

Pragmatic Knowledge and Ability in the Applied Linguistics and Second Language Assessment Literature: A Review

Fred S. Tsutagawa¹

Teachers College, Columbia University

ABSTRACT

Almost forty years after Stalnaker's (1974) initial call for "the development and application of a pragmatic theory in which detailed explanations of phenomena relating to linguistic contexts can be given" (p. 214), a unifying construct definition and theory still does not exist to explain fully pragmatic knowledge and ability (Roever, 2011, 2013). Given that test validity and reliability are crucial concerns in the field of second language (L2) assessment, it is important that we try to develop a more comprehensive theoretical construct, so that tests of pragmatic ability will have more construct validity, better represent the target language use (TLU) domain of the test construct, and thus have results that are hopefully more generalizable and useful to stakeholders. This paper presents a brief review of the extensive pragmatics literature and examines how the construct of *pragmatic knowledge and ability* has been operationalized from the perspective of theoretical and applied linguists. The paper presents (1) an introduction to the theoretical foundations of pragmatics, (2) a brief survey of applied empirical studies in L2 pragmatics, and (3) an overview of how the L2 assessment field has conceptualized pragmatic knowledge and ability, looking also at the various theoretical frameworks that have been proposed and utilized. Validation studies for the assessment models, although scarce, are discussed where appropriate. Lastly, some general ideas and suggestions are offered to help lead us towards a unifying construct definition of *pragmatic knowledge and ability*.

INTRODUCTION

In the introduction to one of the first pragmatics readers compiled, Davis (1991) presents the following puzzle: "We carry on conversation, and most of the time we are able to understand one another. How do we do this? What capacities, faculties, and knowledge enable us to communicate with one another with relative ease?" (p. 4). Part of the answer lies in our knowledge of the syntax (i.e., the linguistic forms, or grammar, of a language), as well as semantics (i.e., how linguistic forms connect with things in the world to create meanings) of language (Levinson, 1983; Yule, 1996). However, syntax and semantics must not explain the whole story, as we frequently *do* things with words (Austin, 1962), sometimes without saying anything explicitly. Thus, another underlying component of language ability exists that enables us to extract meanings from context, especially when words go unsaid; this aspect is termed *pragmatic knowledge and ability*.

¹ Fred Tsutagawa is currently a doctoral student in the Applied Linguistics Program at Teachers College, Columbia University. His research interests include pragmatics, second language assessment, and classroom-based assessment. Correspondence should be sent to fst2102@tc.columbia.edu

Pragmatics was first termed by Charles W. Morris (1938) to describe one of the three major areas of *semiotics*, the study of signs and symbols and their use and interpretation (Levinson, 1983). Since then, pragmatics has evolved, more specifically defined as “the study of the conditions of human language uses as these are determined by the context of society” (Mey, 2011, p. 42)—or more simply, “the study of meaning in context” (Archer & Grundy, 2011, p. 2). A generalized historical review of the pragmatics literature shows that the study of pragmatics moved from primarily philosophical to cognitive concerns, and subsequently to cognitive and sociocultural approaches, with applications to sociocultural concerns giving rise to applied research in second language (L2) pragmatics (Kecskes, 2011). Almost forty years after the initial call for “the development and application of a pragmatic theory in which detailed explanations of phenomena relating to linguistic contexts can be given” (Stalnaker, 1974, p. 214), a complete, unifying construct definition and theory still does not exist (Roever, 2011, 2013). As a result, it is necessary to try to fit the pieces together, bit by bit, from the extensive pragmatics research literature. This paper will review the pragmatics literature by (1) providing a historical review of some of the key theoretical concepts in the study of pragmatics, (2) briefly examining empirical L2 pragmatics studies from the broad field of applied linguistics in order to delineate the ways in which aspects of pragmatic knowledge and ability have been understood and operationalized, and then (3) looking at how the construct of pragmatic knowledge and ability has been conceptualized, measured, and studied in the field of L2 assessment. Finally, these findings will lead to a discussion of the remaining challenges that are left in the assessment of L2 pragmatics, as we strive to create tests that can provide better construct coverage while being able to capture the interactive nature of authentic language in use.

PRAGMATICS IN THEORETICAL LINGUISTICS

As pragmatics is broadly framed as the study of language use in social contexts, there are many lenses upon which it has been approached and studied. The cornerstone of pragmatics is arguably speech act theory, which was founded by the language philosophers Austin (1962), Grice (1957, 1975) and Searle (1965, 1969, 1975, 1979). These seminal discussions covered foundational concepts on the nature of communication including: (1) the argument that utterances can contain underlying *intentions* (Grice, 1957); (2) the notion that language carries *locutionary force* (in the act of saying something), *illocutionary force* (in the act of doing something with an utterance) and *perlocutionary force* (in the act performed by uttering something) (Austin, 1962; Leech, 1983); (3) the idea that propositions, rules, meanings, and facts vary as social situations change (Searle, 1965, 1969), (4) the observation that speech acts can be performed both directly and indirectly (Searle, 1975; Levinson, 1983); (5) the concept that communication is socially constructed such that conversation operates under the Cooperative Principle (CP), in which contributions are “as informative as is required (for the current purposes of the exchange)” (Grice, 1975, p. 308) while concurrently “not more informative than required” (p. 308); and (6) the reflection that speakers can mean more than what is actually said, as utterances can possess primary and secondary aspects of meaning (Searle, 1979). In short, these early conversations revealed that pragmatic knowledge and ability is an inherently complex socially-constructed phenomenon, leading to the development of several sub-fields of pragmatics that investigate a variety of linguistic topics ranging from direct vs. indirect speech acts (Levinson, 1983; Searle, 1965, 1969, 1975), relevance, deixis and indexicality (Levinson, 1983;

Sperber & Wilson, 1981b; Wilson & Sperber, 1986), presuppositions (Karttunen, 1979; Levinson, 1983; Stalnaker, 1974), metaphor and irony (Lakoff & Johnson, 1980; Searle, 1982; Sperber & Wilson, 1981a), and politeness theory (Brown & Levinson, 1987; Leech, 1983).

Leech (1983) and Thomas (1983) made the important distinction between *pragmalinguistics* and *sociopragmatics*, which has played an important role in the field of assessing pragmatics. *Pragmalinguistics* refers to the grammatical and the linguistic resources people utilize to convey “communicative acts and relational and interpersonal meanings” (Kasper & Rose, 2001, p. 2), while *sociopragmatics* is associated with “proper social behavior” (p. 3) and the “social perceptions underlying participants’ interpretation and performance of communicative action” (p. 2). Regarding the assessment of pragmalinguistic and sociolinguistic knowledge, McNamara & Roever (2006) remark, “[b]ecause of the close connection between [them], it is difficult to design a test that tests pragmalinguistics to the exclusion of sociopragmatics or vice versa” (p. 55). Currently, many more studies focus on pragmalinguistic issues over sociolinguistic ones (Trosborg, 2010). A cursory inspection of the *Journal of Pragmatics* and *Pragmatics* show just how robust pragmatics research has become. However, the remainder of this review will focus on applied L2 pragmatics studies, which have been just as fertile and varied as their theoretical counterparts.

As applied linguists realized that “pragmatics can assist people, as language users, in their endeavors to realize their personal goals in the societal setting in which they live” (Mey, 2001, p. 315), these theoretical concepts quickly branched out into two highly productive and closely related applied research strands: *cross-cultural pragmatics* (Blum-Kulka, House, & Kasper, 1989) and *interlanguage pragmatics* (Kasper & Blum-Kulka, 1993). Enriched by several theoretical lines of inquiry, cross-cultural pragmatics looks at intralingual variation in speech acts and how these differences affect communication (Blum-Kulka, House, & Kasper, 1989). Interlanguage pragmatics (ILP) is an “action-theoretical” interdisciplinary field informed from second language acquisition research, which focuses on the developing pragmatic competence of L2 learners. ILP examines, for instance, pragmatic transfer from the native language to the L2, and explores the phenomenon of pragmatic failure (Kasper & Blum-Kulka, 1993). This paper will consider the cross-cultural, ILP, and L2 assessment pragmatics literature. We now turn to empirical studies that have examined pragmatic knowledge and ability in applied linguistics. To help facilitate the discussion we organize the literature based on the following four themes: 1) speech act studies, 2) implicature studies, 3) studies on formulaic expressions/routines, and 4) pragmatic awareness-raising studies.

L2 PRAGMATIC KNOWLEDGE AND ABILITY: EMPIRICAL RESEARCH STUDIES

Speech Acts

Described as “one of the most compelling notions in the study of language use” (Blum-Kulka, House, & Kasper, 1989, p. 1), speech acts are important in pragmatics research, not only because of the influence of speech act theory in the history of the study of pragmatics, but also because speech acts can carry social implications (Ervin-Tripp, 1976) while still being discrete points of contrast and contact between cultures and languages (Blum-Kulka, House, & Kasper, 1989). It is not surprising to find, then, that the bulk of empirical L2 pragmatics research utilize speech acts to investigate various pragmatic phenomena. Requests or directives are the most

commonly studied speech act (e.g. Blum-Kulka, 1989; Blum-Kulka, Danet, & Gerson, 1985; Blum-Kulka & House, 1989; Brown & Levinson, 1978; Clark & Lucy, 1975; Ervin-Tripp, 1976; Faerch & Kasper, 1989; Gibbs, 1985; Hassall, 1999; Omar, 2008; Takahashi, 2001; Weizman, 1989). There are also robust investigations into various aspects of apologies (Lipson, 1994; Olshtain, 1989; Olshtain & Cohen, 1983; Vollmer & Olshtain, 1989; Wolfson, Marmor, & Jones, 1989) and compliments (Nelson, Al-Batal, & Echols, 1996; Manes, 1983; Manes & Wolfson, 1981; Rose & Ng Kwai-fun, 2001; Wolfson, 1981; Wolfson & Manes, 1980). Collecting compliment and compliment response data are noted to be particularly interesting because they usually entail more complex multi-turn sequences that reflect sociocultural values (Trosborg, 2010). Other speech act paradigms such as refusals (Beebe, Takahashi, & Uliss-Weltz, 1990; Keshavarz, Eslami, & Ghahraman, 2008; Kondo, 2008; Kwon, 2004), suggestions (Martínez-Flor, 2006), gratitude (Eisenstein & Bodman, 1993), and complaints (Olshtain & Weinbach, 1987) have also been explored. Additional dimensions of speech acts, including aspects of directness/indirectness (Blum-Kulka, 1989), degrees of politeness (Brown & Levinson, 1978; Brown & Levinson, 1987; Cook, 2001; House, 1989), and instances of pragmatic failure (Bardovi-Harlig & Dörnyei, 1998; Blum-Kulka & Olshtain, 1986; Jaworski, 1994; Kasanga & Lwanga-Lumu, 2007; Nelson, Al-Batal, & Echols, 1996; Yan & Zhuang, 2010) are frequently examined as well.

Criticisms from practitioners of the speech act paradigm note the fact that the cross-cultural and ILP literature contains redundant results of specific speech acts (e.g., requests and compliments), elicitation tasks (e.g., the DCT), and specific research populations (Trosborg, 2010). The greatest recognized weakness, however, is the realization that the construct of *pragmatic knowledge and ability* is too narrow in this framework (Bardovi-Harlig, 2010). As Grabowski (2009) argues, “reducing pragmatics to speech acts, and trying to completely separate it from grammatical knowledge artificially limits a domain of language knowledge that is much more comprehensive and dynamic” (pp. 45-46) in real life interaction.

One of the most important and in-depth studies on cross-cultural and intralingual variation was completed on requests and apologies in the Cross-Cultural Speech Act Realization Project (CCSARP) (Blum-Kulka, House, & Kasper, 1989). The goal of the study was to examine seven languages or varieties—Australian English, American English, British English, Canadian French, Danish, German, and Hebrew—collecting data from both native and non-native university students in an attempt to observe systematic differences in realization patterns in regards to three types of variability: 1) intra-cultural or situational variability, 2) cross-cultural variability, and 3) individual variability (Blum-Kulka, House, & Kasper, 1989; Blum-Kulka & Olshtain, 1984). The instrument used for data collection was a written DCT, with the choice rationalized on both practical and theoretical grounds (Blum-Kulka, House, & Kasper, 1989). This DCT (Blum-Kulka, 1982; Levenston, 1975) consisted of scripted dialogs that included short descriptions of a situation and other contextual information such as the setting and the nature of the relationship between participants; eight situations elicited requests, and eight generated apologies (Blum-Kulka, House, & Kasper, 1989), with all situations containing a missing turn in a dialogue that respondents were asked to complete. Considering the importance of the CCSARP studies, it is surprising that few studies have looked critically at validity and generalizability concerns of its framework and methodology (de Paiva, 2010; Li, 2009; Rose, 1992; Rintell & Mitchell, 1989; Wolfson, Marmor, & Jones, 1989; Wouk, 2006), although these issues were somewhat typical in speech act research.

Implicature

While it will only be discussed briefly here, conversational implicature studies represent a branch of pragmatics research that investigates meanings in language use. Mey (2001) succinctly defines it as such: “A *conversational implicature* is, therefore, something which is implied in conversation, that is, something which is left implicit in actual language use” (p. 99, original emphasis). Not all implicatures found in natural communication are conversational, as meanings can be manifested and attributed to signs and gestures as well (Mey, 2001). Important studies on implicature have been conducted by Bouton (1988, 1994), Roever (2005), and Taguchi (2005). Although not as robust a body of research as the literature on speech acts, these studies have helped us view a part of pragmatic knowledge and ability as the ability to understand meanings and intentions from context. As we will see later when discussing theoretical frameworks of pragmatic knowledge and ability, conversational implicature is a significant part of the construct developed here.

Formulaic Expressions & Routines

According to Bardovi-Harlig (2008), the term *formula* is vaguely defined in the L2 pragmatics literature, and thus can refer to one of three uses: “one which describes a feature of the acquisition process, one which describes the end point, or target, and one which describes components of a speech act” (p. 3). It is therefore important to be aware of the three strands of research. The latter, sociopragmatic usage of *formulaic expressions* is adopted here, and in general, the instruction of such form-function expressions is known to have several effects on learners. First, when taught formulae, learners can typically overgeneralize, undergeneralize, and/or misuse them as their knowledge of the proper use contexts is limited (Bardovi-Harlig, 2008). It is also observed that length of stay in the target language and culture “contributes more to acquisition of target formulas than proficiency” (p. 19; cf. Schauer, 2006). This line of empirical studies on formulaic expressions and routines has focused on the efficacy of explicit and/or implicit, instructional techniques (Bardovi-Harlig & Nickels, 2011; House, 1996; Tateyama, Kasper, Mui, Tay, & Thananart, 1997; Tateyama, 2001; Yoshimi, 2001). Overall, explicit instruction has been shown to be the more effective method, but as Takahashi’s (2010) review of 49 intervention studies shows, this is not always the case. Issues of how researchers operationalize implicit vs. explicit instruction is a concern, and depending on the type of pragmatic knowledge and ability that is being targeted, implicit measures alone may suffice. Such disparate conclusions earmark the need “to study the difficulty of different speech acts” (Liu, 2010) and be able to place them on a continuum of difficulty as we continue to develop our understanding of pragmatic knowledge and ability. Producing such a taxonomy would be beneficial to both instructors and learners, as we would be able to know, for example, why apologies are easier than refusals (Hudson, 2001a).

Awareness-raising

Pellegrino, Chudowsky, and Glaser (2001) comment that “[o]ne of the most important aspects of cognition is metacognition—the process of reflecting on and directing one’s own thinking” (p. 4) (cf. Verschueren, 2000). According to Purpura (2013), metacognitive strategies describe the process of how one thinks and mentally regulates actions or behaviors. A

metacognitive strategy in a language classroom could then look like, for example, a discussion of or reflection upon the rules of language use. Although very few studies have formally looked at raising students' pragmatic awareness through instruction (Eslami & Eslami-Rasekh, 2008; House, 1996), the results so far are promising. For example, Eslami and Eslami-Rasekh (2008) provided pragmatic awareness instruction to advanced learners of English in Iran, and they found that both pragmatic awareness and production were acquired in the FL classroom setting. House (1996) also gave explicit instruction to raise the awareness of advanced German learners of English over the course of 14 weeks, and for all but one feature, found positive gains in pragmatic fluency as well. More studies need to be done, but as Bardovi-Harlig and Nickels (2011) and Dantas-Whitney (2011) provide helpful, ready-to-use classroom activities to raise such language awareness, we hope that additional studies will follow.

To summarize this section, these studies have first shown us that a broader, and validated, construct for pragmatic knowledge and ability is needed beyond that of the speech act. We have also learned that pragmatic meanings are a key component of the construct and that we need to be able to place various pragmatic functions and/or meanings into some kind of hierarchical matrix. Finally, we know that metapragmatic concerns should also be included in the construct.

L2 PRAGMATIC KNOWLEDGE AND ABILITY: CONCEPTUALIZATIONS AND MEASUREMENT

Researchers in language assessment have attempted to provide theoretical constructs for pragmatic knowledge and ability. In language assessment, test validity and reliability are crucial issues that are constantly under discussion, negotiated and built upon when constructing assessment instruments (Bachman & Palmer, 1996; McNamara, 2000). Validity, or more accurately, construct validity, refers to "the extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s), we want to measure" (Bachman & Palmer, 1996, p. 21). Thus, in general, the more valid a test is deemed via the evidence and argumentation specified, the more generalizable the results are. In order for a test to have strong evidence for construct validity, the target language use (TLU) domain of the test construct, i.e. the "aspects of the knowledge or skill possessed by the candidate which are being measured" (McNamara, 2000, p. 13), should also be represented with an appropriate construct definition. As Yamashita (2008) notes, "[d]efining the construct to include only one area of language knowledge is inappropriately narrow, as the ... TLU domain involves other areas of language knowledge" (p. 203). We have been examining the construct of pragmatic knowledge and ability, but as is evident from the general applied linguistics research above, the construct has been underrepresented and is limited to narrow TLU speech act domains such as requests and apologies, and based on very limited sample populations. This section will review the empirical L2 testing literature. We will begin with an overview of how the L2 assessment field has conceptualized pragmatic knowledge and ability and look at the more comprehensive theoretical models that have been proposed. Next, we will examine validation studies for these models, although scarce, and then look at other L2 assessment studies that do not adhere to a strong construct definition of pragmatic knowledge and ability.

Frameworks for Pragmatic Knowledge and Ability

The importance of including pragmatic knowledge and ability as a component of language ability has been noted from the outset in the L2 assessment literature. In Lado's (1961) description of complex systems of communication, he explains that cultural meanings are also encoded within the language, "which are common to other members of the cultural community" (p. 5). For example, the term *football* has different meanings in the U.S. and Brazil. Thus, outside of the cultural boundary, every person forms individual meanings that are shaped by personal life experiences. Lado declared that "cultural meaning" and "linguistic meaning" are the two main parts of human language (pp. 5-6), and thus the idea of cultural meanings as a part of pragmatic knowledge was introduced.

Oller (1979) proposed that pragmatic knowledge and ability were incorporated into the grammatical system of the language user in a theory he called 'pragmatic expectancy grammar.' He claimed that "somehow [linguistic and pragmatic meaning sequences are] indexed with reference to their appropriateness to extralinguistic contexts" (p. 24). While he was neither able to explain the mechanism of how it worked nor give a detailed description of various pragmatic components, this work is important, because it was "the first serious attempt in language testing to define grammar as an integration of linguistic form and pragmatic use as this relates to context" (Purpura, 2004, p. 53).

Although Hymes (1964, 1972) had formally coined the term *communicative competence*, it wasn't until Canale and Swain (1980) published their seminal piece on a theory of communicative competence for L2 teaching and testing that it became commonplace. The theory identified four primary competencies that contribute to overall *communicative competence*: *grammatical competence*, *sociolinguistic competence*, *strategic competence*, and *discourse competence* (Canale, 1983; Canale & Swain, 1980). In this framework for communicative competence, *pragmatic knowledge and ability* was subsumed under *sociolinguistic competence*, describing it as "sociocultural rules of use and rules of discourse" (p. 30). However, specific elements were still not elaborated.

Building from this framework of communicative competence, Bachman (1990) proposed a model for communicative language ability (CLA), in which pragmatic competence was described as one of the two main components of language competence. Bachman (1990) then divided *pragmatic competence* into two further components: 1) *illocutionary competence* and 2) *sociolinguistic competence*. This model was later refined in Bachman and Palmer's (1996) construct of language ability, with the key term of *competence* replaced with *knowledge*. In this revised view, language ability is comprised of *language knowledge* and *strategic competence*, and subsumed under *language knowledge* are organizational and pragmatic knowledge bases. As the first model to specify pragmatic components, the two areas under *pragmatic knowledge* are: 1) *functional knowledge*, which is subdivided into the knowledge of ideational, manipulative, heuristic, and imaginative functions, and 2) *sociolinguistic knowledge*, which includes knowledge of dialects/varieties, registers, expressions, cultural references and figures of speech.

Purpura (2004) recognizes "the close relationship among grammar, semantics and pragmatics" (p. 58), and similar to Bachman and Palmer (1996), identifies that language knowledge for communication is comprised of the two main areas of *grammatical knowledge* and *pragmatic knowledge*. In his model, however, these components are identified as distinct but related, as Oller (1979) had previously observed. While utterances and communicative acts have more obvious literal meanings, Purpura (2004) argues that "on the pragmatic level, grammatical

forms and meanings may carry one or more implied meanings depending on the context of language use” (p. 80) that can involve a wide array of meanings. Pragmatic knowledge is therefore acknowledged as highly complex. It is indelibly linked to grammatical forms and meanings, requiring a deep understanding of the full range of contexts that languages, cultures and societies share. Rather than organize pragmatic knowledge in terms of communicative functions (Bachman & Palmer, 1996), Purpura framed pragmatic knowledge in terms of pragmatic meaning components: *contextual meanings* (e.g., interpersonal, in-group references, and metaphors), *sociolinguistic meanings* (e.g., social identity markers of age, gender and group, social meanings, language variation and registers), *sociocultural meanings* (e.g., cultural meanings, and cultural norms, preferences and expectations), *psychological meanings* (e.g., attitude and affect), and *rhetorical meanings* (e.g., coherence and genres).

McNamara and Roever (2006) argue that Bachman’s (1990) model of communicative language ability is “clearly primarily cognitive and psychological” (McNamara & Roever, 2006, p. 32), despite including various social dimensions. They argue, therefore, for “a theory of the social context in its own right...that is not primarily concerned with the cognitive demands of the setting on the candidate” (pp. 32-33). From this standpoint, it would appear that Purpura’s (2004) model, with his emphasis on pragmatic meanings in various social contexts, better captures the global concept of pragmatic knowledge and ability. His language model also conveniently subsumes Leech’s (1983) popular pragmalinguistic and sociopragmatic distinctions, with the former equivalent to Purpura’s model of *grammatical meanings* (i.e., the learner’s understanding of phonological/graphological, lexical, morphosyntactic, cohesive, and interactional meanings from the same grammatical forms) and the latter to *pragmatic knowledge* (i.e., the learner’s ability to understand degrees of appropriateness, conventionality, naturalness, and acceptability of pragmatic meanings from varied social contexts). Additionally, Purpura’s framework does not seek to separate pragmatic knowledge from grammatical knowledge. Instead, it shows that both interact (*cf.* Oller, 1979), thus falling in line with L2 pragmatic development research that has shown “two seemingly contradictory tendencies in adult learners’ development of pragmatic and grammatical knowledge: evidence that pragmatics precedes grammar, and evidence that grammar precedes pragmatics” (Kasper & Rose, 2002, p. 189). Niezgoda and Röver (2001) also could not determine whether pragmatic competence preceded grammatical competence, or vice versa, in their study of environmental factors that could influence learners’ awareness of pragmatic and grammatical errors. In summary, Purpura’s (2004) model (see p. 91) is comprehensive, as it takes into account grammatical forms, grammatical meanings, and pragmatic meanings. In other words:

...in a comprehensive model of L2 proficiency, grammatical forms together with their literal and intended meanings (i.e., grammatical knowledge) provide the fundamental resources for communicating contextual implicatures; metaphor; poetry; social and cultural identity; social and cultural appropriateness—formality, politeness; affective stance—emotionality, irony, humor, sarcasm; and so forth. (Purpura, in press)

Few studies have been performed to validate Bachman and Palmer’s (1996) model of language ability (see p. 63). A structural equation modeling (SEM) study was found to confirm a part of their theory of strategic competence (Phakiti, 2008). To my knowledge, no other research has been done to affirm other components such as topical knowledge and language knowledge. On the other hand, a growing number of studies are being produced that validate various aspects

of Purpura's (2004) model of language/grammatical ability (Ameriks, 2009; Chang, 2004; Grabowski, 2009; Liao, 2009; Vafae, Basheer, & Heitner, 2012). Chang (2004) discovered that both grammatical form and semantic meaning were part of the overarching construct of grammatical knowledge, yet both aspects remained unique and distinct from one another. Ameriks (2009) examined underlying trait structures in the grammar section of the Examination for the Certificate of Proficiency in English (ECPE) using SEM techniques and also determined that grammatical knowledge can be successfully separated into form and meaning components. Vafae et al. (2012) utilized a multi-trait multi-method (MTMM) confirmatory factor analysis to verify the construct validity of a grammar test based on Purpura's model. More relevant to the current review, Liao (2009) and Grabowski (2009) explored the relationship between grammatical and pragmatic knowledge, and both studies found that both knowledge bases can be measured separately. Grabowski's (2009) study is especially noteworthy. It employed four innovative multi-turn, oral production role-play tasks with one native English speaking test partner initiating the performance assessments with a sample population ($N = 102$) of non-native English speakers at varying proficiency levels (intermediate and advanced levels) and native speakers (expert level). Using a mixed methods approach, both quantitative and qualitative analyses were conducted. For the quantitative analysis, many-facet Rasch measurement (MFRM), classical test theory (CTT), and multivariate generalizability (MG) theory were applied. The qualitative analysis portion looked at the response data for sociolinguistic, sociocultural, and psychological meanings. The results of the study "showed evidence of validity of the underlying test construct" (p. 291) with respect to Purpura's (2004) model. *Grammatical knowledge and ability* in terms of accuracy and meaningfulness, and *pragmatic knowledge and ability* as sociolinguistic, sociocultural and psychological appropriateness, were substantiated.

The remaining L2 assessment studies that will be reviewed follow research paradigms that either adhere to outmoded frameworks of language ability or fall in line under the speech act paradigm. In both cases, the construct of pragmatic knowledge and ability is generally underrepresented (Roever, 2011).

Perhaps the most influential of all early pragmatic studies, Hudson, Detmer, and Brown (1992, 1995) developed a series of six instruments to measure the cross-cultural pragmatic ability of Japanese ESL learners in the use of requests, apologies, and refusals. With the written open-ended DCT "adopted as the motivator for the development of the other test instruments" (Hudson, Detmer, & Brown, 1995, p. 7), the first two test measures used paper and pencil indirect measures and were comprised of: 1) cued, multiple choice items following a DCT-type situational prompt, and 2) open-ended free response items following a DCT-type situational prompt. The next assessments included two oral semi-direct measures: 3) a listening laboratory tape response to descriptions of situations, and 4) a structured oral interview and videotaped role-play responses to scenarios. The final two pragmatic test formats were designed as self-assessment measures where: 5) participants self-assessed their performances on the previously tape-recorded listening lab situation responses, and 6) participants directly observed and evaluated their videotaped role-play and interview. The goal in developing many instruments was to observe potential variability from the different data collection procedures. For all three speech acts, the variables of relative power (P), social distance (D), and the degree of imposition (R) were then selected for detailed examination (Hudson, 2001b; Hudson, Detmer, & Brown, 1992, 1995), with Blum-Kulka, House, and Kasper's (1989) CCSARP coding scheme used for the requests and apologies, and the coding scheme by Beebe, Takahashi, and Uliss-Weltz (1990) selected for refusals. While some strategies such as alerters, downgraders and upgraders were

initially identified as being used by both NS and NNS participants, closer examination of individual situational cells revealed that NSs and NNSs differed greatly in their usages of strategies. They concluded that “the data showed how sociopragmatic judgment affects NNS speech act realization” (Hudson, Detmer, & Brown, 1995, p. 49), suggesting that cross-cultural differences appeared to truly exist between the NS American and NNS Japanese populations in how they accomplished requests, apologies, and refusals with their words, and in how they viewed power and social distance relationships.

Hudson, Detmer, and Brown (1992, 1995) created a novel rubric based on appropriacy (*cf.* Beebe, Takahashi, & Uliss-Weltz, 1990; Blum-Kulka, House, & Kasper, 1989). They proposed rating subject performance on the DCTs in relation to appropriacy along six categories: 1) ability to use the correct speech act, 2) typical expressions used, 3) amount of speech and information given, 4) levels of formality, 5) levels of directness, and 6) levels of politeness. NSs were trained to rate NNS responses on a five-point Likert scale, in which only the extreme score values had written indicators. A score of ‘5’ was designated “Completely Appropriate,” whereas a ‘1’ was labeled “Very Unsatisfactory.” With the exception of assessing whether typical expressions were used or not, it should be noted that five of the components relate to pragmalinguistic considerations (Hudson, 2001a). Variables (1) and (3) relate to language correctness, while (4), (5) and (6) were intended to “represent more social aspects of language, although they, too, are certainly expressed through linguistic conventions” (p. 284).

Since the development of the various testing instruments by Hudson, Detmer, and Brown (1992, 1995), Hudson (2001a, 2001b) and others reported some results as to their efficacy in assessing pragmatic ability. Hudson (2001b) found the multiple choice, cued response DCT instrument to be ineffective and unreliable, as was the case for Yamashita (1996) and Yoshitake-Strain’s (1997)’s replication studies as discussed in Brown (2001). Regarding the investigation into the efficacy of self-assessment measures for DCTs, video ratings, and role play self-ratings, Hudson (2001b) “did not find self-assessment to be particularly useful in the assessment of cross cultural pragmatics of Japanese speakers learning English” (p. 68). One possible explanation for this result was his use of a homogeneous sample population, but Yamashita (1996) had similar findings despite examining a more variegated sample of beginning, intermediate, and advanced learners of Japanese as a foreign language. Hudson (2001a) evaluated the effectiveness of written and recorded DCTs, as well as role-play scenarios, among 25 Japanese ESL learners, as these were the instruments that elicited language production. Conceding that some order and practice effects may have affected the overall results, Hudson (2001a) had each of the participants take the test in the following sequence: the language lab DCT, the open-ended written DCT, and finally, the role-play. The results indicated that the small degree of variation across the sample population produced negatively skewed data that proved highly problematic, as the participants were homogeneous both in terms of ability level and cultural background. Nevertheless, the role-play results rated higher than the language lab DCT responses, and the role-play “performed differently from the other two [DCT] methods” (p. 295). They also found apologies to be less demanding than refusals, and that amongst the NS raters, “*amount of language and appropriateness expressions* were rated more harshly than *formality, directness, or politeness*” (p. 295, emphasis in original), where these two variables represented more grammatical and lexical features. Lastly, they also discovered that for both the written and language lab DCT instruments, situations with higher degrees of imposition were significantly more difficult for the participants. In sum, the instruments developed to measure pragmatic ability are only preliminary suggestions for pragmatics assessments, with many unresolved issues remaining including the

need for further validation studies.

Brown (2001) more carefully examined the statistical reliability and validity arguments for Hudson, Detmer, and Brown's (1992, 1995) battery of six pragmatic assessment instruments by comparing the results of these measures from two different replication studies: Yoshitake-Strain's (1997) research on Japanese EFL learners in Tokyo and Yokohama, and Yamashita's (1996) study on English speakers learning Japanese as a second language (JSL). From these analyses, Brown (2001) determined that "the six original English-language versions of the pragmatics tests generally did not work as well as the Japanese translations of those original versions in terms of the amounts of variance they produced, their reliability, and their intercorrelations" (p. 321). Separately, in examining each test's practical advantages and disadvantages, Brown (2001) also provided useful information regarding each instrument's test characteristics in the event that any of these pragmatic tasks are sought to be used in future studies and research.

Billmyer and Varghese's (2000) subsequent DCT study examined sociolinguistic variability due to contextual situations and instrument-induced variability (Kasper & Dahl, 1991; Labov, 1966). They created two DCT instrument versions with six parallel situations each. Version 1 was formatted as typical DCT prompts: short, concise and lacking detailed contextual information, but well known for their ease of use (see Blum-Kulka, 1982; Rose, 1992; Beebe & Cummings, 1996). The enriched prompt version (Version 2) significantly lengthened the prose by including detailed psychosocial components such as the setting, time and place, participants, speaking purpose, and the requestive goal. While they boldly concluded that "the content-enriched DCT prompts elicit more robust external modification and elaboration than do the archetypal content-poor prompts which most DCT studies to date have used" (p. 543), numerous methodological issues make it extremely difficult to accept this claim. But given that this study had considerable potential to provide important practical details regarding test specifications for pragmatic DCT tasks and the subsequent writing of such prompts, further studies on this topic are recommended to verify these findings.

Liu (2006, 2007) measured the requests and apologies of Chinese EFL university students through the development of his own DCT-based test instruments. In Liu (2006), three test methods, a written DCT (WDCT), a discourse self-assessment test (DSAT), and a multiple-choice DCT (MDCT), were authored through a meticulous process of exemplar generation, where through an iterative process with the test takers themselves, the test developers tried to ensure each scenario was authentic to the target audience. Liu (2007) gives a more detailed account of the development and validation of the pilot MDCT measure. Liu (2006) reports that the WDCT and DSAT were found to be highly reliable, whereas the MDCT was shown to be only reasonably so (Liu, 2006, 2007). These studies were an extension of some of the methods developed by Hudson, Detmer, and Brown (1992, 1995), but "with more rigor and a more pragmatically-informed test development process" (p. 168). However, some limitations remained. First, none of the methods investigated required oral production. Also, the tests themselves may not have much applicability to real ESL settings or even EFL settings outside of China, due to the situational prompts being developed in the context of Chinese university life. Finally, Liu's (2006, 2007) sample population was also highly homogeneous, including only Chinese EFL university students at the upper-intermediate and advanced levels.

Undoubtedly, the DCT has been a highly influential and popular test format because of its practicality, usefulness, and ease in gathering a large amount of data from different nationalities quickly (Trosborg, 2010). But while so many pragmatic research studies have utilized this

format, one major criticism of DCTs as a means to assess pragmatic ability comes from McNamara and Roever (2006) who claim that traditional DCTs are inauthentic in terms of reproducing actual conversation because they try to cram “all possible speech act strategies into one gap” (p. 63) rather than let the conversation be co-constructed between interlocutors.

Although limited, some researchers (Grabowski, 2009, for instance) have begun to elicit such oral performance assessments of pragmatic knowledge and ability. Norris (2001) sought to determine whether one aspect of spoken pragmatic ability, that of correctly using German address forms, could possibly be related to language ability levels. Noting the influence of the American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines towards the “continued use of live and simulated oral proficiency interviews (OPIs and SOPIs)” (p. 249) to make various high-stake decisions, he investigated this one aspect of sociolinguistic competence on the German Speaking Test (GST), a SOPI. He found that the test participants exhibited more sociopragmatically correct German address forms as global ACTFL proficiency levels rose, but naturally occurring performance variability at each ACTFL level made it “by no means a categorical or implicational relationship” (p. 281). In fact, Norris (2001) observed that “examinees who performed in otherwise very proficient ways (and were rated accordingly) did not necessarily exhibit knowledge of or control over norms of address behavior” (p. 282).

Other, more innovative, pragmatic assessments are also on the horizon. Roever (2006) began the exploration of testing L2 pragmalinguistic knowledge reliably and practically through web-based instruments. Recognizing that several measures to assess sociopragmatic knowledge were already developed by Hudson, Detmer, and Brown (1995), the research was designed to only measure pragmalinguistic considerations. As pragmalinguistics is concerned with “linguistic strategies for implementing speech intentions and the items necessary to express these intentions” (p. 230), he deemed that the learner’s ability to understand implicatures (Bouton, 1988; 1994), routines, and speech acts would comprise the construct for pragmalinguistic knowledge. He developed an online test in HTML and JavaScript, and was able to administer it to 267 English language learners from various ESL and EFL settings. Contrary to Bouton’s (1988, 1994) claims, however, Roever (2006) found that proficiency level, rather than length of exposure to English as a second or foreign language, appeared to result in greater pragmalinguistic competence. Roever (2006) also posited that while the test was highly practical, the measuring of only pragmalinguistic knowledge (as sociopragmatic competence is equally important) was a major limitation of the instrument (p. 251). As the test required the reading of various situations (i.e., no audio or video prompts were provided) and oral production was not required on the internet-based test, valuable contextual information and a lack of authenticity in terms of responses was noted, respectively. This study is certainly valuable, however, in opening the door to more creative pragmatic test designs that maximize technology in order to produce more authentic, reliable, and valid measures for pragmatics assessment.

For future assessments of pragmatic knowledge and ability, Yamashita (2008) challenges us to rethink the assumptions we have built about how best to study it. She calls for a widening of the construct and proposes that features such as phatic expressions, prosodics, turn-taking and backchanneling, for example, be added. Brown (2008) challenges us to use “statistical analyses that can be used to improve pragmatics tests” (p. 224), such as generalizability (G) theory and decision (D) studies. Roever (2008) then shows how MFRM can be used to analyze test-taker, item, and rater facets. Roever (2010) used a differential item functioning (DIF) approach to determine that test-taker cultural background had a significant effect of bias on the pragmatic test results. In many ways, it would appear that the future is already here as Purpura’s (2004) model

of language/grammatical ability provides us with an expanded construct of pragmatic ability. Likewise, Grabowski (2009) has also been able to successfully validate aspects of the construct via interactive role-plays and G-theory and MFRM analyses.

CONCLUSION

In order for a unified construct of pragmatic knowledge and ability to be formulated, broader models of pragmatic knowledge and ability need to be hypothesized and then tested, going beyond the mere notion of individual speech acts. The most challenging aspect of this endeavor will be creating models that are comprehensive enough to be able to account for the entire range of pragmatic phenomena that is observable in authentic communication, while also addressing metapragmatic concerns and being able to see how different pragmatic functions and meanings relate to one another. Based on the limited but favorable validation studies done on his framework, Purpura's (2004) components of contextual, sociolinguistic, psychological, and rhetorical pragmatic meanings provide an excellent starting point to further investigate future pragmatic test items for various pragmatic meanings. But this survey of the applied linguistics and L2 assessment literature is far from complete. Much work continues to be done in the fields of cross-cultural and interlanguage pragmatics, intercultural pragmatics, and L2 testing, with many opportunities for growth and development. If, however, pragmatics research can unify itself under the umbrella of an overarching theory of knowledge and ability, future results will undoubtedly prove more constructive and productive.

REFERENCES

- Ameriks, Y. S. (2009). *Investigating validity across two test forms of the Examination for the Certificate of Proficiency in English (ECPE): A multi-group structural equation modeling approach*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York, NY.
- Archer, D., & Grundy, P. (2011). Introduction. In D. Archer & P. Grundy (Eds.), *The pragmatics reader* (pp. 1-10). New York, NY: Routledge.
- Austin, J. L. (1962). *How to do things with words: The William James lectures delivered at Harvard University in 1955*. Oxford, UK: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford, UK: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language Testing in Practice*. Oxford, UK: Oxford University Press.
- Bardovi-Harlig, K. (2008). On the role of formulas in the acquisition of L2 pragmatics. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics language learning*, volume 11 (pp. 199-225). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.
- Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, 32(2), 233-259.

- Bardovi-Harlig, K., & Nickels, E. L. (2011). No thanks, I'm full: Raising awareness of expressions of gratitude and formulaic language. In N. R. Houck & D. H. Tatsuki (Eds.), *Pragmatics: Teaching natural conversation* (pp. 23–40). New York, NY: TESOL.
- Beebe, L. M., & Cummings, M. C. (1996). Natural speech data versus written questionnaire data: How data collection method affects speech act performance. In S. M. Gass & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 65–87). Berlin: Mouton de Gruyter. (Original version 1985).
- Beebe, L. M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E. S. Anderson, & S. D. Krashen (Eds.), *On the development of communicative competence in a second language* (pp. 55-74). New York, NY: Newbury House.
- Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics*, 21(4), 517-552.
- Blum-Kulka, S. (1982). Learning how to say what you mean in a second language: A study of speech act performance of learners of Hebrew as a second language. *Applied Linguistics*, 3(1), 29-59.
- Blum-Kulka, S. (1989). Playing it safe: The role of conventionality in indirectness. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 37-70). Norwood, NJ: Ablex Publishing Corp.
- Blum-Kulka, S., Danet, B., & Gerson, R. (1985). The language of requesting in Israeli society. In J. Forgas (Ed.), *Language and social situation* (pp. 113-141). New York: Springer Verlag.
- Blum-Kulka, S., & House, J. (1989). Cross-cultural and situational variation in requesting behavior. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 123-154). Norwood, NJ: Ablex Publishing Corp.
- Blum-Kulka, S., House, J., & Kasper, G. (1989). Investigating cross-cultural pragmatics: An introductory overview. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 1-34). Norwood, NJ: Ablex Publishing Corp.
- Blum-Kulka, S., & Olshtain, E. (1984). Requests and apologies: Cross-Cultural Study of Speech Act Realization Patterns (CCSARP). *Applied Linguistics*, 5(3), 196-213.
- Blum-Kulka, S., & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Studies in Second Language Acquisition*, 8(2), 165-179.
- Bouton, L. F. (1988). A cross-cultural study of ability to interpret implicatures in English. *World Englishes*, 7(2), 183-196.
- Bouton, L. F. (1994). Conversational implicature in the second language: Learned slowly when not deliberately taught. *Journal of Pragmatics*, 22(2), 157-167.
- Brown, J. D. (2001). Pragmatic tests. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 301-325). Cambridge, Cambridge University Press.
- Brown, J. D. (2008). Raters, functions, item types and the dependability of L2 pragmatics tests. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 224-248). Bristol, UK: Multilingual Matters.
- Brown, P., & Levinson, S. C. (1978). Universals of language usage: Politeness phenomena. In E. Goody (Ed.), *Questions and politeness* (pp. 56-324). Cambridge, UK: Cambridge University Press.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.

- Canale, M. (1983). On some dimensions of language proficiency. In J. W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 333-342). Rowley, MA: Newbury House.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Chang, J. (2004). *Examining models of second language knowledge with specific reference to relative clauses: A model-comparison approach*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York, NY.
- Clark, H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 14, 56-72.
- Cook, H. M. (2001). Why can't learners of JFL distinguish polite from impolite speech styles? In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 80-102). Cambridge, UK: Cambridge University Press.
- Dantas-Whitney, M. (2011). Developing students' language awareness. In N. R. Houck & D. H. Tatsuki (Eds.), *Pragmatics: Teaching natural conversation*. New York, NY: TESOL.
- Davis, S. (1991). Introduction. In S. Davis (Ed.), *Pragmatics: A reader* (pp. 3-13). New York, NY: Oxford University Press.
- De Paiva, B. M. M. (2010). Theoretical and methodological approaches in interlanguage pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 261-286). Berlin, Germany: De Gruyter Mouton.
- Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.
- Eisenstein, M., & Bodman, J. (1993). Expressing gratitude in American English. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 64-81). Oxford, UK: Oxford University Press.
- Eslami, Z. R., & Eslami-Rasekh, A. (2008). Enhancing the pragmatic competence of non-native English-speaking teacher candidates (NNESTCs) in an EFL context. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 178-197). Bristol, UK: Multilingual Matters.
- Faerch, C., & Kasper, G. (1989). Internal and external modification in interlanguage request realization. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 221-247). Norwood, NJ: Ablex Publishing Corp.
- Gibbs, R. (1985). Situational conventions and request. In J. Forgas (Ed.), *Language and social situations* (pp. 97-113). New York: Springer Verlag.
- Goffman, E. (1967). On facework: An analysis of ritual elements in social interaction. In E. Goffman, *Interaction ritual: Essays in face-to-face behavior* (pp. 5-46). Chicago: Aldine Publishing Co.
- Grabowski, K. C. (2009). Investigating the construct validity of a test designed to measure grammatical and pragmatic knowledge in the context of speaking (Doctoral dissertation, Teachers College, Columbia University, 2009). *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 70(7), 2487.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377-388.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.

- Harley, B., Allen, P., Cummins, J., & Swain, M. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 7-25). Cambridge, UK: Cambridge University Press.
- House, J. (1989). Politeness in English and German: The functions of *please* and *bitte*. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 96-119). Norwood, NJ: Ablex Publishing Corp.
- House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition*, 18, 225-252.
- Hudson, T. (2001a). Indicators for pragmatic instruction: Some quantitative tools. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 283-300). Cambridge, Cambridge University Press.
- Hudson, T. (2001b). Self-assessment methods in cross-cultural pragmatics. In T. Hudson & J. D. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (pp. 57-74). Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawai'i at Mānoa.
- Hudson, T., Detmer, E., & Brown, J. D. (1992). *A framework for testing cross-cultural pragmatics* (2nd. ed.) (Technical Report #02). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.
- Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics* (Tech. Rep. No. 7). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Hymes, D. (1964). Introduction: Towards ethnographies of communication. *American Anthropologist*, 66(6, part 2), 1-34.
- Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269-293). Harmondsworth, UK: Penguin Books.
- Jaworski, A. (1994). Pragmatic failure in a second language: Greeting responses in English by Polish students.
- Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1, 181-194.
- Kasper, G., & Blum-Kulka, S. (1993). Interlanguage pragmatics: An introduction. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 3-17). Oxford, UK: Oxford University Press.
- Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, 13(2), 215-247.
- Kasper, G., & Rose, K. R. (2001). Pragmatics in language teaching. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 1-9). Cambridge, UK: Cambridge University Press.
- Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Malden, MA: Blackwell Publishing, Ltd.
- Kecskes, I. (2011). Intercultural pragmatics. In D. Archer & P. Grundy (Eds.), *The pragmatics reader* (pp. 372-384). New York, NY: Routledge.
- Keshavarz, M. H., Eslami, Z. R., & Ghahraman, V. (2008). Pragmatic transfer and Iranian EFL refusals: A cross-cultural perspective of Persian and English. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics language learning*, volume 11 (pp. 359-402). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.

- Kondo, S. (2008). Effects on pragmatic development through awareness-raising instruction: Refusals by Japanese EFL learners. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 153-177). Bristol, UK: Multilingual Matters.
- Kwon, J. (2004). Expressing refusals in Korean and in American English. *Multilingua*, 23(4), 339-364. *IRAL: International Review of Applied Linguistics in Language Teaching*, 32(1), 41-55.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London, UK: Longman.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Leech, G. N. (1983). *Principles of pragmatics*. Essex, UK: Longman Group UK Limited.
- Levenston, E. (1975). Aspects of testing the oral proficiency of adult immigrants to Canada. In L. Palmer & B. Spolsky (Eds.), *Papers on language testing 1967-1974*. Washington: TESOL.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge, UK: Cambridge University Press.
- Li, C. (2009). Indirectness scale in Chinese request speech act. *US-China Foreign Language*, 7(10), 51-60.
- Liao, Y. (2009). *A construct validation study of the GEPT reading and listening sections: Re-examining the models of L2 reading and listening abilities and their relations to lexicogrammatical knowledge*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York, NY.
- Lipson, M. (1994). Apologizing in Italian and English. *IRAL: International Review of Applied Linguistics in Language Teaching*, 32(1), 19-39.
- Liu, J. (2006). *Measuring interlanguage pragmatic knowledge of EFL learners*. Frankfurt, Germany: Peter Lang.
- Liu, J. (2007). Developing a pragmatics test for Chinese EFL learners. *Language Testing*, 24(3), 391-415.
- Liu, J. (2010). Testing interlanguage pragmatic knowledge. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 467-488). Berlin, Germany: De Gruyter Mouton.
- Manes, J. (1983). Compliments: A mirror of cultural values. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 96-102). New York, NY: Newbury House.
- Manes, J., & Wolfson, N. (1981). The compliment formula. In F. Coulmas (Ed.), *Conversational routine: Explorations in standardized communication situations and prepatterned speech* (pp. 115-132). The Hague, Netherlands: Mouton.
- Martínez-Flor, A. (2006). The effectiveness of explicit and implicit treatments on EFL learners' confidence in recognizing. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics language learning*, volume 11 (pp. 199-225). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.
- McNamara, T. F. (2000). *Language Testing*. Oxford, UK: Oxford University Press.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Malden, MA: Blackwell Publishing.
- Mey, J. L. (1993). *Pragmatics: An introduction*. Oxford, UK: Blackwell Publishers.

- Morris, C. W. (1938). Foundations of the theory of signs. In O. Neurath, C. Carnap, & C. W. Morris (Eds.), *International encyclopedia of unified science* (pp. 77-138). Chicago, IL: University of Chicago Press.
- Nelson, G. L., Al-Batal, M., & Echols, E. (1996). Arabic and English compliment responses: Potential for pragmatic failure. *Applied Linguistics*, 17(4), 411-432.
- Niezgoda, K., & Röver, C. (2001). Pragmatic and grammatical awareness. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 63-79). Cambridge, UK: Cambridge University Press.
- Norris, J. M. (2001). Use of address terms on the German Speaking Test. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 248-282). Cambridge, Cambridge University Press.
- Oller, J. W., Jr. (1979). *Language tests at school: A pragmatic approach*. London, UK: Longman Group Limited.
- Olshtain, E. (1989). Apologies across languages. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 155-173). Norwood, NJ: Ablex Publishing Corp.
- Olshtain, E., & Cohen, A. (1983). Apology: A speech act set. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and second language acquisition* (pp. 18-35). New York, NY: Newbury House.
- Olshtain, E., & Weinbach, L. (1987). Complaints: A study of speech act behavior among native and non-native speakers of Hebrew. In M. B. Papi & J. Verschueren (Eds.), *The pragmatic perspective: Selected papers from the 1985 International Pragmatics Conference*. Amsterdam, Netherlands: Benjamins.
- Omar, A. S. (2008). Kiswahili requests: Performance of native speakers and learners. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics language learning*, volume 11 (pp. 227-252). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.
- Pellegrino, J.W, Chudowsky, N. & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Phakiti, A. (2008). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language Testing*, 25(2), 237-272.
- Purpura, J. E. (forthcoming). Language learner strategies and styles. In M. Celce-Murcia, D. Brinton, & A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed.). Boston, MA: Heinle Cengage Learning.
- Purpura, J. E. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.
- Purpura, J. E. (in press). Assessing grammar. In A. Kunnan (Ed.), *Companion to language assessment*. Oxford, UK: Wiley-Blackwell.
- Rintell, E. M., & Mitchell, C. J. (1989). Studying requests and apologies: An inquiry into method. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 248-272). Norwood, NJ: Ablex Publishing Corp.
- Roever, C. (2005). *Testing ESL pragmatics: Development and validation of a web-based assessment battery*. Frankfurt, Germany: Peter Lang.
- Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, 23(2), 229-256.

- Roever, C. (2008). Rater, item and candidate effects in discourse completion tests: A FACETS approach. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 249-266). Bristol, UK: Multilingual Matters.
- Roever, C. (2010). Effects of cultural background in a test of ESL pragmalinguistics: A DIF approach. In G. Kasper, H. t. Nguyen, D. R. Yoshimi, & J. K. Yoshioka (Eds.), *Pragmatics language learning*, volume 12 (pp. 187-212). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.
- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4), 463-481.
- Roever, C. (2013). Assessment of pragmatics. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). John Wiley & Sons, Inc. Retrieved from <http://onlinelibrary.wiley.com/book/10.1002/9781405198431/titles?page=2&filter=>
- Rose, K. R. (1992). Speech acts and questionnaires: The effect of hearer response. *Journal of Pragmatics*, 17(1), 49-62.
- Rose, K. R., & Ng Kwai-fun, C. (2001). Inductive and deductive teaching of compliments and compliment responses. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 145-170). Cambridge, UK: Cambridge University Press.
- Schauer, G. A. (2006). The development of ESL learners' pragmatic competence: A longitudinal investigation of awareness and production. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics language learning*, volume 11 (pp. 135-163). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.
- Schegloff, E. A. (1999). Discourse, pragmatics, conversation, analysis. *Discourse Studies*, 1(4), 418-426.
- Searle, J. R. (1965). What is a speech act? In M. Black (Ed.), *Philosophy in America* (pp. 221-239). Ithaca, NY: Cornell University Press.
- Searle, J. R. (1969). *Speech acts*. Cambridge, UK: Cambridge University Press.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 59-82). New York: Academic Press.
- Searle, J. R. (1979). Referential and attributive. *Monist*, 62, 190-208.
- Searle, J. R. (1982). In A. Ortony (Ed.), *Metaphor and thought* (pp. 92-123). Cambridge, UK: Cambridge University Press.
- Sperber, D., & Wilson, D. (1981a). Irony and the use-mention distinction. In P. Cole (Ed.), *Radical pragmatics* (pp. 295-318). New York: Academic Press.
- Sperber, D., & Wilson, D. (1981b). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Stalnaker, R. C. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and philosophy* (pp. 197-214). New York: New York University Press.
- Taguchi, N. (2005). Comprehending implied meaning in English as a foreign language. *The Modern Language Journal*, 89, 543-562.
- Takahashi, S. (2001). The role of input enhancement in developing pragmatic competence. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 171-199). Cambridge, UK: Cambridge University Press.
- Takahashi, S. (2010). Assessing learnability in second language pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 391-421). Berlin, Germany: De Gruyter Mouton.

- Tateyama, Y. (2001). Explicit and implicit teaching of pragmatic routines: Japanese *sumimasen*. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 200-222). Cambridge, UK: Cambridge University Press.
- Tateyama, Y., Kasper, G., Mui, L., Tay, H., & Thananart, O. (1997). Explicit and implicit teaching of pragmatic routines. In L. Bouton (Ed.), *Pragmatics and language learning*, monograph series vol. 8 (pp. 163-178). Urbana-Champaign, IL: Division of English as an International Language, University of Illinois, Urbana-Champaign.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics* 4(2), 91-112.
- Trosborg, A. (2010). Introduction. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 1-39). Berlin, Germany: De Gruyter Mouton.
- Vafaei, P., Basheer, N., & Heitner, R. (2012). Application of confirmatory factor analysis in construct validity investigation: The case of the grammar sub-test of the CEP Placement Exam. *Iranian Journal of Language Testing*, 2(1), 1-19.
- Verschuere, J. (2000). Notes on the role of metapragmatic awareness. *Pragmatics*, 10(4), 439-456.
- Vollmer, H. J., & Olshtain, E. (1989). The language of apologies in German. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 197-218). Norwood, NJ: Ablex Publishing Corp.
- Weizman, E. (1989). Requestive hints. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 71-95). Norwood, NJ: Ablex Publishing Corp.
- Wilson, D., & Sperber, D. (1986). Inference and implicature. In C. Travis (Ed.), *Meaning and interpretation* (pp. 45-75). Oxford, UK: Basil Blackwell.
- Wolfson, N. (1981). Compliments in cross-cultural perspective. *TESOL Quarterly*, 15(2), 117-124.
- Wolfson, N., & Manes, J. (1980). The compliment as a social strategy. *Paper in Linguistics*, 13(3), 389-410.
- Wolfson, N., Marmor, T., & Jones, S. (1989). Problems in the comparison of speech acts across cultures. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 174-196). Norwood, NJ: Ablex Publishing Corp.
- Wouk, F. (2006). Strategies of apologizing in Lombok Indonesia. *Journal of Politeness Research*, 2(2), 277-311.
- Yan, L-M., & Zhuang, Y. (2010). *US-China Foreign Language*, 8(1), 1-7.
- Yamashita, S. (2008). Investigating interlanguage pragmatic ability: What are we testing? In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 201-223). Bristol, UK: Multilingual Matters.
- Yoshimi, D. (2001). Explicit instruction and JFL learners' use of interactional discourse markers. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 223-244). New York: Cambridge University Press.
- Yoshitake-Strain, S. (1997). *Measuring interlanguage pragmatic competence of Japanese students of English as a foreign language: A multi-test framework evaluation*. Unpublished doctoral dissertation, Columbia Pacific University, Novato, CA.
- Yule, G. (1996). *Pragmatics*. Oxford, UK: Oxford University Press.