
Assessing Meaning

James Enos Purpura

Contents

Introduction	2
Early Developments	4
Major Contributions	11
Work-in-Progress	24
Problems and Difficulties	25
Future Directions	26
Related Articles in the Encyclopedia of Language and Education	26
Cross-References	27
References	27

Abstract

The quintessential quality of communicative success is the ability to effectively express, understand, dynamically co-construct, negotiate and repair variegated meanings in a wide range of language use contexts. It stands to reason then that meaning and meaning conveyance should play a central role in L2 assessment. Instead, since the 1980s, language testers have focused almost exclusively on functional proficiency (the conveyance of functional meaning – e.g., can-do statements), to the exclusion of the conveyance of propositional meanings or implied pragmatic meanings. While the ability to use language to get things done is important, excluding propositional content from the assessment process is like having language ability with nothing to say, and excluding pragmatic meanings guts the heart and soul out of communication.

In this chapter, I review how L2 testers have conceptualized “meaning” in models of L2 proficiency throughout the years. This logically leads to a discussion of the use of language to encode a range of meanings, deriving not only from

J.E. Purpura (✉)

Teachers College, Columbia University, New York, NY, USA

e-mail: jp248@tc.columbia.edu

an examinee's topical knowledge but also from an understanding of the contextual factors in language use situations. Throughout the discussion, I also highlight how the expression and comprehension of meaning have been operationalized in L2 assessments. Finally, I argue that despite the complexities of defining and operationalizing meaning in assessments, testers need to seriously think about what meanings they want to test and what meanings they are already assessing *implicitly*.

Keywords

Assessing meaning • Assessing pragmatics • Assessing content

Introduction

Nonnative speakers use second or foreign languages (L2) to get/give information at school, to create and maintain relationships online, to get a glimpse into other cultures, or, more subtly, to decipher intentions in political discourse. In other words, they use their L2 to express a wide range of meanings within social-interpersonal contexts (e.g., a friend recounting a subway story), social-transactional contexts (e.g., a client resolving a problem with a bill), academic contexts (e.g., a student writing a term paper), professional contexts (e.g., a scientist giving a talk), and literary or imaginative contexts (e.g., a poet writing/reciting a poem at a poetry slam). Since the ability to effectively express, understand, co-construct, negotiate, and repair meanings is the quintessential quality of communicative success, it stands to reason then that meaning and meaning conveyance should play a central role in L2 assessment (Purpura 2004).

In the L2 use domains mentioned above, language serves to generate messages that embody a variety of simultaneously occurring meanings. First and foremost, messages contained in utterances or texts encode *propositional* or *topical content*. Thus, the *propositional* or *topical meaning* of utterances or texts is said to convey factual information, ideas, events, beliefs, conjectures, desires, and feelings and is presumed to be context-free or decipherable apart from a communicative situation (Gibbs 1994). These propositional utterances are open to scrutiny in terms of their factual accuracy or their true-value¹. Propositional meanings in the literature have also been referred to as the literal, semantic, sentential, compositional, grammatical, linguistic, inherent, conventional, or locutionary meaning of utterances and are generally considered a reflection of an individual's substantive, topical, or disciplinary, domain specific, subject matter, or content knowledge. They are fundamental to all language use.

The expression of propositions in messages is also used to assert a person's agency and express his intentionality in communicative interactions (e.g., to persuade) (Bloom and Tinkler 2001). By encoding intended meanings, these messages are used by interlocutors to *perform speech acts* or *communicative functions* with reference to some language use context. Thus, messages in utterances or texts *also* encode a user's *intended* or *functional meanings*. We can say then that the

¹See Donald Davidson's essays for a fascinating discussion of truth and meaning.

propositional content of a message conveys more than what is said with words; it also communicates intended or functional meaning relevant to a language use context. Intended or functional meanings have been referred to in the literature as conveyed, interactional, illocutionary, or speaker's meaning. Unlike propositional meanings, functional meanings depend on the context of language use for successful interpretation. Similar to propositional meanings, however, functional meanings are fundamental to all language use as they represent an individual's *functional proficiency*.

Finally, while messages emerge from, depend on, and embody representations of an individual's internal mental content and serve as a reflection of personal agency and intentionality, they do not occur in isolation; they exist within a given sociocultural and interactional context and are thereby shaped by and interpreted within that context. Given that communication depends on the participants' shared presuppositions, experiences, and situational associations, much of what occurs in language use is unstated or implied. As a result, these same messages embody yet other layers of meaning, referred to as *implied* or *implicational pragmatic meanings*.

Implied pragmatic meanings emerge, for example, when someone is offered red wine and the acceptance response is: *Hey, I'm Italian*. Explicit in this response is the expression of propositional content – nationality. However, the response is also used *in this context* to communicate the respondent's functional meaning (i.e., my interlocutor made an offer; I'm *accepting*). Conjointly with the propositional and functional meanings, the response subtly encodes layers of other implied meanings including (1) *situational meanings* (i.e., the response reminds my interlocutor of my ethnicity and the role of red wine in my culture and presupposes my interlocutor will interpret my indirect response as an acceptance *in this situation*, even though not explicitly stated), (2) *sociolinguistic meanings* (i.e., the response conveys familiarity), (3) *sociocultural meanings* (i.e., the response presupposes what is common knowledge about Italians in *our culture*), and (4) *psychological meanings* (i.e., the response conveys playfulness). These implied pragmatic meanings have been referred to as socio-pragmatic, figurative, extralinguistic, or implicational meanings.

Implied pragmatic meanings can also emerge as a simple function of word order. Consider the propositional, sociocultural, and psychological meanings associated with the utterance “My niece got married and had a baby” as opposed to “My niece had a baby and got married.” Consider also how these meanings might vary across different social contexts.

In sum, language is efficiently designed to convey propositional meanings through topical content, together with functional meanings and layers of implied pragmatic meaning relevant to some language use context (Purpura 2016). The interaction among topical knowledge, language knowledge, and context and, I would add, the sociocognitive features of task engagement enable nuanced communication. And while these simultaneous encodings of meaning joyfully provide the basis for humor or poetry, they also increase the risk of communication breakdowns or the miscommunication of intent. They also present L2 learners with daunting challenges and heartwarming joys of learning to use an L2.

In L2 assessment, especially with nonreciprocal tasks, the propositional messages conveyed by an interlocutor, along with other pragmatic meanings, might be

considered a manifestation of a person's topical and language knowledge, her understanding of context, and her sociocognitive abilities. This is especially true if the propositions are true and faithful representations of the external world, if communication goals are met, and if the language output is grammatically precise and appropriate for the situation. With reciprocal tasks, however, these same messages serve only to initiate the establishment of joint understandings, followed by the co-construction of meanings relevant to the context. Communicative success thus is a *joint* product of the co-construction of variegated meanings. Finally, for communication to be successful, interlocutors need to express their own representations of mental content, reconstruct mental content representations of their interlocutors, and jointly co-construct these meaning representations synchronically and diachronically in verbal or nonverbal behavior. As Bates (1976 cited in Seliger 1985) stated:

Meaning is a set of mental operations carried out by the speaker, which the speaker intends to create in the mind of the listener by using a given sentence. Whether or not the speaker actually succeeds is a separate issue. (p. 4)

Although the communication of meaning through propositional content and context plays a central role in L2 communicative success, L2 testers have devoted surprisingly little empirical attention to this topic. Instead, they continue to produce assessments, which, in my opinion, over-attribute value to the well-formedness of messages and to the completion of the functional acts, and they under-attribute importance to the conveyance of substantive, relevant, or original content, the development of topical progressions, and the conveyance of implied pragmatic meanings. This, by no means, is meant to diminish the significance of linguistic well-formedness in contexts where communicative precision is needed, or the importance of ascertaining L2 functional ability; it is simply a reminder that the primary aim of communication is the exchange of meanings in context. Thus, language, meaning through content, contextual considerations, and the socio-cognitive considerations of task engagement should figure prominently in the design and validation of all L2 assessments.

In this chapter, I will review how testers have conceptualized “meaning” in models of L2 proficiency, describing the role that meaning conveyance through content and context has played in L2 assessment. I will argue for a reprioritization of meaningfulness over well-formedness in L2 test design since the exclusion of meaning from models of L2 ability likens to having language ability with nothing to say. Finally, I will highlight some of the problems and challenges in assessing meaning.

Early Developments

Although some early language testers have purposefully disregarded the importance of meaning in models of L2 proficiency, others have clearly acknowledged the critical role it plays in communication and have addressed meaning and meaning conveyance in characterizations of L2 proficiency. This reflects the fact that people

use language in systematic ways to exchange messages on a variety of topics in a wide range of contexts, and in that way, they use language to get things done.

In 1961 Lado proposed a model of L2 proficiency based on a conceptualization of “language” as linguistic forms, occurring in some variational distribution, that are needed to convey linguistic, cultural, and individual meanings between individuals. *Linguistic meanings* referred to the denotative or the *semantic meaning* of “dictionaries and grammars” and were “interpretable without recourse to full cultural reference” (p. 3). Currently, linguistic meanings would be referred to as the literal, semantic, or propositional meaning of a form, utterance, or text. Linguistic meanings were said to reside in the use of phonology, sentence structure, and the lexicon and context limited to that contained within a sentence. *Cultural meanings* referred to concepts or notions that are culturally bound and only interpretable within a specific speech community or culture (e.g., *tapas*, *English breakfast*). Currently, these would be referred to as pragmatic meanings. Finally, *individual meanings* for Lado referenced words or concepts that lay outside the culture per se, indexing personal associations, such as when the word *dog* carries positive or negative connotations based on an individual’s past experiences. With respect to these meanings, Lado argued that language is based initially on the linguistic meanings of structures and their combinations in an utterance, followed by other contextually bound meanings (p. 6). His schematization appears in Fig. 1.

Despite Lado’s visionary depiction of “language” as a system of meaningful communication among individuals, he prioritized discrete linguistic elements (phonology, syntax, lexicon) of language use (reading, listening, speaking, writing) when it came to assessment design. He thus organized assessments around discrete forms, rather than around rich communicative situations in which layers of meanings could be elicited and measured. Consequently, assessing the extent to which messages are

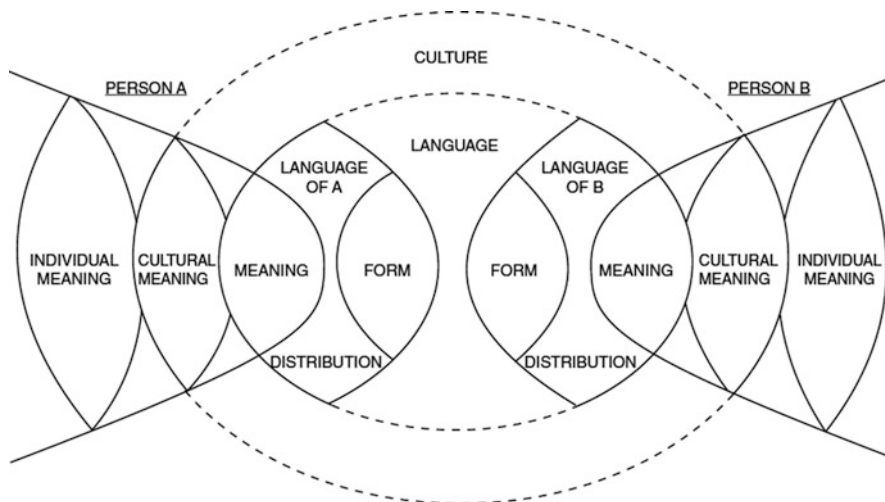


Fig. 1 Language, culture, and the individual (Lado 1961, p. 6)

encoded semantically and communicated socially was secondary to the measurement of the linguistic resources used to deliver these messages. This perspective resulted in discrete-point assessments of forms and their associated meanings, instead of assessments eliciting propositionally accurate, topically elaborated, and situationally appropriate responses.

For example, in assessing phonological awareness through lexis, Lado presented students with two pictures, each representing a lexical item chosen because they happened to be minimal pairs (e.g., ship/sheep). Students then heard a word and selected the correct answer. In testing the meaning of counterfactual *if*-clauses, he presented examinees with a sentence and asked them to infer the correct propositional meaning of the sentence, based solely on the linguistic context, as seen below:

If the windows were closed, I would ask you to open them.

- A) The windows are closed.
- B) The student goes to the windows and opens them.
- C) The student remains seated. (p. 158)

Given the minimal context, the inferencing needed to relate option (C) to the stem seems greater than the inferencing needed to understand the stem.

When it came to assessment, Lado generally preferred restricting test input to information that was “common knowledge in the culture where the language was spoken” (p. 205) and restricting the questions to selected-response items. This was based on the conviction that such restrictions would reduce the risk of introducing extraneous factors into the measurement process through situational context. However, when it came to extended production tasks, he argued that extraneous factors could be controlled to some extent by the use of rating scales revolving around linguistic difficulties and the success of meaning conveyance. Interestingly, the following language and meaning-based descriptors were used to rate the ability to narrate a story based on a picture:

2 – Conveys a simple description completely and correctly.

Conveys the simple description completely and correctly, but elaborates, and in so doing, makes some error, or error of vocabulary, grammar, or pronunciation – errors which interfere little with the understandability of the utterance. (Lado 1961, p. 240)

0 – Conveys very little meaning.

Conveys the wrong meaning.

Makes errors, which obscure the meaning.

Says nothing. (Lado 1961, p. 241)

While Lado is best known for the measurement of linguistic forms with discrete-point tasks, his conceptualization of L2 proficiency is clearly broader than that. From the onset, he recognized the importance of meaning in communication and provided recommendations for its measurement, not only in selected-response tasks, where meanings associated with grammatical forms and sentential propositions were assessed, but also in extended production tasks, where consideration was given to the overall conveyance of meaning in responses and to the extent that grammatical

inaccuracy detracted from meaning conveyance. While much remained unspecified in Lado's model regarding the types of meanings that were assessed or even the role of topical knowledge in enabling meaning conveyance, his ideas on this topic were insightful and, in my opinion, should have had a greater impact.

Carroll (1961/1972, 1968) also highlighted the role that meaning plays in language assessment. He defined "language" as:

A system of 'rules' for generating utterances (or written representations thereof) that will be accepted by members of a given speech community as 'correct or 'grammatical' and understood by them as having a possible semantic interpretation. (Carroll 1968, p. 47)

Like Lado, Carroll recommended that L2 knowledge be specified in terms of linguistic forms, complemented by a semantic component. Unlike Lado, however, he recommended that less attention be paid to discrete morphosyntactic and lexical forms than to the "total effect of an utterance" or the "total meaning of the sentence" (p. 37). As a result, he proposed that that measurement of discrete components of L2 knowledge be supplemented by performance tasks that require the integration of components through connected discourse. Unfortunately, Carroll's inclusion of a meaning component in assessment was inconsistently applied in construct definition and operationalization.

Oller (1979) significantly advanced the conversation on "meaning" by describing "language" as both the interpretation and conveyance of factual content and the transmission of emotive or affective meanings in language use. He maintained "language is usually used to convey information about people, things, events, ideas, states of affairs, and [emphasis in the original] attitudes toward all the foregoing" (p. 17). He referred to the literal propositional meanings as the *factive information* of language use expressed by "words, phrases, clauses and discourse" (p. 33) and the psychological meanings of language use as *emotive* or *affective* information often carried by phonology or gestures. The emotive features were seen to "convey attitudes toward the asserted or implied state of affairs, [which] may further code information concerning the way the speaker thinks the listener should feel about those states of affairs" (p. 18). Furthermore, according to Oller, the factive and emotive information of communication was highly dependent on the context of language use, which he referred to as (1) the *linguistic context*, consisting of the *verbal* and *gestural contexts* of language use, and (2) the *extralinguistic context*, involving the subjective and objective realities of "things, events, persons, ideas, relationships, feelings, perceptions memories, and so forth" (p. 19). He then asserted, "linguistic contexts are pragmatically mapped onto extralinguistic contexts, and vice versa" (p. 19). In other words, the meaning of the combined linguistic forms of an utterance (i.e., literal propositional meaning) is shaped by the pragmatic context in which the utterance or text is expressed. Also, the pragmatic mappings are a result of relating the propositional meanings of linguistic forms to extralinguistic context or human experience. In sum, Oller's farsighted conceptualization of L2 proficiency took into account linguistic knowledge, factual or topical knowledge, pragmatic

knowledge (including emotive), and contextual features (extralinguistic) – a view strikingly similar to some current conceptualizations of L2 proficiency.

The following meaning recognition item illustrates how Oller attempted to measure the examinee’s ability to decipher the meaning of the word *dropped* when the literal meaning was extended to suit the context:

John dropped the letter in the mailbox.

- A) John sent the letter.
- B) John opened the letter.
- C) John lost the letter.
- D) John destroyed the letter. (Oller 1979, p. 46)

In other words, “*dropping* a letter in a mailbox” is assumed to mean *sent*, not *let fall*, based on information in the context. Meaning extension here again derives from the available distractors, as an option such as “*put* the letter in the mailbox” would have been closer in meaning to the stem. Oller attempted to do the same in the following inferencing item.

Man’s voice: Hello Mary. This is Mr. Smith at the office. Is Bill feeling any better today?

Woman’s voice: Oh yes, Mr. Smith. He’s feeling much better now. But the doctor says he’ll have to stay in bed until Monday.

Third voice: Where is Bill now?

- A) At the office.
- B) On his way to work.
- C) Home in bed.
- D) Away on vacation. (Oller 1979, p. 47)

Response (C) was also designed to measure the ability to decode meaning by mapping it onto an extralinguistic context (i.e., implied pragmatic meanings) as Bill’s location cannot be derived *solely* from the linguistic context of the input, but from the presupposition that Bill’s bed is in his home (i.e., he *could* have a bed in his office). Nonetheless, this item could have been a clearer example of meaning extension had distractor (C) been worded *At home*.

This approach to assessment supported Oller’s proposal to use “integrative” or “pragmatic” tests to measure a learner’s “pragmatic expectancy grammar,” defined as a “psychologically real system that sequentially orders linguistic elements in time and in relation to the extralinguistic contexts in meaningful ways” (p. 34). An examinee would then display knowledge of pragmatic expectancy grammar by “relating these sequences of linguistic elements via pragmatic mappings to extralinguistic context” (p. 38). Importantly, pragmatic expectancy grammar aimed to connect the grammatical forms of an utterance, and the meaning expressed by this utterance in context, to some extralinguistic reality by inferential (i.e., cognitive) processes, thereby linking the utterance, I believe, to the individual’s prior experience, knowledge, agency, and intentionality. Oller’s position demonstrates a strong rejection of the then-current Bloomfieldian (1933) approach to linguistic analysis

and formalism, reified by Chomsky (1957), where meaning was completely disregarded from linguistic analysis,² in favor of a communication-based approach to language use.

In terms of measurement, Oller also recommended scoring protocols that specified not just “how well the text conforms to discrete points of morphology and syntax, but how well it expresses the author’s intended meaning” (p. 386) in a given context, since judges always consider the communicative effectiveness of responses, whether or not they are scored.

Oller can be credited for highlighting not only the literal propositional (factual) and psychological (emotive/affective) content of utterances, encoded by linguistic forms, but also how these utterances relate to both internal mental states (i.e., cognition) and extralinguistic context. His work is also credited for specifying scoring methods that operationalize the assessment of propositional meaning conveyance in a variety of task types. Unfortunately, Oller never provided detailed theoretical or operational definitions of factive and emotive meanings conveyed in language use so that the quality of the factive information or the appropriateness of the emotive information in responses could be systematically assessed. Nor did he specify how test design could systematically account for extralinguistic context or the cognitive components of L2 proficiency in response elicitation. Nonetheless, Oller’s insightful and forward-thinking ideas on meaning foreshadowed later conceptualizations of L2 proficiency.

Other testers have also highlighted the importance of meaning in language assessment. Inspired by Hymes (1967, 1972), Savignon (1972), Halliday (1973), Van Ek (1976), and Munby (1978), among others, Canale and Swain (1980) argued that language competence should be conceptualized within a framework of communication, where the functional meaning of utterances is central to L2 proficiency. In other words, priority was placed more on an individual’s ability to achieve a communicative goal – to convey intended or functional meanings in context, than on the capacity to communicate accurate or relevant propositional content within the function. Secondary priority was given to an individual’s ability to communicate with grammatical accuracy in ways that are socioculturally appropriate.

Canale and Swain’s model conceptualized communicative competence as “a synthesis of knowledge of basic grammatical principles, knowledge of how the language is used in social contexts to perform communicative functions, and knowledge of how utterances and communicative functions can be combined according to the principles of discourse” (p. 20). They defined this construct in terms of grammatical, sociolinguistic, and strategic competence – later discourse competence was added (Canale 1983). While not the primary focus, the importance of meaning was noted in many parts of the model. For example, grammatical competence was

²Surprisingly, the commitment to a syntactocentric approach to assessment, where only features of the language are assessed for accuracy, complexity, range, and fluency, has persisted in many assessments. As a result, the effective communication of propositions and the communicative meanings associated with these propositions are often ignored in the measurement process.

defined in term of rules of semantics associated with “word meaning and sentence meaning” or the notion of “getting one’s point across” (p. 10) (i.e., propositional meaning), and sociolinguistic competence was described as the sociocultural rules of language use and the rules of discourse (i.e., pragmatic meaning) (see Halliday and Hassan 1976; van Dijk 1977; Widdowson 1978). Canale and Swain further argued that learners need to know both sets of rules in order to appropriately express and understand meanings, especially when there is a “low level of transparency between the literal meaning of an utterance and the speakers’ intended meaning” (p. 30) – in other words, in situations where the propositional content of utterances along with the communicative intents can be derived only from situational factors. Canale and Swain explained that the sociocultural rules of language use made possible the expression and interpretation of appropriate attitudes and registers within sociocultural contexts and that the discourse rules³ allowed for the expression and interpretation of cohesion and coherence. Cohesive rules related forms to different types of referential meanings in texts,⁴ while coherence rules related propositions and their communicative functions in sequenced discourse to implied rhetorical meanings in text. To exemplify, consider the implied rhetorical meanings created in following discourse sequence.

Dialogue	Functional (and propositional) meanings	Implied rhetorical meanings (coherence)
A) That’s the telephone	Device identification (<i>The phone is ringing</i>)	Implied request (<i>Can you answer the phone?</i>)
B) I’m in the bath	Expression of location (<i>I’m in the tub, presumably taking a bath</i>)	Implied refusal (<i>I’m taking a bath so I can’t answer the phone</i>)
A) OK (Data from Widdowson 1978, p. 29)	Acknowledgment (<i>I acknowledge you are in the tub taking a bath</i>)	Implied acceptance of refusal (<i>I acknowledge you can’t answer the phone</i>); implied response to request (<i>I’ll answer it</i>)

Canale and Swain’s widely accepted model significantly broadened our understanding of the individual components of communicative competence and helped further the shift in assessment from a focus on grammatical forms to an emphasis on functional meanings in social interaction. It also highlighted, at least theoretically, the need to consider the sociolinguistic meanings carried by utterances, where an assessment might measure sociocultural appropriateness. Finally, it underscored the need to account for the rhetorical meanings encoded in cohesion and coherence. Although this model downplayed the role of topical knowledge and context in

³Canale (1983) later recognized that the rules of discourse might better be separated from the sociocultural rules of language use. Thus, he broadened the original conceptualization of communicative competence to include grammatical, sociolinguistic, and discourse competence and the cognitive component of language use, strategic competence.

⁴For example, anaphoric reference to relate the pronoun, *him*, to a referent, *boy*, or the logical connector, *then*, to relate temporality between clauses.

functional proficiency, it still inspired other testers to refine later notions of communicative competence as a basis for assessment.

Major Contributions

Influenced by Canale and Swain (1980) and many others, Bachman (1990) proposed a model of communicative language ability framed within the notion of language use. This model was later refined in Bachman and Palmer (1996, 2010). In this model, meaning played a prominent role. Bachman and Palmer (2010) defined “language use”:

... as the creation and interpretation of intended meanings in discourse by an individual, or as the dynamic and interactive negotiation of meaning between two or more individuals in a particular situation. In using language to express, interpret, or negotiate intended meanings, language users create discourse. This discourse derives meaning not only from utterances or texts themselves, but, more importantly, from the ways in which utterances and texts relate to the characteristics of a particular language use situation. (p. 14)

While Canale and Swain limited their discussion to a language user’s “communicative competence,” defined in terms of language knowledge components and strategic competence, Bachman and Palmer (2010) significantly broadened the construct by arguing that in addition to *language knowledge*, language users in the act of communication need to engage their topical knowledge, affective schemata, and strategic competence when presented with some real-life or assessment task. They further argued that it is the interaction between an individual’s language knowledge and these other factors that enable the user to create and understand meanings through discourse. While Bachman and Palmer never really provided an explicit definition of “meaning” in their model, they engaged in a compelling discussion of the knowledge components underpinning the creation and comprehension of meanings in discourse.

Bachman and Palmer’s (2010) comprehensive description of language use consisted of language knowledge, topical knowledge, affective schemata, strategic competence, and other personal attributes; however, I will limit this discussion to an examination of language and topical knowledge given their role in the communication of meaning.

Bachman and Palmer conceptualized *language knowledge* as the interactions between organizational and pragmatic knowledge. *Organizational knowledge* was defined as (1) the knowledge that users need to produce or interpret spoken and written utterances to construct meaning – i.e., *grammatical knowledge*, or knowledge of vocabulary, syntax, phonology, graphology, and (2) the knowledge they need to organize these utterances into coherent spoken or written texts – i.e., *textual knowledge*, or knowledge of cohesion and rhetorical/conversational organization. Although they did not explicitly frame organizational knowledge in terms of forms and their associated meanings, they alluded to these two dimensions in discussing

scoring. For example, when a Spanish learner says **hers dogs* instead of *her dogs*, the incorrect utterance reveals his knowledge of *cohesive meaning* (correct reference to a female) and lack of knowledge of *cohesive form* (possessive adjectives do not agree with nouns in number in English). Therefore, in these cases, they recommended assigning one point to meaning and zero to form.

The second component of language knowledge in this model, *pragmatic knowledge*, was defined as the mental representations needed to “enable users to create or interpret discourse by relating utterances or sentences and texts to their meanings, to the intentions of language users, and to the relevant characteristics of the language use setting” (Bachman and Palmer 2010, p. 46). Pragmatic knowledge was further defined in terms of *functional* and *sociolinguistic knowledge*. Both components deal with meaning on some level.

Functional knowledge was said to “enable us to [express and] interpret relationships between utterances or sentences and texts and the intentions of language users” (Bachman and Palmer 2010, p. 46) in order to accomplish some communicative goal in context. Interestingly, this definition characterizes functional knowledge as a feature of the co-construction of communicative goal between two or more individuals, rather than as an attribute of a single user’s communicative intentionality. As a result, a learner might be seen as demonstrating evidence of functional knowledge by responding to a friend’s question *Can I give you some more wine*, with *Sure, pour away*, instead of *Yes, you are strong enough to lift the bottle*. In this context, the learner interpreted her friend’s question as an indirect offer, rather than a query about ability, thereby achieving communicative success. If she had responded with a confirmation of ability, rather than an acceptance, this might have confused the interlocutor, resulting in communicative failure (i.e., lack of functional knowledge), unless, of course, she was intentionally being sarcastic. Thus, functional knowledge enables users to utilize context, as minimal as it is, to reassign meaning from a literal proposition (*can* = ability) to an intended meaning (*can* = request), or even to an implicated meaning (*can* = sarcasm) based on the communicative function of the utterance in discourse. Functional knowledge is thus seen as enabling users to *get things done through language* (van Dijk 1977).

Drawing on Halliday (1973) and Halliday and Hasan (1976), Bachman and Palmer identified four categories of functional knowledge that permit users to communicate joint intentions: *knowledge of ideational functions* (i.e., use of functions to relate ideas related of the real world – informing), *knowledge of manipulative functions* (i.e., use of functions to impact the world around us – requesting), *knowledge of heuristic functions* (i.e., use of functions to extend their knowledge of the world – problem-solving), and *knowledge of imaginative functions* (i.e., use of functions related to imagination or aesthetics – joking). In each case, a user would be judged on her ability to perform these functions.

Functional knowledge thus embodies the mental structures needed to communicate contextually relevant intentions between users with respect to the four communicative goals. It also enables users to get things done through language, thus explaining its operationalization in assessments as *can-do* statements. What remains unclear is the role that propositional content plays in expressing the four functions. It

would not be hard to imagine a situation in which learners can use the L2 accurately (grammatical knowledge) to summarize a story (functional knowledge), but the information in the summary (propositional knowledge through content) is inaccurate. In other words, it seems possible to demonstrate functional knowledge without displaying topical knowledge. Also unclear is the role that context plays in the expression or interpretation of functional knowledge. For instance, the interpretation of an indirect request (manipulative function) or a joke (imaginative function) depends on context for meaning conveyance, in addition to topic. So, given that meaning in these instances is derivable primarily, and sometimes uniquely, from features of context, is it possible to communicate functional knowledge without accurate or relevant topical content related to these contextual features? I would argue then that assessments based solely on functional proficiency provide only a partial estimate of a person's proficiency and one that can result in miscommunication.

Bachman and Palmer then defined *sociolinguistic knowledge* as the mental structures required to “enable us to create and understand language appropriate to a particular language use settings” (p. 47). Sociolinguistic knowledge targeted the user's capacity to use genres, dialects/varieties, registers, natural or idiomatic expressions, and cultural references or figures of speech appropriately in context. Thus, users able to use register appropriately and flexibly in formal contexts would be scored high for appropriate and wide knowledge of registers. The sociolinguistic component emphasized a user's “sensitivity” to register variations, natural or conventional expressions, and other linguistic features with relation to their appropriate use in context. Of note, however, is that this component is framed in terms of user sensitivity to these features, rather than in terms of the user's ability to recognize and transmit these implicit meanings in context.

Implicit in Bachman and Palmer's notion of sociolinguistic knowledge is first the inherent potential that users have for extending meaning beyond what is literally indexed in discourse. For example, the ability to use the expression *Your wish is my command* appropriately in context extends beyond an understanding of the literal propositional or functional meanings of the expression; it also presupposes an understanding of the context of language use as it relates to the transmission of sociocultural meaning (genie in a bottle) and sociolinguistic meaning – power (unequal), imposition (no limits), and distance (near). Although Bachman and Palmer did not frame sociolinguistic knowledge in terms of meaning, their model clearly highlighted the importance of sociolinguistic knowledge as a feature of language use and provided a basis for further work on the assessment of pragmatic ability.

One of the most interesting features of Bachman and Palmer's (2010) work in terms of meaning, however, was their discussion of topical knowledge as a consideration in assessment design and operationalization. While previous researchers vaguely referred to propositional content encoded in messages, Bachman and Palmer provided a compelling discussion of what topical knowledge refers to, how it interacts with other features of language use, and how it might be assessed.

They defined *topical knowledge* (also referred to in the literature as content knowledge, knowledge schemata, real-world knowledge, overall literal semantic

meaning, propositional content, or background knowledge) as knowledge structures in long-term memory (LTM) – unfortunately without further specification. They argued that topical knowledge is critical to language use because it provides the information needed to use language with reference to the real world and, I would add, with reference to an individual’s internal world, as in creative expression. They stated that while topical knowledge is separate from language ability, it is still “involved in all language use” (p. 41) and is a factor in all test performance. They also maintained that since “it may not be possible to completely isolate language ability from topical knowledge in some test tasks, no matter how we score test takers’ responses” (p. 325), testers should consider topical knowledge in assessment. Finally, they added that when an individual’s topical knowledge interacts with the topical content in task completion, it impacts difficulty.

To disentangle the relationship between language ability and topical knowledge in test design, Bachman and Palmer offered three specification options:

1. Define the construct solely in terms of language ability.
2. Define language ability and topical knowledge as a single construct.
3. Define language ability and topical knowledge as separate constructs (p. 217).

Option 1 refers to assessment contexts making claims only about a component of L2 ability – e.g., knowledge of form. This might involve tasks focusing only on the measurement of form (with the topical meaning dimension being controlled) – e.g., when examinees are asked to choose among allophones (/t/, /d/, /id/) or among different verb forms (enjoy + *work*, *works*, *working*). In these cases, most testers would argue that topical knowledge is not part of the construct; thus, only one component of L2 knowledge (i.e., knowledge of form) would be scored. I would argue, however, that topical knowledge, in the form of metalinguistic knowledge, would be engaged – even if it is implicit knowledge. Option 2 refers to contexts making claims about L2 ability and topical knowledge as part of the same construct – e.g., when an international teaching assistant, *assumed to have the required topical knowledge for task completion*, must give a presentation in the L2. Only one score is taken and interpreted as the ability to use L2 and topical knowledge to teach. This option confounds language ability and topical knowledge, as scores could be affected by deficiencies in either. Finally, option 3 refers to contexts making claims about both L2 ability and topical knowledge as different constructs – e.g., in language for specific purposes (LSP) contexts, where examinees need to display their ability to use L2 ability to communicate disciplinary content – e.g., an analysis of food chains in an ecosystem. In this case, topical knowledge is conceptualized as drawing on explicit declarative memory or, I might add, a network of facts, concepts, principles, and rules in semantic memory that are assumed to be separate from language ability. Bachman and Palmer provided an example of a rubric [only partially presented] designed to measure topical knowledge in this context.

Levels of knowledge/ mastery	Description
4 complete	Evidence of: <i>complete knowledge of relevant topical information</i> <i>Range:</i> evidence of unlimited range of relevant topical information <i>Accuracy:</i> evidence of complete accuracy throughout range
2 moderate	Evidence of: <i>moderate knowledge of relevant topical information</i> <i>Range:</i> medium <i>Accuracy:</i> moderate to good accuracy within range (Bachman and Palmer 2010, p. 352)

While Bachman and Palmer’s model greatly advanced our understanding of topical knowledge in L2 assessment, several questions remain. The first relates to the composition of knowledge structures related to topical knowledge in LTM. Are these knowledge structures limited to a semantic memory for factual or disciplinary knowledge (Dehn 2008), or might these assessments require examinees to draw on other memory sources in task completion? A second question concerns the relationship between topical and L2 knowledge. If topical knowledge is needed to generate and understand propositions encoded in language, is it actually possible to communicate without topical knowledge? And is it ever possible to assess L2 knowledge without some form of topical knowledge? Similarly, is it possible to have pragmatic ability without knowledge of the contextual situation (episodic memory) (Dehn 2008)? Finally, if communicative language ability *always* includes topical knowledge on some level, along with L2 knowledge, contextual understandings, and cognitive processing factors, then shouldn’t these four features *always* be specified in assessment tasks involving communication? After all, each can potentially moderate L2 performance. In Fig. 2 I have attempted to schematize design considerations relating to context, topical content, language, and cognition/disposition as potential moderators of L2 proficiency in task engagement.

Building on Bachman and Palmer (1996) and Chapelle’s (1998) interactionist approach to construct definition, Douglas (2000) reexamined the role of meaning by problematizing the relationship between background knowledge and L2 ability in the context of LSP assessment. He argued that in LSP contexts, an examinee’s background knowledge was, in addition to L2 knowledge and strategic competence, a critical contributor to *specific purpose language ability* (SPLA). As a result, he defined background knowledge as part of the SPLA construct. Douglas defined *background knowledge* as “frames of reference based on past experience” (p. 35) within a discourse domain – a conceptualization reminiscent of what Baddeley et al. (2009) refer to as semantic declarative memory (i.e., factual knowledge) and episodic memory (i.e., experiential knowledge) associated with past contexts, events, and episodes related to LSP contexts. This insight, in my view, extends to all L2 assessments, as prior topical knowledge *on some level* is fundamental to meaning making. What, I believe, will fluctuate in L2 use (and in assessments) is the *type* of

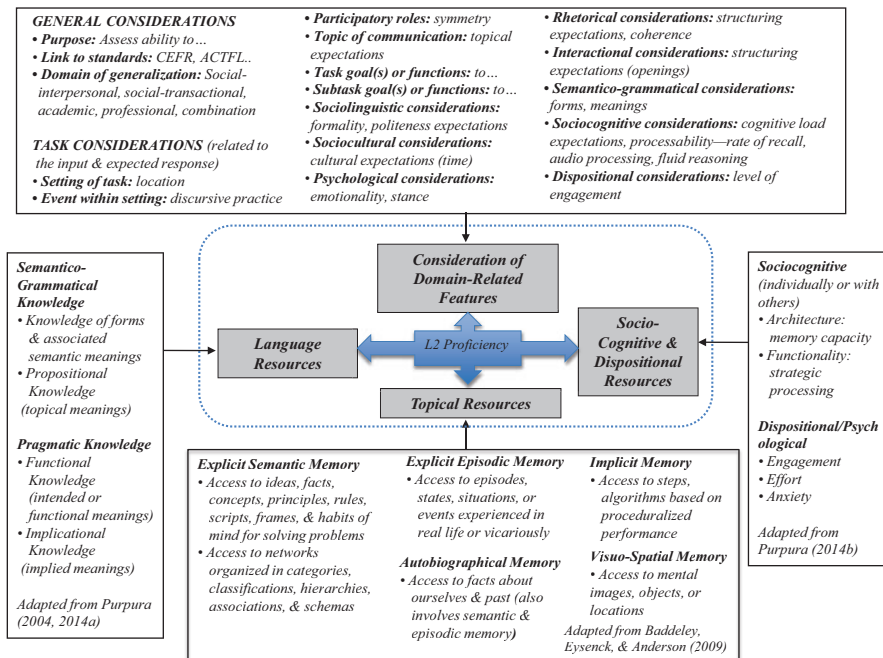


Fig. 2 Representations of context, content, language, and cognition as potential moderators of L2 proficiency in task engagement

topical knowledge needed to communicate (e.g., disciplinary knowledge, autobiographical knowledge, situational knowledge), but certainly communication is impossible with no topical knowledge or flawed topical knowledge.

Finally, Douglas noted that even when LSP assessments are successful in engaging SPLA, performance on these tests is far too often based on scoring criteria revolving around the grammatical features of the output (e.g., intelligibility, fluency, grammatical accuracy, lexical appropriateness), rather than on “aspects of communicative language ability [...] deemed to be important” (p. 279) for inferences about SPLA. In other words, assessment criteria failed to target the examinees’ ability to perform functions in LSP contexts that measure L2 ability in conjunction with critical aspects of disciplinary knowledge.

Building on prior research, Purpura (2004, 2014a, 2016) offered a slightly different conceptualization of L2 proficiency in which the ability to communicate meanings in some domain of L2 use depends upon the interaction between the context of language use, language knowledge, topical knowledge, and the socio-cognitive and dispositional resources of task engagement, as seen in Fig. 2. In this conceptualization, meaning and meaning conveyance are seen as the cornerstone of L2 proficiency. This depiction of L2 proficiency is based on the assumption that proficiency, sampled as the simple utterance of a sentence with beginners to highly nuanced communication, requires a network of resources that enable users to

express, understand, dynamically co-construct, negotiate, and repair meanings, knowledge, and action, often in goal-oriented interaction. It also acknowledges the risks associated with meaning-related conversational breakdowns or flat-out miscommunications, due not only to semantico-grammatical deficiencies but also and more insidiously to pragmatic infelicities, which can easily lead to mutual misjudgment of intentions and abilities, miscommunication, and even cultural stereotyping (Gumperz 1999), which could ultimately promote linguistic manipulation, discrimination, and social inequity.

For example, considering a situation in which two L2 colleagues are preparing a presentation together in a café, successful communication would require (1) an understanding of the communicative goals and the sociocultural context of the meeting (situational understandings), (2) the use of semantico-grammatical resources (forms and semantic meanings), (3) the exchange of topical information (propositional meanings), (4) the accomplishment of interactional goals in talk-in-interaction (functional meanings), and (5) the nuanced communication of other implicated meanings relevant to the context (pragmatic inferences), such as a sense of camaraderie, collaboration, and comity. Finally, the ability to integrate these components in the goal achievement depends upon the users' (6) socio-cognitive mechanisms relating to the brain's architecture (e.g., memory), its functionality through processing (e.g., strategies), and (7) other dispositional factors (e.g., engagement, effort, attitude) (Purpura 2014b). In sum, successful communication in this context involves a complex network of interacting competencies, which can be assessed independently or as a whole, but each can potentially contribute to score variability.

In this *meaning-oriented model of L2 proficiency*, L2 knowledge depends on two mental assets: semantico-grammatical knowledge and pragmatic knowledge, both inextricably linked at the level of meaning in communication. *Semantico-grammatical knowledge* involves a user's knowledge of *grammatical forms* and their associated *semantic meanings* on the one hand and their ability to use these forms together to convey *literal propositional* or *topical meaning*. *Knowledge of grammatical forms* involves linguistic features at both the (sub)sentential (i.e., phonological/graphological, lexical, morphosyntactic forms) and the discourse levels (i.e., cohesive, information management, interactional forms). Knowledge of these forms has often been assessed in terms of *accuracy* or *precision*, *range*, or *complexity* or can also be inferred through characterizations of L2 production (i.e., percentage of error-free clauses) (see Ellis and Barkhuizen 2005).

Semantic meaning (also referred to as grammatical or literal meaning) is more complex. At the subsentential level, it encompasses the literal or propositional meaning(s) associated with individual forms. For example, semantic meaning on the subsentential level can be associated with the dictionary meaning of a lexical item, the morphosyntactic meaning of a past tense form (= past time, completed action), the referential meaning of a cohesive form (*hence* = conclusion), or the interactional meaning of a discourse marker (*anyway* = topic shift marker).

At the sentential level, however, grammatical forms along with their semantic meaning(s), arranged in syntax, conspire to produce the *literal propositional* or

topical meaning of the utterance. Literal propositional meaning encodes the topical content of a message and is often referred to as *factual, literal, topical, subject matter, domain specific, or disciplinary content*. Propositional meaning references subject matter literality, truth-conditional literality, or context-free literality (Gibbs 1994) and is available in LTM through topical knowledge by accessing (1) explicit semantic memory of facts, concepts, ideas, principles, rules, scripts, frames, or algorithms; (2) explicit episodic memory of states, episodes, situations, or experienced events; (3) autobiographic memory (Baddeley et al. 2009); and so forth (See Fig. 2). Some testers have vaguely referred to this as “general background knowledge.” Interestingly, the literal propositional meaning of an utterance is its default meaning, especially when insufficient extralinguistic context is available for interpretation. Literal propositional meaning can be a source of ambiguity in indirect speech and is, amusingly, a critical part of puns (e.g., *A boiled egg in the morning is hard to beat*). With additional context, however, ambiguous propositional meanings often give way to the speaker’s functional meaning in context for interpretation. Finally, the ability to convey propositional meaning depends on the user’s ability to relate conceptual mappings available in LTM to situative contexts in order to generate propositional content (Pellegrino et al. 2001).

The propositional meaning of utterances or texts is often measured in terms of *meaningfulness* or *content control*, referring to the extent to which a user *gets her message across*, or the degree to which the topical content is accurate, relevant, sufficiently elaborated, and original. Propositional meaning can also be measured through *comprehension*, or the extent to which the *topical meaning of the message or text is understood*. Thus, the propositional or topical meaningfulness of utterances or texts encodes the user’s expression or comprehension of content as it reflects a felicitous representation of the real world. Finally, in some assessment contexts, propositional meaningfulness is assessed via L2 production features such as the number of idea units encoded in texts (see Zaki and Ellis 1999).

The current scoring guide for the speaking section of the *TOEFL Primary* (ETS) provides a good example of how propositional knowledge has been operationalized in their scale descriptors.

Language use, content, and delivery descriptors (TOEFL Primary)

The test taker fully achieves the communicative goal

A typical response at the 5 levels is characterized by the following

The meaning is clear. Grammar and word choice are effectively used. Minor errors do not affect task achievement. Coherence may be assisted by the use of connecting devices

The response is full and complete. Events are described accurately and are easy to follow

Speech is fluid with a fairly smooth, confident rate of delivery. It contains few errors in pronunciation and intonation. It requires little or no listener effort for comprehension (italics added)

Pragmatic knowledge is the second component in this model and refers to knowledge structures that enable learners to utilize contextual factors such as speech acts, indexicals, presuppositions, situational and cultural implicatures, and conversational and textual structuring to understand, express, co-construct, or negotiate

meanings *beyond what is explicitly stated* by the propositional meaning of the utterance. Pragmatic knowledge is multifaceted, but, for measurement purposes, can be defined in terms of the mental resources related to the communication of *functional* and *implied or implicated meanings* in language use. Thus, pragmatic knowledge depends on both a person's *functional knowledge* and her *implicational knowledge*. So, when a person wanting salt in a restaurant decides to formulate a message about this desire, her linguistic expression of it encodes the propositional meaning of the utterance. *Simultaneously*, her message in this context functions as a request, thereby encoding her agency and intentionality (Bloom and Tinkler 2001); it encodes functional meaning. The ability to understand and comprehend functional meanings in talk and text then depends on a person's *functional knowledge*, a critical component of pragmatic knowledge. Finally, as functional knowledge allows us to use messages to *get things done* in communication, this core competence has been operationalized to generate functional performance descriptors of L2 proficiency as seen in the *can-do* statements of the CEFR (Council of Europe 2001) (<http://www.coe.int/en/web/portal/home>), the TESOL Pre-K-12 Proficiency Standards Framework (TESOL 2006) (<http://www.tesol.org/advance-the-field/standards/prek-12-english-language-proficiency-standards>), the American Council on the Teaching of Foreign Languages guidelines (<https://www.actfl.org/publications/guidelines-and-manuals/actfl-performance-descriptors-language-learners>), and the Canadian Language Benchmarks (2012) (<http://www.language.ca>).

More interestingly, *pragmatic knowledge* also involves knowledge structures that enable learners to simultaneously encode, onto these same utterances or texts, a wide range of meanings that are implicated by shared presuppositions, experiences, and associations with reference to the communicative situation. This can be done through the select use of verbal and nonverbal resources in conjunction with a range of contextual factors. The ability to understand and comprehend these implied meanings in talk and text then depends on a person's *implicational knowledge*, another critical component of pragmatic knowledge. For example, the person in the restaurant, mentioned above, had a choice of making her request for salt in several ways. She could have been friendly, patient, and witty or aloof, demanding, and snide. These meanings can all be encoded in the simple request for salt. Given the complexities of pragmatic inference, these meanings often pose a serious challenge to L2 speakers and are clearly associated with L2 proficiency. In Fig. 3, I have identified the following seven types of implied pragmatic meanings encoded in talk and text (adapted from Purpura 2004, p. 91):

- ***Situational meanings***⁵: based on understandings of the local context of situation (i.e., how to communicate meanings *specific to a given situation*) – e.g., acceptable, appropriate, natural, and/or conventional use of indirect functions,

⁵In Purpura (2004) the term *contextual meanings* was used. The term *situational meaning* is now preferred as it attempts to codify meaning extensions derivable only from the local speech event (i.e., *you had to be there to get it*).

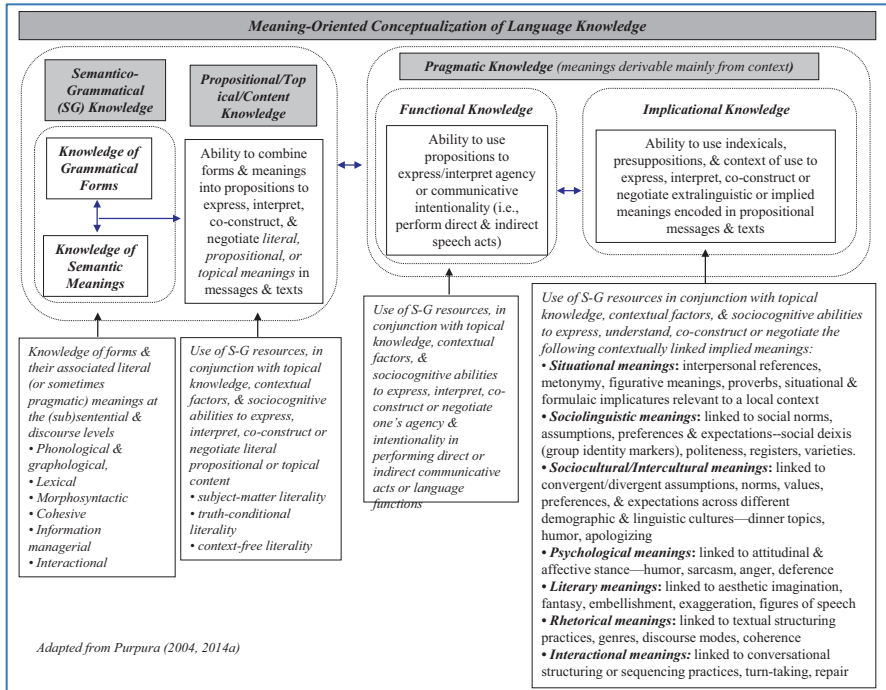


Fig. 3 Meaning-oriented model of L2 knowledge (Adapted from Purpura 2004)

interpersonal references or associations, figures of speech, proverbs, and situational and formulaic implicatures

- **Sociolinguistic meanings:** based on understandings of the social norms, assumptions, preferences, and expectations within a specific speech community (i.e., how to communicate with a given person *in a given social context*) – e.g., acceptable, appropriate, and conventional use of social deixis (group identity markers), politeness (relative power, degree of imposition, social distance), registers, varieties, etc.
- **Sociocultural/intercultural meanings**⁶: based on understandings of the convergent or divergent assumptions, norms, values, preferences, and expectations *across* different demographic and linguistic cultures (how to communicate *within a given culture* or *across cultures*) – e.g., acceptable, appropriate, and conventional use of topic, humor, gratitude, regret, and criticism; avoidance of taboos; etc.

⁶Purpura (2004) specified only *sociocultural meanings*; however, as L2 communication in global contexts often involves speakers from diverse languages and cultures, the ability to understand and express *intercultural, cross-cultural, or transcultural meanings* was considered a pragmatic resource for intercultural communication.

- **Psychological meanings:** based on understandings of affective stance (how to communicate mood, attitudes, feelings, emotionality, and other dispositions) – e.g., acceptable, appropriate, or conventional use of humor or sarcasm or the conveyance of anger, deference, patience, affection, self-importance, etc.
- **Literary meanings:** based on understandings linked to aesthetic imagination, fantasy, embellishment, exaggeration, and figures of speech – e.g., appropriate, creative, and original use of literary conventions
- **Rhetorical meanings:** based on understandings of textual structuring practices, genres, discourse modes, and coherence – e.g., acceptable, appropriate, and conventional use of organizational patterns
- **Interactional meanings:** based on understandings of conversational structuring practices, sequencing practices, turn-taking practices, and repair practices – e.g., acceptable, appropriate, natural, and conventional practices associated with conversational norms, assumptions, and expectations

To summarize, pragmatic knowledge refers to the mental structures underlying the ability to communicate functional and other implicational meanings. The ability to utilize these structures in the task completion, however, is more complex, as it involves *pragmatic ability*, or the capacity to draw on semantico-grammatical resources to express or interpret propositional meanings, which, when used in situated interaction, carry contextually relevant layers of implicational meaning. Since pragmatic knowledge is a fundamental component of L2 knowledge, pragmatic ability is elicited in *all* contextualized language use *no matter the level of L2 proficiency*. The components of pragmatic knowledge can be assessed separately, or in combination, for situational, sociolinguistic, sociocultural/intercultural, psychological, rhetorical, or interactional appropriateness, acceptability, naturalness, or conventionality.

In most assessments, the contextual features needed for tasks to *systematically* elicit implicational pragmatic meanings are often insufficient. An exception to this is Grabowski's study (2009), which investigated examinees' ability to use three implicational meanings (sociolinguistic, sociocultural, psychological) in the context of a reciprocal role-play speaking task. The task prompt specified a particular social role for each interlocutor to assume (e.g., neighbor-neighbor), a communicative goal (e.g., get the neighbor to turn down the music at night), background information on the relationship between the speakers (e.g., persistent tensions over noise), culturally relevant information (e.g., values related to territorial rights, noise, and social harmony), and information relevant to the interlocutors' affective dispositions (e.g., frustrated). Thus, the sociolinguistic considerations of task design involved power distributions, social distance relationships, and absolute ranking of imposition; the sociocultural considerations addressed cultural norms, assumptions, and expectations of the situation in the local culture; and the psychological considerations involved a directive to assume a particular affective stance (e.g., frustration). The test taker responses were scored for grammatical accuracy, semantic (propositional) meaningfulness, and sociolinguistic, sociocultural, and psychological appropriateness based on an analytic rubric. The results showed that, in fact, highly

contextualized tasks could be used to systematically elicit propositional meanings alongside a range of implicated pragmatic meanings, which could be consistently scored and scaled across multiple proficiency levels.

The studies presented thus far have conceptualized meaning and meaning transmission mostly from a sociocognitive approach, describing mental representations of meaning (semantic, propositional, functional, and implicational) in the heads of interlocutors as they communicate, so that individual performance consistencies can be scored independently. However, according to proponents of the *socio-interactional approach* to construct definition (described in Purpura 2016), the sociocognitive approach fails to fully account for communicative success, since communicative success involves the *joint co-construction of relevant and appropriate meanings* that emerge from individuals interacting on a moment-by-moment basis to perform some goal-oriented activity (McNamara 1997; He and Young 1998). In the sociointeractional approach, the capacity to communicate meaning is not so much seen as ability within an individual than as the co-construction of meanings created between interlocutors in interaction. Evidence of this is seen, for example, when one interlocutor collaboratively finishes another's sentence or when interlocutors jointly contribute to the development of a topic when telling a story. While it is true that the creation of meanings in interaction is often a joint product of both interlocutors, it is also true that interlocutors avail themselves of individual resources in the co-construction of these meanings. If one interlocutor has fewer resources, the joint product is likely to suffer. Similarly, if sociocultural or intercultural norms of participation require an asymmetrical pattern of interaction (e.g., teacher-student), the joint co-construction of meanings is unlikely to emerge effectively, possibly affecting test performance. Thus, the sociointeractional approach might be better characterized as both a sociocultural *and* psychological phenomenon, where successful meaning conveyance in interaction is located *within* and *across* individuals *inside* sociocultural contexts.

These observations present testers with the conundrum of what to assess in interaction. Do we attempt to assess each interlocutor's capacity to express or comprehend meanings; do we assess the meaningful product of co-construction achieved by interlocutors; or do we assess both? While the idea of assessing only the joint co-construction of meanings is problematic in most assessment contexts, this approach has succeeded in highlighting the need to assess interactional practices related to turn-taking, conversational structure, and so forth. In the end, the ability to use these interactional practices appropriately (or not) in interaction encodes, as we have seen, a host of pragmatic meanings (e.g., the sociocultural meaning of interrupting inappropriately or the intercultural meaning associated with translanguaging).

Finally, a focus on meaning has been the cornerstone of a task-based language assessment (TBLA) approach to construct definition, where assessment revolves around the examinee's ability to use language meaningfully to accomplish tasks, designed as contextualized, real-world activities (e.g., give a presentation). According to Norris et al. (2017), these activities are also designed to require learners to draw on complex cognitive skills and domain-related knowledge, typically

aligned with a task-based language teaching (TBLT) pedagogical framework (Norris 2009). In TBLA, the competences needed to perform tasks are not drawn from a theoretical model of L2 proficiency, but rather are taken from the specific knowledge, skills, and abilities needed to accomplish the task at different performance levels, what Jacoby and McNamara (1999) call “indigenous assessment criteria.”

Task accomplishment in TBLA has been assessed in many ways. Skehan (1998) and most other SLA researchers have evaluated the extent to which the language produced by examinees in task completion displays the linguistic features of complexity, accuracy, and fluency. This syntactocentric focus is, in my view, confusing given TBLA’s focus on meaning in task accomplishment and would be more consistent with task-based pedagogy if this linguistic focus were complemented by a meaning focus involving an examination of the propositional features of L2 production together with judgments relating to the examinee’s communicative functional ability through the successful exchange of meaningful, relevant, and original content. If more subtle characterizations of task completion were needed or if the results of these assessments were used for formative purposes, then TBLA rubrics would need to consider a pragmatic component. After all, we might have completed the task, but in the process offended our interlocutors.

Finally, an excellent example of a task-based approach to measuring functional communicative ability is seen in the English Language Section of Hong Kong’s Curriculum and Development Institute, where assessment is organized around criteria related to the accomplishment of a sequence of goal-oriented tasks. These tasks required examinees to use language *meaningfully* to accomplish tasks they would likely perform in real life. Interestingly, the assessment explicitly specified general and task-specific assessment criteria related to the conveyance of meaning. Evidence of general content control was defined in terms of *topical relevance*, *propositional appropriateness*, *topical coverage*, and *ideational creativity/originality*, as seen in Fig. 4.

General and Task-specific Criteria for Assessing Task 1 – The Most Beautiful Cities in the World Subtask 2: Writing back to your email pal (Writing)	
General criteria for assessing writing	Task-specific criteria
<p><u><i>Content—demonstrating</i></u></p> <ul style="list-style-type: none"> • relevance of ideas to the topic • appropriateness of ideas • substantive coverage • creativity and originality of ideas 	<p><u><i>Content</i></u></p> <ul style="list-style-type: none"> • writer starts by thanking email pal for information on Seattle and asking for missing details • writer describes Hong Kong • no irrelevant or inappropriate content • substantive content

Fig. 4 Task-based assessment criteria for content (http://cd1.edb.hkedcity.net/cd/eng/TBA_Eng_Sec/web/seta.htm)

Work-in-Progress

Several researchers are currently working on the role of meaning in language assessments. Bae et al. (2016) have just published an interesting, although somewhat controversial, paper on the role of content in writing assessment. Defining “content” as “ideas or meaning expressed in writing” and as “the extent to which those ideas are elaborated, developed, logical, consistent, interesting, and creative as well as relevant to the task requirements” (p. 6), they examined the extent to which the content ratings on an L2 writing assessment could be explained by vocabulary diversity, text length, coherence, originality, and grammar. Modeling the direct and indirect effects of these variables by means of structural equation modeling, they found that a substantial proportion of the variability associated with original, reflective, and interpretative content could be explained by the sum of these five elements. Thus, examinees displaying higher levels of content control produced more “original, reflective, and interpretive texts,” thereby conveying greater levels of topical understanding. Bae et al. concluded that in summative assessment contexts, where practicality is always a concern, the assessment of the content *alone* provided an empirically sound, meaningful, and sufficient measure of writing ability.

Timpe Laughlin et al. (2015), interested in developing an interactive pragmatics learning tool for L2 learners of English in the workplace, provided a systematic and comprehensive review of the role of pragmatics as a component of L2 communicative language ability. This review offered a basis for rethinking the pragmatic competence construct. Influenced by a meaning-oriented approach to pragmatic competence, they proposed a model that addressed two fundamental features of communication: interactive construction and context. They then explicitly specified a meaning space in which two interlocutors in a given sociocultural and situational context can be assessed on their display of five distinct but interrelated dimensions of L2 knowledge. These include sociocultural knowledge, pragmatic-functional knowledge, grammatical knowledge, discourse knowledge, and strategic knowledge. Finally, they provided several interesting examples of task types that could be used in the measurement and ultimate development of pragmatic-functional awareness of L2 learners.

Finally, drawing on the Cognitively Based Assessment *of, for, and as* Learning (CBAL™) project (Bennett 2010; Bennett and Gitomer 2009) at ETS and on Sabatini and O’Reilly’s (2013) application of this work to reading literacy assessments, Sabatini et al. (2016) proposed a technique for organizing online assessments to measure the students’ ability to *display and develop* language and topical knowledge while performing a tightly structured and topically coherent sequence of tasks designed to guide them through the resolution of a goal-oriented problem within a real-life scenario (Sabatini and O’Reilly 2013). These scenario-based assessments thus endeavor to measure the extent to which learners, with different levels of background knowledge, understand topical content in written (reading ability) and spoken text (listening ability), develop deep language and topical understandings with targeted assistance (the development of language and topical knowledge), and then use the newly acquired topical information to perform writing and speaking

tasks related to the scenario goal. This assessment is designed to reflect the multifaceted processes people use when working in a group to research and solve a complex problem. For example, a scenario might ask an examinee, along with his virtual group members, to enter a travel contest in which they have to submit a video-recorded pitch of two possible educational trips. To complete this task, examinees have to research websites and summarize the advantages and disadvantages of taking these excursions, learn about their misunderstandings of the texts, remediate these misunderstandings, synthesize the findings, prioritize the advantages over the disadvantages, and provide *meaningful and content-responsible* recommendations for the best trip. This assessment thus provides a perfect opportunity for assessing the display *and development* of L2 proficiency, topical knowledge, and reasoning skills in which contextual factors, L2 resources, topical resources, sociocognitive, and dispositional resources convene to play an explicit role in task achievement.

Problems and Difficulties

While many testers have recognized the critical role of language in expressing meaning in assessments, only a few have endeavored to define the construct in ways that would allow it to be measured systematically and meaningfully. This comes as no surprise as researchers have had difficulty defining meaning and its relationship to L2 proficiency. After all, two broad fields of linguistics, semantics and pragmatics, have grappled with the meaning of meaning for centuries, with no one coherent model. The fundamental challenge with meaning, in my opinion, is that utterances expressed in context do not encode one meaning; they naturally encode several layers of meaning as we have seen. Nonetheless, we all seem to recognize successful communication when we see it.

To illustrate these complexities, consider, at the subsentential level, we can assess the meaning of a phonological form (e.g., rising intonation to encode curiosity) or the meaning of a morphosyntactic form (e.g., past conditional form to encode regret). At the sentential level, we can assess the meaning of a proposition – a statement that can be true or false. However, this becomes really interesting in L2 contexts when similar meanings across languages are not expressed in the same way. For example, *I dropped my pen* in English would be *My pen fell from me* in Spanish and *I let my pen drop* in French. Then, when these same messages are uttered in context with other interlocutors, the mutually conveyance of meanings becomes much more nuanced and complex. *What the speaker said* (propositional meaning) and *intended to communicate with the message* (intended or functional meaning) is overshadowed by *what was achieved by the message* (functional meaning) and *what was implied by it* (implicational meanings). At this point, meanings depend on pragmatic inferences based on contextual factors. On one level, meanings are contingent upon the mental common ground they have established regarding a set of propositions each speaker takes for granted in that context (Portner 2006) or a set of shared contextual associations. For example, a speaker might use a proposition to accomplish some action (e.g., invite), thereby encoding propositional and functional

meaning. Similarly, she might also use the proposition to communicate nuanced subtexts relating to the social or cultural context or to speaker's psychological state of mind. What is complex is that these meanings are simultaneously encoded in contextualized utterances or texts.

The challenge then for testers is what meanings to assess and how to assess them. The answer, of course, depends on the assessment purpose. However, as Grabowski taught us, we can rest assured that by specifying the appropriate amount of context in the input, it is indeed possible to assess only one or all these layers of meaning systematically and meaningfully. The critical takeaway, then, is for testers to think about what meanings they want to test and to score appropriately. Testers need also to be conscious of the meanings they are assessing *implicitly*.

Future Directions

The field of L2 assessment has long engaged in debates about how to define L2 knowledge and what components, other than L2 knowledge, contribute to L2 proficiency. Over the years, testers have learned to acknowledge how tasks, similar to those examinees are likely to encounter in the real world, have served to engage examinees cognitively in L2 use. This led Chapelle (1998) and Chalhoub-Deville (2003) to conclude that in addition to trait considerations (L2 knowledge and strategic competence), L2 performance assessments needed to seriously consider context and interaction. This paper carries this a step further, arguing that topical knowledge expressed through meaning conveyance is equally important and should *always* be specified on some level. It also maintains that the complexity of the construct and the challenges in eliciting meanings systematically should be no excuse for ignoring one of the most fundamental features of communication and therefore of L2 proficiency. In the end, we need to think about meaning in ways that move beyond simple measures of vocabulary knowledge. L2 learners really need to know if *what* they said, *how accurately* they said it, and *what they accomplished* in saying it were effective or not. They also need to know if, in communicating it, they were contextually, socially, culturally, emotionally, and interactionally appropriate.

Related Articles in the Encyclopedia of Language and Education

Lorena Llosa: [Assessing Students' Content Knowledge and Language Proficiency](#).

In Volume: Language Testing and Assessment

Frank Hardman: [Guided Co-construction of Knowledge](#). In Volume: Discourse and Education.

Silvia Valencia Giraldo: [Talk, Texts and Meaning-making in Classroom Contexts](#). In Volume: Discourse and Education.

Nick C. Ellis: [Implicit and Explicit Knowledge about Language](#). In Volume: Language Awareness and Multilingualism.

Anne-Brit Fenner: [Cultural Awareness in the Foreign Language Classroom](#). In Volume: Language Awareness and Multilingualism.

Cross-References

- ▶ [Assessing Students' Content Knowledge and Language Proficiency](#)
- ▶ [Cognitive Aspects of Language Assessment](#)
- ▶ [Content-Based Instruction](#)
- ▶ [Critical Language Testing](#)
- ▶ [Discourse and Second Language Learning](#)
- ▶ [History of Language Testing](#)
- ▶ [Task and Performance Based Assessment](#)
- ▶ [The Common European Framework of Reference \(CEFR\)](#)

References

- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford: Oxford University Press.
- Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. New York: Psychology Press.
- Bae, J., Bentler, P. M., & Lee, Y.-S. (2016). On the role of content in writing assessment. *Language Assessment Quarterly*, 13(4), 1–27. doi:10.1080/15434303.2016.1246552.
- Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning: A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives*, 8, 70–91. doi:10.1080/15366367.2010.508686.
- Bennett, R. E., & Gitomer, D. H. (2009). Transforming K-12 assessment. In C. Wyatt-Smith & J. Cumming (Eds.), *Assessment issues of the 21st century* (pp. 43–61). New York: Springer.
- Bloom, L., & Tinkler, E. (2001). The intentionality model and language acquisition: Engagement, effort, and the essential tension in development. *Monographs of the Society for Research in Child Development*, 66(4), 1–91.
- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.
- Canadian Language Benchmarks. (2012). For English as a second language. <http://www.cic.gc.ca/english/pdf/pub/language-benchmarks.pdf>.
- Canale, M. (1983). On some dimensions of language proficiency. In J. Oller (Ed.), *Issues in language testing research* (pp. 333–342). Rowley: Newbury House.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1–47.
- Carroll, J. B. (1961/1972). Fundamental considerations in testing for English language proficiency of foreign students. Paper presented at the conference on Testing the English Proficiency of Foreign Students, Washington, DC, May 11–12, 1961. Reprinted in H. B. Allen & R. N. Campbell (Eds.), *Teaching English as a second language: A book of readings* (pp. 313–321). New York: McGraw Hill.
- Carroll, J. B. (1968). The psychology of language testing. In A. Davies (Ed.), *Language testing symposium: A psycholinguistic approach* (pp. 46–69). London: Oxford University Press.

- Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends. *Language Testing*, 20(4), 369–383.
- Chapelle, C. (1998). Construct definition and validity inquiry in SLA research. In L. Bachman & A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 32–70). Cambridge: Cambridge University Press.
- Chomsky, N. A. (1957). *Syntactic structures*. The Hague: Mouton.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Dehn, M. J. (2008). *Working memory and academic learning*. Hoboken: Wiley.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, UK: Cambridge University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Gibbs, R. J. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.
- Grabowski, K. (2009). *Investigating the construct validity of a test designed to measure grammatical and pragmatic knowledge in the context of speaking* (Unpublished doctoral dissertation). Columbia University, New York.
- Gumperz, J. J. (1999). On interactional sociolinguistic method. In S. Sarangi & C. Roberts (Eds.), *Talk, work, and institutional order* (pp. 453–471). Berlin: Mouton de Gruyter.
- Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Edward Arnold.
- Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. London: Edward Arnold.
- He, A. W., & Young, R. (1998). Language proficiency interviews: A discourse approach. In R. Young & A. W. He (Eds.), *Talking and testing* (pp. 1–24). Philadelphia: John Benjamins.
- Hymes, D. (1967). Models of the interaction of language and social setting. In J. Macnamara (Ed.), *Problems of bilingualism. Journal of Social issues*, 23, 8–28.
- Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Middlesex: Penguin.
- Jacoby, S., & McNamara, T. (1999). Locating competence. *English for Specific Purposes*, 18, 213–241.
- Lado, R. (1961). *Language testing*. London: Longman.
- McNamara, T. F. (1997). Interaction' in second language performance assessment: Whose performance? *Applied Linguistics*, 18, 446–466.
- Munby, J. (1978). *Communicative syllabus design*. Cambridge, UK: Cambridge University Press.
- Norris, J. M. (2009). Task-based teaching and testing. In C. Doughty & M. Long (Eds.), *The handbook of language teaching* (pp. 578–594). Malden: Blackwell.
- Norris, J. M., David, J., & Timpe Laughlin, V. (2017). *A framework for designing second language educational experiences for adult learners*. New York: Routledge.
- Oller, J. (1979). *Language tests in schools: A pragmatic approach*. London: Longman.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- Portner, P. (2006). Meaning. In R. W. Fasold & J. Connor-Linton (Eds.), *An introduction to language and linguistics* (pp. 137–168). Cambridge: Cambridge University Press.
- Purpura, J. E. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.
- Purpura, J. (2014a). Assessing grammar. In A. J. Kunnan (Ed.), *Companion to language assessment* (pp. 100–124). Oxford: Wiley. doi:10.1002/9781118411360.wbcla147.
- Purpura, J. (2014b). Cognition and language assessment. In A. J. Kunnan (Ed.), *Companion to language assessment* (pp. 1452–1476). Oxford: Wiley.
- Purpura, J. E. (2016). Second and foreign language assessment. *Modern Language Journal*, 100 (Suppl), 190–208.
- Sabatini, J., & O'Reilly, T. (2013). Rationale for a new generation of reading comprehension assessments. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components* (pp. 100–111). Baltimore: Brookes Publishing.

- Sabatini, J., O'Reilly, T., & Purpura, J. E. (2016). *Scenario-based language assessments in educational settings: Theoretical foundations, prototype examples and future applications*. Workshop presented at LTRC 2016, Palermo.
- Sauvignon, S. J. (1972). *Communicative competence: An experiment in foreign-language teaching*. Philadelphia: Center for Curriculum Development.
- Seliger, H. W. (1985). Testing authentic language: The problem of meaning. *Language Testing*, 2 (1). doi: [10.1177/026553228500200102](https://doi.org/10.1177/026553228500200102).
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- TESOL (2006). TESOL Pre-K-12 Proficiency Standards Framework (<http://www.tesol.org/advance-the-field/standards/prek-12-english-language-proficiency-standards>)
- Timpe Laughlin, V., Wain, J., & Schmidgall, J. (2015). *Defining and operationalizing the construct of pragmatic competence: Review and recommendations (ETS Research Report No. RR-15-06)*. Princeton: Educational Testing Service. doi:[10.1002/ets2.12053](https://doi.org/10.1002/ets2.12053).
- Van Dijk, T. A. (1977). *Text and context: Explorations in the semantics and pragmatics of discourse*. London: Longman.
- Van Ek, J. A. (1976). *Significance of the threshold level in the early teaching of modern languages*. Strasbourg: Council of Europe.
- Widdowson, H. G. (1978). *Teaching language as communication*. London: Oxford University Press.
- Zaki, H., & Ellis, R. (1999). Learning vocabulary through interacting with a written text. In R. Ellis (Ed.), *Learning a second language through interaction*. Amsterdam: John Benjamins.