

論文の英文要旨	
論文題目	Determining Factors of Good Example Selection for EFL Teaching and Learning: Toward Methodologies and Evaluation Criteria
氏名	Kawamoto, Naho

This dissertation aims to empirically investigate the criteria for selecting and evaluating “good” example sentences for English teaching and learning. Although example sentences are commonly provided in various language learning materials such as textbooks, grammar books, and vocabulary exercise books, no set of criteria has been established for evaluating these example sentences. Previous research in fields such as pedagogical lexicography, Natural Language Processing (NLP), Computer-Assisted Language Learning (CALL), and Data-Driven Learning (DDL) has focused mainly on linguistic features contributing to determining difficulty levels of example sentences; however, factors determining the quality of example sentences extend beyond these linguistic features. Yet, empirical research focusing comprehensively on various factors determining the quality of example sentences is scarce.

After a brief introduction in Chapter 1, Chapter 2 first reviews previous studies in the field of pedagogical lexicography, NLP, CALL, and DDL to identify key variables that contribute to determining the quality of example sentences and the effectiveness of their usage. Based on this information, Chapter 3 proposes a tentative framework titled “the Model of Example Use and Evaluation.” Within this model, the terms of important factors for evaluating example selection and usage are introduced: the type of linguistic information each example is expected to deliver ([INFORMATION]), the users who utilize those examples ([USER]), the various language teaching materials through which the examples are provided ([MEDIUM]), and the intended purposes behind providing examples ([PURPOSE]). In addition to these factors, the model sets overarching criteria for evaluating example sentences ([CRITERIA]), which are determined through the mutual interaction of [USER], [MEDIUM], and [PURPOSE]. These CRITERIA determine the usefulness of the examples. Among the CRITERIA, there is a set of linguistic features [Criterial Linguistic Properties] that contribute to the determination of the quality of examples. This is distinct from the type of information that examples are intended to convey ([INFORMATION]), which refers to various types of information presented by the examples for their intended purposes. Criterial Linguistic Properties, on the other hand, define linguistic features contributing to sentence quality such as the sentence length or the average vocabulary level in a sentence. Using this framework, three research questions are established:

- RQ1 What is the nature of “criterial” linguistic properties of examples for evaluating the quality and difficulty of sentences?
- RQ2 How should sentence characteristics be properly weighted in light of other factors such as PURPOSE, USER, and MEDIUM?
- RQ3 How do actual users use examples under specific conditions?

To address these research questions, four empirical studies were conducted, and this dissertation attempts to refine the “Model of Example Sentence Usage and Evaluation” based on their findings. The empirical studies investigated the relationships between various factors examined in the previous chapters.

Three empirical studies (Chapters 4 through 6) were conducted to investigate the

interrelationships among the aforementioned factors. Chapter 4 examined the relationship between Criterial Linguistic Properties and the purpose of providing example sentences ([Criterial Linguistic Properties] × [PURPOSE]). Chapter 5 explored the relationship between Criterial Linguistic Properties and the annotated difficulty level of example sentences in CEFR-based inventories ([Criterial Linguistic Properties] × [USER]). Chapter 6 focused on how example sentences set on Criterial Linguistic Properties are perceived by actual learners ([Criterial Linguistic Properties] × [USER] ↔ [REAL USER]). Building upon these findings, Chapter 7 shifted attention to examining how all of these linguistic features contribute to the quality of example sentences revealed in the previous chapters and conducted an integrated observational study of example sentence usage by actual users.

Chapter 4 specifically focuses on the interplay between criterial linguistic properties of examples and the intended purposes of providing examples—in particular, the illustrative examples provided within the existing monolingual and bilingual dictionaries. Linguistic features expected as contributing factors (such as sentence length, use of phrasal vs. sentence examples, vocabulary level, etc.) were annotated on each sentence, and criterial linguistic properties distinguishing between the purposes of providing examples (e.g., for encoding vs. decoding) were explored using a statistical method, decision trees. The results revealed that, especially in bilingual dictionaries, formal properties of examples (phrases vs. sentences) contributed most to distinguishing between examples intended for decoding and encoding. Conversely, in monolingual dictionaries that tended to provide full-sentence examples, the length of examples (i.e., the number of words in the sentence) contributed to distinguishing between examples for reception and production. Additionally, lexical features such as the average vocabulary level of examples and the percentage of B1- and B2-level words within sentences partially contributed to this distinction.

Chapter 5 investigated the criterial linguistic properties contributing to the difficulty level of examples. Using examples provided in CEFR-based inventories (*Core Inventory*, *Global Scale of English*, *English Grammar Profile*), the study examined which linguistic features determined the difficulty level of examples through ordinal logistic regression analysis. The results of Chapter 5 showed that criterial linguistic properties that consistently contributed to the determination of example sentences in every inventory were quite limited. The statistical analysis identified variables such as the number of A1-level words within a sentence, the maximum depth of syntactic structures, the number of nouns per sentences, the average CEFR-J grammatical levels of examples, and the number of A2.2-level grammatical features within sentences as variables contributing to example difficulty. With few exceptions, these criterial linguistic properties showed positive coefficients in ordinal logistic regression models, indicating their effectiveness as features for level advancement.

Chapter 6 experimentally investigated how examples contained in the CEFR-based inventories were perceived by actual users. While expert-created CEFR inventories provide valuable insights into features contributing to the difficulty of examples, learners may not necessarily perceive the difficulty level of these examples as specified by the inventories. Therefore, using examples in the *Core Inventory* and *Global Scale of English*, the study examined learners' perceived difficulty levels. The study operationalized learners' perceived example difficulty as the accuracy of translation into Japanese and investigated whether there was a correlation between learners' perceived difficulty levels and the difficulty levels assigned by the inventories using Rasch analysis. The features correlated with the rank order determined by Rasch analysis were the presence of A2-level, B2-level, and proper nouns as well as syntactic related features, such as the maximum depth of syntactic features and the length of examples. Furthermore, qualitative analysis revealed that linguistic properties other than the annotated linguistic features contributed to perceived difficulties. They included verb tenses and the presence of polysemous words

Chapter 7 observed how learners as example users interacted with example sentences in actual language use situations. By synthesizing results from the previous chapters, an observational study was conducted to explore the criterial linguistic properties and interface features in the context of

dictionary use by Japanese learners of English. This involved considering six variables as criterial linguistic properties (phrases vs. full sentences, length of examples, position of headwords within examples, presence of difficult words within examples, grammar structure of examples) and five variables related to the dictionary interface (number of examples, presentation order, presence of grammar codes, presence of Japanese translations, presentation method [KWIC vs. sentence presentation]). Tasks included English sentence completion, error correction, and Japanese-English translation based on provided dictionary information. Cognitive processes of learners were elicited using think-aloud protocols. The provided dictionary information was organized into controlled sets to observe learners' performance across different sets. The observation results revealed complex mutual influences between variables related to criterial linguistic properties and interface on example usage difficulty.

Chapter 8 discusses the outcomes of the individual studies conducted in Chapters 4 to 7, with the objective of refining the “Model of Example Sentence Usage and Evaluation” introduced in Chapter 3. This dissertation makes methodological contributions by striving to systematize the purposes, conditions, and evaluation criteria of dictionary example sentences through a mixed-method approach. Theoretical contributions entail proposing a framework for example sentence research applicable to interdisciplinary fields like lexicography, language learning, and natural language processing. Pedagogically, the study clarifies the utilization of example sentences for specific purposes and target users, and suggests more effective methods for creating and selecting example sentences based on this clarification.

*Keywords: example sentences, good examples, CEFR, criterial linguistic properties, user perspectives, pedagogical lexicography, CEFR inventory*