic knowledge and experience contribute to the development of new communication – then Ito's (2017) comment on errors arising out of "transference (interference) from…native language" is actually due to a mistaken interpretation. Furthermore, as noted in Section 2.2, I use Ito's (2019a) definition of "being able to read." If we assume that mistakes are not restricted to "misreading," then we come closer to an orthoepic competence descriptor that better conforms to the reality of – or rather, a correct understanding of – the CEFR.

Table 4: Revisions to the Descriptor for B1 Orthoepic Competence

Level	Ito (2017)	Revisions based on the current analysis
B1	The learner makes mistakes in terms of mispronunciation and transference (interference) from his/her native language, but he/she is able to read the material with which he/she comes into contact in daily life.	While there are cases where [learners] may make obvious mistakes, they are able to relatively accurately read characters related to a broad range of material, including everyday topics.

5.2   Orthoepic Competence Descriptors for Level B2

Next, if we turn to level B2, Ito (2017) provides the following descriptor:

> The learner misreads hardly any characters due to transference (interference) from his/her native language. While there are characters he/she cannot read due to a lack of orthoepic knowledge, he/she is familiar with enough characters to ensure that he/she does not struggle with communication in daily life.

However, unlike Ito (2017) or the examination of level B1, Figure 3 does not portray a strong co-occurrence relationship between "mistake" (間違う) and other terms.
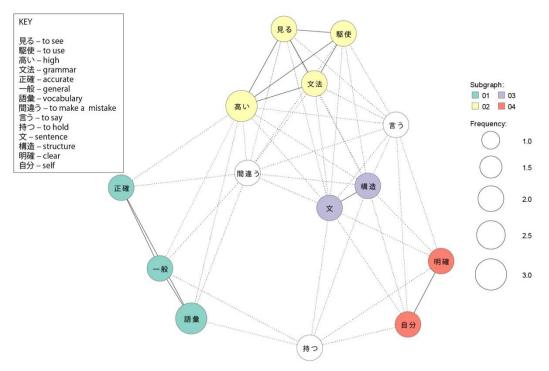
Figure 3: B2 Co-occurrence Network

Furthermore, with respect to "vocabulary" (語彙), in case of level B1, there is a co-occurrence relationship with "everyday topics" (日常的な話題), yet in case of level B2, there are instead co-occurrence relationships with "general" (一般) and "accurate" (正確). Thus, the breadth of vocabulary in level B2 seems slightly more extensive, making it a level where a language learner is "able to accurately use [用いる] words if they are common" (一般的な語について正確に用いることができる);[3] this segment represents another group in level B2 that displays a co-occurrence relationship. In this group, a strong, reciprocal co-occurrence relationship is found between "high" (高い), "to see" (見る), "use" (駆使), and "grammar" (文法). From this observation, we can infer that B2 users possess an extensive command of grammar. I propose that we combine this conclusion with the earlier statement that B2 users can "accurately use words if they are common." Thus, we attain a descriptor of orthoepic competence for this level, whereby the individual "has a high level of orthoepic competence, and is able to accurately read [読む] words if they are common" (高い読字能力を持っており、一般的な語であれば正確に読むことができる).

---

[3] Regarding this observation about the "expansion of vocabulary" at the B2 level, Ito (2019b), in his investigation of the abilities considered important for visual receptive activities, similarly notes that from B2 onward, "not only the type of reading or the method of reading, but also the depth of understanding of the content receives more emphasis at this stage, with a higher degree of competency expected" (p. 73). There may be a connection with my results.

Table 5: Revisions to the Descriptor for B2 Orthoepic Competence

Level	Ito (2017)	Revisions based on the current analysis
B2	The learner misreads hardly any characters due to transference (interference) from his/her native language. While there are characters he/she cannot read due to a lack of orthoepic knowledge, he/she is familiar with enough characters to ensure that he/she does not struggle with communication in daily life.	[The individual] has a high level of orthoepic competence, and is able to accurately read words if they are common.

5.3 Orthoepic Competence Descriptors for Level C1

Ito (2017) offers the following descriptor for level C1: "The learner makes few mistakes and can read almost all characters without referring to a dictionary, including vocabulary related to his/her field of expertise." Figure 4 displays a group centered on "expression" (表現) and one centered on "mistake" (誤り), "unimportant" (些細), and "accurate" (正確). Of these, the "mistake," "unimportant," and "accurate" group shares a commonality with the section of Ito's (2017) descriptor, which states that the individual "makes few mistakes" (literally "mistaken readings," 読み間違え).
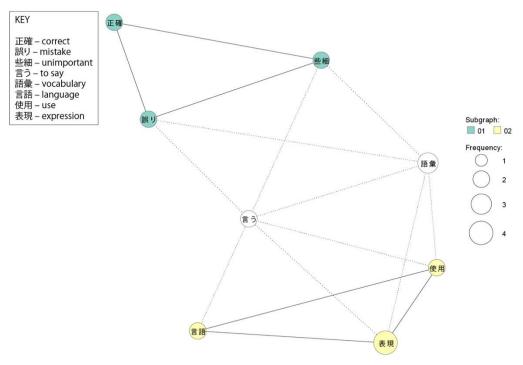

Figure 4: C1 Co-occurrence Network

Furthermore, "mistake," "unimportant," and "vocabulary" (語彙) are connected via a weak co-occurrence relationship. This connection points at users making "slight" (i.e., unimportant) "mistakes" at the level of "vocabulary." On the other hand, as with our observation regarding level B1, "mistakes" are not limited to "mistaken readings." For this reason, I suggest that Ito (2017)'s section of "makes few mistakes" be revised to "makes slight mistakes at the level of vocabulary, but can read accurately." Regarding the group centered on "expression," "expression," "language" (言語), and "use" (使用) are connected through a strong co-occurrence relationship. Meanwhile, "to speak" (言う) is connected to the three abovementioned words, and "vocabulary" has a weak co-occurrence relationship with "expression" and "use." We do not see keywords such as "field of expertise" or "dictionary" given by Ito (2017). However, in terms of C1 visual reception, Ito (2019b) notes that the framework emphasizes a language learner's ability to have "a detailed understanding of opinions (etc.) regarding complex, difficult texts related to [his/her] own [field of] specialty" (自分の専門) (p. 74). We can surmise that "field of expertise" (専門分野) is an important part of C1. Hence, I altered this section to read: "can accurately read linguistic expressions and vocabulary such as [those] used in [his/her] field of expertise." Combining this section with the above mentioned first part, my suggested reform of the orthoepic competence descriptor for C1 now reads as follows: "While [learners] may make slight mistakes at the level of vocabulary, they can accurately read linguistic expressions and vocabulary, such as [those] as used in their field of expertise" (語彙レベルの些細な誤りもあるが、専門分野などで使用される言語表現や語彙を正確に読むことできる).

Table 6: Revisions to the Descriptor for C1 Orthoepic Competence

Level	Ito (2017)	Revisions based on the current analysis
C1	The learner makes few mistakes and can read almost all characters without referring to a dictionary, including vocabulary related to his/her field of expertise.	While [learners] may make slight mistakes at the level of vocabulary, they can accurately read linguistic expressions and vocabulary, such as [those] as used in their field of expertise.

5.4   Orthoepic Competence Descriptors for Level C2

For the C2 level Ito (2017) offers the following descriptor: "A learner can consistently read accurately and is even deeply familiar with difficult kanji outside of his/her field of expertise." Since I was unable to create a co-occurrence network, I attempted to use an alternative approach to revision. My strategy involves revealing information about C2 orthoepic competence by looking more closely at the capacity valued with respect to C2 visual reception, as outlined by Ito (2019b). According to Ito (2019b), the skill valued in terms of C2 visual reception is the language learner's ability to "understand

and savor complex meanings in sources such as literature" (pp. 74–75). First, since "understanding complex meanings in sources such as literature" is expected, we can revise the section "outside of his/her field of expertise" by adding "such as literature" (文学など), which would be more suitable. Being able to read "difficult kanji" is not necessarily required for "understanding and savoring complex meanings in sources such as literature." However, since "difficult kanji" (literally "kanji difficult to read" or 難読漢字) is indicated, and because I was unable to conduct a detailed analysis into the C2 level, I deemed it inappropriate to make large changes. Therefore, regarding a modification to C2 orthoepic competence, I propose the following: "The learner can consistently read accurately, and can besides read difficult kanji such as those found in literature" (一貫して正しい読みをし、文学などで使われる難読漢字なども読むことができる).

Table 7: Revisions to the Descriptor for C2 Orthoepic Competence

Level	Ito (2017)	Revisions based on the current analysis
C2	The learner can consistently read accurately and is even deeply familiar with difficult kanji outside of his/her field of expertise.	The learner can consistently read accurately, and can besides read difficult kanji such as those found in literature.

5.5   Revised Orthoepic Competence Descriptors

This concludes the overview of my revisions using co-occurrence network analysis regarding the orthoepic competence descriptors for levels B1 to C2, as provided by Ito (2017). Table 8 lists the revised orthoepic competence descriptors for levels B1 and C2 as outlined in this paper, alongside the re-considered orthoepic competence descriptors for levels A1 and A2 advanced by Ito (2019a).

Table 8: Revised Orthoepic Competence Based on Current Research
(B1-C2) and Ito (2019a).

Level	Revised Orthoepic Competence
C2	The learner can consistently read accurately, and can besides read difficult kanji such as those found in literature.
C1	While [learners] may make slight mistakes at the level of vocabulary, they can accurately read linguistic expressions and vocabulary, such as [those] as used in their field of expertise.
B2	[The individual] has a high level of orthoepic competence, and is able to accurately read words when common.
B1	While there are cases where [learners] may make obvious mistakes, they are able to relatively accurately read characters related to a broad range of material, including everyday topics.

Level	Revised Orthoepic Competence
A2	There are many cases where the learner may need to re-read a section of text or reads incorrectly; however, if he/she has the necessary basic vocabulary, then he/she is able to read material encountered in daily situations.
A1	The learner is able to read a section of text if he/she has studied the material and has a basic, concrete, and limited repertoire (words and expressions, etc.) that relates to his/her personal information.

6   Conclusion and Future Tasks

Building upon the earlier work of Ito (2017, 2019a) – that is, the creation of suggested orthoepic competence descriptors for the CEFR, as well as a re-investigation of levels A1 and A2 – I re-examined the descriptors for levels B1 to C2. My findings did not return characteristic words determined on the basis of frequency of appearance. However, by utilizing co-occurrence network analysis, I was able to extract co-occurrence networks for levels B1 to C1. By re-scrutinizing Ito's (2017) proposal for orthoepic competence descriptors, I was able to identify a more objective method than that utilized by Ito (2017). Employing this technique, I outlined revised orthoepic competence descriptors for levels B1 to C1. Regarding C2, a small number of descriptors for this level meant that co-occurrence network analysis was unfeasible. For this reason, I only made out light changes grounded in inferences drawn from the competences valued for visual reception, as shown by Ito (2019b). By combining my outcomes with those of the re-examination conducted by Ito (2019a) into levels A1 and A2, I was able to present orthoepic competence descriptors that are more objective than those initially put forth by Ito (2017).

Nevertheless, a number of issues remain that should be addressed in future research. First, as noted above, since co-occurrence network analysis was not possible for the C2 descriptors, it is similar to a tentative plan. Furthermore, the B1 level descriptor was altered from "The learner makes mistakes in terms of mispronunciation and transference (interference) from his/her native language" to "there are cases where [learners] may make obvious mistakes." This modification represents a higher degree of abstraction, and may be considered an orthoepic competence descriptor that more closely aligns with the reality of – or, rather, a correct understanding of – the CEFR. Nevertheless, at the same time, we cannot deny the possibility that for B2, the resulting descriptor – "[The individual] has a high level of orthoepic competence, and is able to accurately read words if they are common" – may in fact be too abstract. Hence, some further consideration as to whether this descriptor can be used in its current formn is needed. Going forward, relying on the analysis of descriptors alone may not be appropriate for further investigating problems such as the above (that is,

the inability to conduct analysis for C2 due to a small number of descriptors, and the feeling that the descriptor for B2 is too abstract). Concerning the orthoepic competence descriptors presented in this paper, it will be necessary to focus on administering a Japanese language test and self-evaluation questionnaire to Japanese language learners in order to thoroughly comprehend the relationship between orthoepic competence descriptors and Japanese language ability. At present I do not have a concrete research plan; I can only offer these orthoepic competence descriptors. Future studies should explore the relationship between Japanese language ability and the self-evaluation of orthoepic competence.

References

Bellassen, J., & L. Zhang. (2008). Ou zhou yu yan gong tong can kao kuang jia xin li nian duihan yu jiao xue de qi shi yu tui dong– chu yu jue ze guan tou de han yu jiao xue [The CEFR: New concept and its implications and impetus – Chinese language teaching at its critical moment]. *Shi jie han yu jiao xue [Chinese Teaching in the World]*, *85*, 58–73.

Chéng, Y. (2017). *Chūka sekai ni okeru CEFR no juyō to bunmyakuka [The introduction and contextualization of CEFR in the Chinese speaking world]*. Tokyo, Japan: Coco Publishing.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge, UK: Cambridge University Press.

Higuchi, K. (2014). *Shakaichōsa no tameno keiryō tekisuto bunseki—naiyō bunseki no keishō to hatten o mezashite [Quantitative textual analysis for social surveys: Aiming for the succession and development of content analysis]*. Sapporo, Japan: Nakanishiya Publishing.

Ito, H. (2017). Kakuchō seichika no tameno dokuji nōryoku kijutsubun shian sakusei -CEFR/JFS no gengo kōzōteki nōryoku o sankō ni- [Devising tentative descriptors of orthoepic competence for extension and refinement: With reference to CEFR/JFS linguistic competence]. *Nihongo kyōiku [Journal of Japanese Language Teaching]*, *168*, 55–62.

Ito, H. (2019a). Dokuji nōryoku no hyōka shakudo no saikō- "kiso dankai no gengo shiyōsha" ni chūmoku shite- [Reconsidering can-do descriptors for orthoepic competence with a focus on the CEFR's "Basic Language User]. *Kiso kyōiku hoshōgaku kenkyū [The Journal of the JASBEL], 3*, 72–86.

Ito, H. (2019b). CEFR no juyōteki katsudō dewa naniga jûshi sareteirunoka- shikakutekina juyōteki katsudō no kaku reberu no bunseki- [What is important in receptive activities of CEFR? An analysis of visual receptive activities for different levels]. *Nihongo kyōiku [Journal of Japanese Language Teaching], 173*, 69–76.

Kumano, N., Ito, H., & Hachisuka, M. (2013). JFS/CEFR ni motozuku JFS nihongo kōza reberu nintei shiken (A1) no kaihatsu [Level certification test (A1) development for JFS Japanese-language courses based on JFS/CEFR]. *Kokusai kōryū kikin nihongo kyōiku kiyō [The Japan Foundation Japanese-Language Education Bulletin]*, *9*, 73–88.

Länsisalmi, R. (2012). Conquering the world with "cutting-edge curricula": Global Citizens learning East Asian Languages. In D. Smakman & L. Willemsen (Ed.), *Proceedings of the 2012 "Van Schools tot Scriptie" Colloquium* (pp. 99-111). Leiden, Netherlands: University Library, Leiden University.

Majima, J. (2018). CEFR no kokunaigai no nihongo kyōiku eno inpakuto [The domestic and international impact of CEFR on Japanese language education]. In H. Sensui (Ed.), *Kotoba o oshieru kotoba o manabu fukugengo fukubunka yōroppa gengo kyōtsū sanshōwaku [Teaching language, learning language: Language education and the Common Framework of Reference for Languages for a multi-lingual, multi-cultural Europe (CEFR)]* (pp. 249–274). Shiga, Japan: Kōrosha.

Meyer, F. (2010, August). *Some considerations on the new HSK. Does it help spreading Chinese as a foreign language (CFL)?* Paper presented at the 16th Meeting on Chinese as a Foreign Language. Zürich, Switzerland: University of Zürich.

North, B. (2014). *The CEFR in Practice.* Cambridge, UK: Cambridge University Press.

Shigemori Bučar, C., Ryu, H., Moritoki Škof, N., Hmeljak Sangawa, K. (2014). The CEFR and teaching Japanese as a foreign language. *Linguistica, 54*(1), 455-469. https://doi.org/10.4312/linguistica.54.1.455-469

Ushizawa, K. (2018). *Yatte miyō tekisuto mainingu jiyū kaitō ankēto no bunseki ni chōsen! [Let's try text mining: Learn free answer questionnaire analysis!]* Fukuoka, Japan: Asakura Shoten.