REVIEW OF THE LITERATURE

In order to assess speaking ability, it is first necessary to discuss the nature of speaking a second language. Lado (1961) notes that, “Language is more than the apparently simple stream of sound that flows from the tongue of the native speaker… It is a complex system of communication with various levels of complexity involving intricate selection and ordering of meanings, sounds, and larger units and arrangements” (p. 2). Untangling the intricacies of communication is imperative if any assessment of speaking ability endeavors to measure speaking ability with validity and reliability.

Of course, there is a wide variety of features a speaker must know in order to communicate verbally, and as Fulcher (2003) concedes, “no operational construct definition can ever capture the richness of what happens in a process as complex as human communication” (p. 19). However, by incorporating the following three components of communicative competence into the speaking ability construct, we argue that these features align with the goals of, and services offered by, Tools for Clear Speech. The following sections provide in-depth backgrounds of the central components of the Oral Communication Assessment for English Language Learners: grammatical knowledge, phonological knowledge, and pragmatic knowledge.
Grammatical Knowledge

Over the course of the last sixty years, a number of notable models have attempted to account for a clear explanation of the components of grammatical knowledge. These models can be broadly divided into two epistemologically distinct categories: structuralist and communicative (Purpura, 2004). Early conceptualizations of grammar were typically syntactocentric, and continue to exert a strong influence on second language (L2) assessments.

Chomsky (1965), in his proposal of transformational-generative grammar, and later Universal Grammar (UG), provides one of the best-known notions of form-based grammar. This theory of UG claims that knowledge of a language is composed of both universal principles consistent with all languages and the parameters that are unique to an individual language or variety of a language. However, due to its inability to account for the prominent role of semantics, and its overall exclusion of pragmatics, numerous researchers have criticized UG (Hymes, 1971; Halliday, 1994; Purpura, 2004).

Lado (1961), another early proponent of structuralist theory, conceives of grammatical knowledge as being narrowly defined through a “skills-and-elements” model of language proficiency. In this model, elements consist of “pronunciation, grammatical structure, the lexicon, and cultural meanings,” while skills are defined as reading, writing, listening, and speaking (Lado, 1961, p. 25). Lado further divides grammatical structure into morphology and
syntax. In line with structuralist theory, absent in this model is any mention of knowledge of grammatical meaning. Furthermore, with respect to L2 assessment, the model posits that discrete-point tests measuring one grammatical form are adequate, while knowledge of grammatical meaning is disregarded. This weakness was later challenged by Carroll (1968), who broadens grammatical knowledge by connecting grammar with use in a real-life task.

More recent models of grammatical knowledge embody a communicative approach, and have been informed by linguistic philosophers dating back to the 1960s who illustrated the idea that “grammar…stands not only for form, but also for meaningfulness and pragmatic appropriacy, acceptability or naturalness” (Purpura, 2004, p. 16). Austin (1962), for example, argues that an utterance extends far beyond the literal meaning of the utterance and deconstructed utterances into three related speech acts (see the section on pragmatic knowledge for further discussion on speech acts). In essence, Austin (1962) maintains the importance of context in understanding an utterance, thus demonstrating that underlying meanings may be intertwined with linguistic forms. Later, Hymes (1972) extended the theory of communication beyond speech act theory and UG by proposing that speakers must have both linguistic competence and communicative competence. Appropriacy, as a component of communicative competence, is defined by Hymes (1967) as relating to, among others, the participants, roles, setting, form of the message, topic, and purpose.

Oller (1979), another critic of Lado’s (1961) elements-and-skills approach to proficiency,
argues for *pragmatic expectancy grammar*, in which linguistic forms are attributed to contextual meanings. Building upon Halliday and Hasan’s (1976) conceptions of coherence and cohesion, Oller’s notion of grammatical form is evaluated on both the sentential level and discourse level, and is credited with being the first attempt to link pragmatic use to linguistic form (Purpura, 2004).

Canale and Swain (1980) and Canale (1983) provide the first comprehensive model of communicative competence, intended for both instructional and assessment purposes (Celce-Murcia et. al, 1995). The model consists of four components: *grammatical competence, strategic competence, sociocultural competence, and discourse competence*. Grammatical competence, as defined by the authors, constitutes the “language code,” including grammatical rules, vocabulary, pronunciation, and spelling, among others. Bachman (1990) and Bachman and Palmer (1996) further elaborate Canale and Swain’s (1980) model by dividing language knowledge into *organizational knowledge* (textual and grammatical knowledge) and *pragmatic knowledge* (lexical, functional, and sociocultural knowledge). Bachman’s (1990) and Bachman and Palmer’s (1996) notion of grammatical knowledge closely mirrors that of Canale and Swain’s (1980) grammatical competence, where one’s grammatical knowledge is restricted to the subsentential and sentential levels. Textual knowledge, similar to Canale and Swain’s (1980) discourse competence, is concerned with the structure of language at the discourse level. Similar to earlier models, this model places knowledge of grammatical meaning as a separate component of
pragmatic knowledge.

Celce-Murcia et. al (1995) refine Canale and Swain’s (1980) and Bachman and Palmer’s (1996) models by depicting communicative competence as a triangle with *discourse competence* bounded by the three points of the triangle. Each point denotes a particular competence: *sociocultural competence, linguistic competence, and actional competence*. Finally, *strategic competence* is portrayed as a circle surrounding the triangle. Bidirectional arrows between the competencies in the points of the triangle and discourse competence illustrate how the “lexico-grammatical building blocks, the actional organizing skills of communicative intent, and the sociocultural context come together and shape the discourse, which, in turn, also shapes each of the other three components” (p. 9). Unlike previous models, the one proposed by Celce-Murcia et. al (1995) connects Bachman and Palmer’s (1996) grammatical knowledge and lexical knowledge to form *linguistic competence*. It remains unclear, however, exactly how knowledge of grammatical meaning operates under linguistic competence, particularly when lexical knowledge “appropriately belongs to more than one area” (p. 17). Similarly, Larsen-Freeman (1991, 1997) conceptualizes grammar as three interrelated components consisting of form (syntax), meaning (semantics), and use (pragmatics). As intuitive as the model appears, Purpura (2004) remarks that the distinction between grammar and language is not made clear, nor is the model supported by empirical findings.

A more recent model proposed by Purpura (2004) clearly delineates the nature of...
knowledge of grammatical form and knowledge of grammatical meaning. Knowledge of grammatical form and knowledge of grammatical meaning are further divided into subsentential, sentential, suprasentential and discourse levels. At the subsentential and sentential levels are phonological and graphological forms and meanings, lexical forms and meanings, and morphosyntactic forms and meanings. Cohesive forms and meanings, information management forms and meanings, and interactional forms and meanings comprise the suprasentential and discourse levels. In Purpura’s model, grammatical form refers to the syntactocentric approaches included in previous models. Grammatical meaning, on the other hand, is considered the literal meaning of the linguistic forms. Arguing that pragmatic knowledge is a separate, but related component of language knowledge, Purpura positions pragmatics as an independent feature that nevertheless interacts with grammar knowledge. Finally, Purpura distinguishes between ability and knowledge, where knowledge “refers to a set of informational structures” and ability “involves the capacity to use these informational structures” (p. 86). Therefore, Purpura’s notion of grammatical ability incorporates strategic competence of prior models.

In defining the construct of grammatical knowledge in our own model of speaking ability, we draw on Purpura’s (2004) distinction between knowledge of grammatical form and knowledge of grammatical meaning. As the Oral Communication Video Assessment seeks to diagnose general and specific areas of improvement in speaking ability, lexical forms and meanings, morphosyntactic forms and meanings, cohesive forms and meanings, and information
management forms and meanings are pertinent to our purposes. However, due to the primarily monologic task types of the assessment, interactional forms and meanings are not included in our construct. Moreover, as will be discussed in the following section, phonological forms and meanings are operationalized as an independent component, much like pragmatic knowledge.

**Phonological Knowledge**

Defining phonological knowledge is an unusually elusive endeavor, especially since emphasis placed on pronunciation pedagogy has “been left behind” other areas of L2 instruction (Isaacs, 2014, p. 143). In fact, the difficulty in defining phonological knowledge appears to be closely aligned with historical perspectives on L2 pedagogical approaches. Yoshida (2004) writes of three stages in pronunciation-based teaching: “(1) the outset period, (2) the neglected period, and (3) the resurgence period and the present” (p. 3). L2 English instruction in the 1940s, 1950s, and 1960s placed a great deal of value in pronunciation; instruction included a range of pronunciation exercises, including drills, dialogues, and imitation. Focus was on both segmental and suprasegmental features, and corrective feedback was extensive (Morley, 1991). The shift away from audiolingualism and the oral approach in the 1940s and 1950s, however, resulted in less emphasis placed on native-like pronunciation, with a focus primarily directed towards communicative competence. Due to beliefs about the importance of grammar and vocabulary, in addition to the difficulty in acquiring native-like pronunciation, pronunciation instruction was
ignored in many language courses (Celce-Murcia et al., 2010). As Isaacs (2014) notes, the role of pronunciation in L2 English pedagogy, as well as pronunciation research, has only recently been reexamined in a more generous light. Arguing for the importance of intelligible pronunciation on the part of non-native speakers, Morley (1991) asserts that teaching pronunciation is intrinsic to “communicative empowerment- effective language use that will help them not just to survive, but to succeed” (p. 489). Indeed, a plethora of variables, including age, aptitude, native language, and sociocultural and psychological factors can play a role in L2 phonology acquisition (Yoshida, 2004).

Given the historical neglect of L2 pronunciation instruction, it may be unsurprising that there is little consensus as to how to appropriately define phonological knowledge. Levis (2006) notes two competing ideologies: the “nativeness principle” and the “intelligibility principle” (p. 370). “Nativeness” refers to helping L2 learners obtain native-like pronunciation through removing L1 accent qualities from their speech. While popular in the age of audiolingualism, this form of pronunciation pedagogy has fallen out of favor. Instead, most researchers encourage the intelligibility principle, but even this may be particularly problematic because researchers have defined intelligibility in a multitude of ways (Isaacs 2008a). Indeed, theorists as early as Lado (1961) have noted underlying challenges in defining intelligibility, especially since intelligibility can be dependent on whom the listener is. Recent definitions of intelligibility have ranged from broad interpretations to tightly narrowed frameworks. For example, while Morley (1994)
includes listener effort and distracting features as part of intelligibility, Derwing and Munro (1997) view intelligibility as the extent to which a listener understands a speaker’s speech independent of effort required.

Clearly, if many theorists define intelligibility differently, then assessing phonological knowledge will lead to myriad results. Unremarkably, this is the case: determining exactly what phonological knowledge is has led to multiple forms of measurement, often assessing different constructs. Derwing and Munro (1997) found that rubrics that incorporate accentedness (i.e., how much a non-native speaker’s speech resembles a native speaker norm) into intelligibility ratings are actually confounding two separate dimensions. Isaacs (2008b) notes that other rating scales opt for “comprehensibility,” a typically scalar assessment that measures “listener perceptions of ease of understanding non-native speech” (p. 558). Intelligibility, on the other hand, can be more objectively scored; a common form of intelligibility assessment has been to transcribe as much non-native speech as possible in a recorded sample (Isaacs, 2008a, 2008b).

One way to more objectively define and assess intelligibility has been to view it through the lens of segmental and suprasegmental features. Early pronunciation instruction tended to rely more on segmental features through practicing difficult consonant and vowel contrasts. Minimal pair exercises that try to encourage awareness by L2 learners of two similar segmental features offer one example. The notion of functional load, or the “gauge of the frequency which two phonemes contrast in all possible environments” (King, 1967, p. 831), factored prominently in
pronunciation pedagogy. Further refining the functional load principle, applied linguists such as Brown (1988) and Catford (1987) have ranked the importance of specific phonemic contrasts in English pronunciation.

On the other hand, more recent research has suggested that suprasegmental features heavily influence overall intelligibility (Isaacs, 2008b). Hanh (2004) found evidence that improper placement of primary stress in syllables in discourse by international teaching assistants impeded communication. Likewise, McNerney and Mendelsohn (1992) argue that suprasegmental features should take precedence over segmental features in accelerated pronunciation courses, because they “have the greatest impact on the comprehensibility of learners’ English” (p. 186). Anderson-Hsieh et al. (1992) had ESL teachers rate speech samples of 60 speakers from 11 different language groups and found that “deviant” pronunciation areas included segments, syllable structure, and prosody.

However, more recent research indicates that, although suprasegmental features are a fundamental component to pronunciation instruction, explicit segmental instruction is better than none at all (Derwing, Munro, & Wiebe 1998). In fact, current views on pronunciation instruction with respect to segmental and suprasegmental features has taken a more balanced view (Celce-Murcia et al., 2010). To add empirical evidence to this claim, Isaacs (2008b) found that segmental and suprasegmental features factored into the overall intelligibility of an L2 English speaker. Perhaps even more significantly, intelligibility, defined through being able to correctly
transcribe words of L2 English speakers, is a satisfactory form of measuring L2 English speaker pronunciation, as opposed to comprehensibility.

A final word about pronunciation is determining how phonological knowledge fits into a theoretical model. Yoshida (2004) observes that the resurgence of pronunciation pedagogy starting in the 1970s has allowed pronunciation a place in every model that explains communicative competence. Probably the most influential of all models describing pronunciation, Lado’s (1961) skills-and-elements model of language knowledge labeled phonology as one of three elements within the context of listening, reading, speaking, and writing. However, even though Lado (1961) wrote extensively about pronunciation assessment, the model only addressed form, and not meaning. Indeed, although Lado’s (1961) model lacks current perspectives on pronunciation instruction, Isaacs (2014) notes that Lado’s (1961) model is the “most comprehensive treatment of L2 pronunciation assessment to date” (p. 142).

More recent models have also sought to incorporate phonology with varying results. Canale and Swain (1980) and Canale (1983), in their model of communicative competence, included a category labeled grammatical competence. However, Celce-Murcia et al. (1995) noted that the label grammatical, while certainly accounting for syntax and morphology, left room for ambiguity with respect to phonology. In the model proposed by Celce-Murcia et al. (1995), linguistic competence replaced grammatical competence. Under this competency, phonology was broken down into segmental (phonemes) and suprasegmental (rhythm, stress, and intonation)
subcomponents. A key issue with where to put pronunciation into a theoretical model has been made evident by Bachman and Palmer (1982), who claim that phonology and graphology are viewed “more as channels than as components” (p. 450). Despite pronunciation being described as a channel as opposed to a component of communicative competence, Bachman’s (1990) model still incorporates phonology/graphology. Unlike previous models, Purpura’s (2004) model built upon Bachman’s (1990) and Bachman and Palmer’s (1996) in that it separated form and meaning. Purpura (2004) positioned phonological and graphological form and phonological and graphological meaning within the sentential levels of knowledge of grammatical form and knowledge of grammatical meaning, respectively.

In light of the above models, we have chosen to adapt and modify Purpura’s (2004) model with respect to the positioning of phonological knowledge. While we recognize that phonological knowledge would normally fall under knowledge of grammatical form and knowledge of grammatical meaning, for purposes of constructing a simpler model strictly addressing the goals of the assessment, the current model places phonological knowledge as an overarching component (Celce-Murcia et al., 1995). Knowledge of phonological form and knowledge of phonological meaning are represented in the model, and each knowledge type consists of segmental and suprasegmental features. In each area, the segmental level is comprised of individual speech sounds, while the suprasegmental level consists of word and sentence stress, intonation, and tone. Segmental and suprasegmental features are present in the
model in accordance with research demonstrating the importance of both in successful intelligibility of L2 English speakers. Furthermore,

Although not present in the model, our framework for interpreting phonological knowledge derives from Isaacs’ (2008a, 2008b) definition of intelligibility, which focuses on perceiving and producing phonological features regardless of listener effort or accent.

Pragmatic Knowledge

Pragmatic knowledge cannot be observed directly, but we can infer its properties by observing how it manifests itself in contextual, sociolinguistic, sociocultural, and psychological meanings. Before operationalizing these meanings into test tasks, it is necessary to define the properties that make up pragmatic knowledge. By having a clear definition of pragmatic knowledge, we can begin to think of how the various pragmatic meanings can be organized into test content and test tasks. The following is a short review from some of the major theories concerning pragmatic knowledge over the last 50 years. Through this review, a selection will be put together to build a construct most suitable for the test takers of the Oral Communication Video Assessment.

Arguably, ever since Austin’s (1962) notion of performative utterances and Searle’s (1976) development of illocutionary speech acts, along with Grice’s (1975) conversational implicatures and Cooperative Principle, the domain of language learning in terms of pragmatics has never
been the same. A performativive does not merely describe reality but performs some action on it. For example, *I hereby pronounce you husband and wife* is not describing some state of affairs in the world but acting upon it *vis-à-vis* words and utterances. Within this world of performatives there are indirect illocutionary speech acts that we perform on a daily basis with little thought, but for second language learners, these indirect illocutionary speech acts can be potentially problematic. Consider the following examples: 1. *Do you have a pen?* and 2. *Is Jeff there?* Beyond the propositional content of the question or its locutionary force requiring a simple yes or no answer, an intended meaning in the form of a request is contained within the utterance. It would appear odd if, for example, an interlocutor answered the phone to a request, *Is Jeff there?* with only a *Yes, he’s here* without the understanding that the requester desired to speak to Jeff.

Intended meanings such as these examples are not limited to indirect illocutionary speech acts but can also be seen in conversational implicatures when we violate Grice’s Cooperative Principle. Consider the following example:

A: *Is New York expensive?*
B: *Is the pope Catholic?*

B’s response implies that New York is an expensive place to live in, and that even though B violated Grice’s (1975) Cooperative Principle, B can be understood. This may not be necessarily true for second language learners. Together with Searle’s (1976) speech act theory, learners need to know how language works from the point of view of users, the choices they make, the
constraints they encounter in using language in social interaction, and the effects of their
language use on others in the act of communicating (Crystal, 1985). Echoing this idea and
expanding on it, Yule (1996) defines pragmatics as being concerned with meaning as
communicated by a speaker and interpreted by a listener. Its chief aim is to discern what people
mean by their utterances rather than what the words or phrases in those utterances might mean by
themselves.

Building on Searle’s (1976) speech act theory and Grice’s (1975) conversational
implicatures, Leech & Thomas (1983) define pragmatics through two subdivisions:
pragmalinguistics and sociopragmatics. Pragmalinguistic competence is concerned with the
forms necessary to achieve communicative ends while sociopragmatic competence focuses on
appropriateness in social context. Roever (2004) elaborates on this distinction in that in order to
be pragmatically competent, learners must map their sociopragmatic and pragmalinguistic
knowledge onto each other, and be able to use their knowledge under the constraints of a
communicative situation. While pragmalinguistic competence may implicate the other and vice
versa, mapping the two together is easier said than done. Purpura (2004) writes:

“Pragmatic knowledge presents a major challenge for test developers given that one
utterance can simultaneously encode multiple pragmatic meanings, and many times, without
asking the speaker, it is difficult to determine which meanings were implied, and in fact,
without asking the interlocutor, it is difficult to determine which meanings were actually understood” (p. 77).

This is difficult given that both components of pragmatics are equally necessary, and, according to Leech (1983), both pragmalinguistic and sociopragmatic knowledge are needed by native and non-native speakers alike. Language users who know target language sociopragmatic norms but have no pragmalinguistic knowledge will not be able to participate in a meaningful conversation. On the other hand, by having only pragmalinguistic knowledge, language users may potentially offend others in a speech community that is governed by local pragmatic norms. Because these two components of pragmatics are inextricably bound together, it is difficult to determine the source of pragmatic failure. For example, if an utterance is missing mitigation strategies in performing a request, it is difficult to know if this lack of strategy is due to a pragmalinguistic deficit (lacking knowledge of the forms necessary to achieve communicative ends) or if one knows the correct forms but might not know the forms may be needed in a given context (sociopragmatic deficit). This situation is further complicated by identity maintenance, in which an interlocutor may have both sufficient pragmalinguistic and sociopragmatic knowledge, but resists in order to maintain some subjective belief about how language ought to be used (Ishihara & Tarone, 2009; Siegal, 1994).

Pragmatics is an important component embedded within a larger framework of communicative language ability (Bachman, 1990; Bachman & Palmer, 1996; Canale & Swain,
1980; Purpura, 2004). Having the ability to map sociopragmatic and pragmalinguistic knowledge onto each other is what Bachman and Palmer (1996) call sociolinguistic knowledge. In their model, they define pragmatic knowledge as the ability to create and interpret language in context. One must know the conventions that determine the appropriate use of dialects or varieties, registers, natural or idiomatic expressions, cultural references, and figures of speech (Bachman & Palmer, 1996). In Purpura’s (2004) pragmatic component, sociolinguistic meaning consists of gender, age, status, power, politeness, preferences, or expectations. As for sociocultural meanings, this is the domain of cultural norms, preferences, naturalness, or formulaic expressions. Psychological meaning encompasses affective stance such as sarcasm and attitudinal stance indexes the disposition of the speaker towards the utterance. Finally, contextual meaning is governed by in-group interpersonal meanings.

This overview of pragmatics provides a foundation for operationalizing the pragmatic knowledge component in the Oral Communication Video Assessment. Purpura’s (2004) contextual control is used to recognize how in-group and interpersonal relationships govern the use of Grice’s (1975) Cooperative Principle, as well as when the Cooperative Principle is violated via a conversational implicature that conveys an intended meaning. We applied Purpura’s sociolinguistic control, which encompasses gender, age, status, power, preferences, or politeness markers that determine communicative strategies. Pragmalinguistics in this test construct denotes the grammar forms necessary to achieve communicative ends. Purpura’s
sociocultural control indexes American pragmatic norms found in the local speech community where its members regulate what is or is not appropriate language use. Sociopragmatics is concerned with how social relationships in context determine which pragmalinguistic forms to encode. Finally, Purpura’s psychological control consists of the attitudinal and affective stance of the speaker.

A Theoretical Model of Speaking Ability

This literature review provides a rationale for the operationalization of the speaking ability construct. Speaking ability is composed of three observable components. First, grammatical control is defined by knowledge of grammatical form and knowledge of grammatical meaning. Second, phonological control consists of knowledge of phonological form and knowledge of phonological meaning. Finally, pragmatic control is explained by pragmalinguistic and sociopragmatic knowledge. The theoretical model of speaking ability used in the Tools for Clear Speech Oral Communication Video Assessment is represented in Figure 1 below.
REFERENCES


