Chapter 10: Vocabulary and speaking – Current research, tools, and practices

Takumi Uchihara
Orcid.org/0000-0003-4476-534X

Abstract
Vocabulary knowledge is considered an essential component of second language (L2) proficiency, and there has been a growing interest in the relationship between vocabulary and speaking. This line of research generally takes two approaches: the first compares lexical use with general oral proficiency, both of which derive from the same spoken responses (i.e., vocabulary and speaking data are dependent); the second compares lexical knowledge with oral proficiency, measured separately (i.e., vocabulary and speaking data are independent). This chapter focuses on the second approach and gives an overview of research examining the vocabulary-speaking link, highlighting four aspects of L2 speech: temporal, lexical, phonological, and global features. I also discuss how L2 speech was measured in previous studies to explore the relationship between vocabulary and speaking. Finally, I suggest the importance of exploring the role of vocabulary in various aspects of L2 speech in a way that sheds further light on the vocabulary-speaking relationship. This chapter ends with further reading and useful information for readers who are interested in the measurement of different L2 speech aspects.
Introduction

The relationship between vocabulary and speaking has received recent attention in L2 research (e.g., Koizumi & In’nami, 2013; Kyle & Crossley, 2015), the area which used to be neglected in comparison to, for example, the relationship between vocabulary and reading (Uchihara & Saito, 2019). Researchers generally take two approaches when investigating the vocabulary-speaking link. The first approach is to elicit speech samples through oral tasks (e.g., oral narrative), assess the samples holistically (e.g., native judgements of communicative adequacy), score the same samples lexically (e.g., number of infrequent words), and examine the extent to which multiple lexical measures predict general speaking proficiency using multivariate statistical analysis (e.g., Kyle & Crossley, 2015; Saito, Webb, Trofimovich, & Isaacs, 2016). The second approach, unlike the first, elicits vocabulary and speaking data separately, and therefore the two elicited samples are not dependent. Focusing on studies taking the second approach, this chapter reviews research on vocabulary and four aspects of L2 speech—fluency, lexical richness, pronunciation, and global features—and introduces tools commonly used for measuring oral proficiency in this field. Due to space limitation, this chapter will focus exclusively on reviewing and discussing speaking measurement (for those who are interested in tools for measuring vocabulary knowledge, refer to Read, 2000; Schmitt, 2010).

Current research on vocabulary and speaking

Oral fluency is probably the most extensively researched aspect of L2 speech in relation to vocabulary knowledge (e.g., De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2013; Hilton, 2008; Koizumi & In’nami, 2013; Uchihara & Saito, 2019; Uchihara, Saito, & Clenton, this volume). From a theoretical standpoint, the speech production model posits that speaking is lexically driven, to the extent that learners with rich lexicons are hypothesized to retrieve lemmas efficiently, making their overall speech production fast (Kormos, 2006). Research supports this view, as studies consistently report medium-to-large correlations between vocabulary knowledge and fluency—particularly, with speed fluency (e.g., articulation rate, mean length of run), $r = .34$ to .67 (De Jong et al., 2013; Hilton, 2008; Uchihara & Saito, 2019).

Exploration of lexical richness in spoken responses is another way of evaluating L2 speaking proficiency. Lexical richness is commonly defined as lexical sophistication and diversity (see Kyle, 2019 and Read, 2000 for discussion of the construct), and the former is often measured using word frequency information and the
latter is measured with a simple type token ratio or advanced variants of it (Kyle, 2019). In principle, learners producing more low frequency words and fewer repetitions are more proficient than those who do the opposite. Recent studies examine the relationship between vocabulary knowledge (size test scores) and use (lexical richness) in order to test the hypothesis that learners with rich lexical knowledge show lexically rich language use in speech (Uchihara & Clenton, 2018; Uchihara et al., this volume). Their findings show significant correlations between the two, but a closer examination of their data suggest a complexity of the relationship, indicating that speakers with larger vocabulary sizes might not necessarily produce lexical richer words.

Pronunciation, perhaps receiving the least attention in vocabulary research, is also one of the important aspects of speaking proficiency. Although our understanding of the vocabulary-pronunciation link is limited, research has begun to shed some light on the role that vocabulary plays in phonological development. Bundgaard-Nielsen, Best, and Tyler (2011) suggest that learners with larger vocabulary sizes have finer-tuned phonological representations, enabling more accurate phonological perception. In Uchihara and Saito (2019), however, vocabulary knowledge did not seem relevant to the ability to pronounce the L2 in a target-like manner. More research is needed to advance our understanding of how vocabulary knowledge relates to pronunciation ability at both perception and production levels.

Finally, some studies examine vocabulary knowledge and global aspects of L2 speech, in addition to specific oral aspects (e.g., fluency, lexical richness, and pronunciation). The global construct includes perceived comprehensibility (i.e., ease of understanding; Saito et al., 2016) and communicative/functional adequacy (i.e., success of task achievement; De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2012). Research indicates some indirect relationship between vocabulary knowledge and comprehensibility (Uchihara & Saito, 2019; Uchihara et al., this volume), whereas De Jong et al. (2012) suggest that learners with large vocabulary sizes are more likely to complete oral tasks successfully.

**Tools and practices for measuring speaking proficiency**

Given the many different approaches and tools to assess L2 speaking proficiency, this section focuses mainly on the tools and practices that are commonly employed in studies examining the relationship between vocabulary knowledge and speaking. Approaches to assessing L2 speech is broadly divided into the following two, human rating and objective measurement employing acoustic and corpus-based analysis tools.
Human rating

One typical approach to assessing L2 speech is to employ listener judgements. After speech samples are elicited through oral tasks (e.g., picture narration, interview), trained or untrained raters listen to each of the elicited samples and rate these based on numerical scale points while referring to holistic or analytic language descriptors. By way of an example of holistic rating, in De Jong et al. (2012), four native speaking non-expert raters were selected to evaluate speech data on the communicative/functional adequacy of the oral responses. Recruiting untrained raters was an important decision in their study because the researchers did not want their raters to pay special attention to specific linguistic errors (e.g., lexical and grammatical errors). As for analytic rating, on the other hand, raters are encouraged to attend to specific linguistic features. For instance, in Uchihara and Saito (2019), five native speaking raters were trained to refer to a fluency descriptor stating various kinds of temporal information, including the number of filled/silent pauses and repetition. In Uchihara and Clenton (2018), three native speaking expert raters were instructed to focus on lexical use of L2 speakers in reference to the vocabulary component of the IELTS speaking band descriptors containing various vocabulary-related information.

Objective measures

Another approach to L2 speech assessment is quantifying linguistic features in question by means of, for example, counting the number of lexical, grammatical, and phonological errors. This approach provides insight into learners’ L2 use objectively, rather than relying on human rating. For instance, De Jong et al. (2013) used acoustic analysis tools, PRAAT (Boersma & Weenik, 2013), and calculated number of silent pauses, total duration of speaking time, and total duration of pausing time for the purpose of assessing oral fluency. For measuring lexical sophistication, Uchihara and Clenton (2018) used a corpus-based lexical analysis tool, TAALES (Kyle & Crossley, 2015), in order to calculate average frequency scores given to individual words used in oral responses per speaker. For measuring lexical diversity, Uchihara et al. (this volume) adopted the measure of textual lexical diversity (MTLD) automatically produced by a text analysis tool, Coh-Mertix (McNamara, Graesser, McCarthy, & Cai, 2014).

Conclusion

This chapter provided a brief overview of research on vocabulary and speaking, and introduced existing tools and practices adopted for measuring speaking proficiency in
this area of research. Readers should note that the information provided in this chapter is far from exhaustive and many additional speaking measures are available to examine the vocabulary-speaking relationship, such as syntactic complexity (e.g., number of clauses per speech unit; Koizumi & In’nami, 2013), discourse competence (e.g., number of cohesive devices; Saito, Trofimovich, & Isaacs, 2017), and pronunciation accuracy (e.g., vowel and consonant errors; Suzuki & Kormos, 2019). In order to advance our understanding of the relationship between vocabulary and speaking, it is important to investigate the relative contribution of learners’ vocabulary knowledge to different aspects of L2 oral proficiency.

**Further reading and useful information**


- Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics, 38*, 439–462. *This article provides insight into multifaceted aspects of oral proficiency measured globally (e.g., comprehensibility) and analytically (e.g., temporal, lexical, and phonological features), using various speech analysis techniques, such as human rating, acoustic, and corpus-based measures.*

**References**


Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-


language speaking ability. *Language Teaching Research.*
