

# Identity and Proficiency: Meaningful Approaches to Learning and Assessment

Trevor A. HOLSTER

## Abstract

Proficiency is central to both learning and assessment, yet proficiency testing and communicative teaching typically address not proficiency, but behavioristic notions of linguistic performance, failing to address whether learners are authentic users of language and constraining them to mimicry rather than language use. This paper argues that the ability to create and project a range of identities is fundamental to becoming a proficient user, and that teaching and assessment must address this use of language. The primacy of language use over behavioristic performance has profound implications for both materials writers and test designers, and it is argued that the nature of commercial texts and tests precludes this use of language, so it is the role of the classroom teacher to fill this void.

## Discussion

Any discussion of language teaching or assessment ultimately rests on defining a conception of language proficiency, as Hadley (2001, p. 8) reminds us, “In one sense, a focus on proficiency has always driven language learning and teaching. Obviously, no program has ever claimed to be oriented toward ‘non-proficiency’ or incompetence.” Improving students’ English is an uncontroversial goal, the problem lies in agreeing what proficiency is and the mechanisms by which it can be

taught or acquired, or in technical terms, in defining a teaching approach. Richards and Rogers' (2001) well known framework defines a teaching approach as including a theory of language and a theory of learning, and similarly we can define a testing approach as including a theory of language and a theory of measurement. From this we can see that a coherent curriculum will define integrated teaching and testing approaches based around a shared theory of proficiency, which in turn means that a theory of proficiency must precede instruction or assessment. Designing curriculums or tests without first establishing agreement on a conception of proficiency is a recipe for confusion.

Clearly there are many different things needed for language proficiency, such as fluency, grammar, vocabulary knowledge, and so on, but determining their relative importance, relationships, and organizing them into a practical framework for classroom use is difficult. One historically influential view is behaviorism, which, as Searle (1998, p. 46) explains, holds that "...the mind reduces to behavior and dispositions to behavior. For example, to be in pain is just to engage in pain behavior or to be disposed to engage in such behavior.", so behaviorists think that we should focus only on objectively observable behavior, not internal mental events because these are subjective, so language proficiency means responding appropriately to stimuli and teaching language means teaching students to respond appropriately to the stimuli they can expect to encounter. The implication of this is that curriculum planners, materials writers, and test designers must base their decisions on analysis of large, representative corpuses of language in use so as to sample the full range of contexts and stimuli and appropriate responses for every conceivable language use task.

Besides the impracticality of such an undertaking, behaviorism has fundamental problems, one being that we do intend our words to convey inner mental states to other people, even if there is uncertainty about how accurately this occurs, as Quine (1970) shows, so behaviorism simply misses the point. Another problem, as Pinker (1994) points out, is creativity in language, because children don't just mimic language they've heard, but quickly begin to create original expressions. Rather than being simple phrasebook memorization, language proficiency involves creative production based on systems of generative rules, or grammars, that

constrain what is meaningful and what is not. This generative grammar view, as Pinker (1994) explains, revolves around a “deep grammar” that defines the rules that allow us to generate sentences that other speakers will comprehend, so we can create new expressions and communicate new meanings, and proficiency is defined by the precision and variety of meanings a speaker can communicate and comprehend.

Unfortunately there are problems with the generative model, one being variability of language use, such as dialects and registers, so a “prototypical native speaker” is impossible to define precisely, and writing a comprehensive list of generative rules does not seem to be a practical undertaking, as Wardhaugh (2006) explains. Also, as Ellis (1994) explains, variation in learner interlanguage is central to language learning, so the generative model may not account for language learning mechanisms, even if it can define what is grammatical.

However, if we define proficiency in terms of communication, as in the currently popular communicative approach (Richards, 2006), then it may not be necessary to formally define all the rules of language because proficient language users can recognize other proficient language users, in a manner similar to the well known Turing test, a proposal for

“...a test of a machine’s capability to perform human-like conversation. ... a human judge engages in a natural language conversation with two other parties, one a human and the other a machine; if the judge cannot reliably tell which is which, then the machine is said to pass the test.” (Wikipedia, 2004)

Communicative views can thus base proficiency on human judgments, but this leaves the problem that we can only judge a speaker’s observable performance, which may show great variation from moment to moment, leading to the distinction between competence and performance, where competence describes our underlying knowledge, which is stable but impossible to measure directly, while performance describes our use of knowledge at specific time, which is objectively observable but variable (ACTFL-ALC Press, 1996; Hadley, 2001). Communicative views and behaviorism both rely on observing behavior, therefore,

but communicative views make inferences about internal mental states, although the reliability of these judgments is a major difficulty in testing (J. D. Brown, 2005; Hughes, 2003).

Given that competence judgments are based on performance, the question becomes what types of performance are important and how these should be categorized (Bachman, 1995; Bachman & Palmer, 1996; Morrow, 1981). Grammar must be a central consideration in any model of proficiency, but there are many other language features that contribute, and it's difficult to isolate one component of proficiency from all the others in dealing with actual language use. Grammatical accuracy, fluency, or vocabulary in isolation are of little communicative value, leading to the view of proficiency as a holistic trait composed of a number of components, and thus describable only in terms of all the components in combination, not in isolation.

Canale and Swain (1980) developed an influential model of proficiency, where communicative competence is divided into grammatical, contextual, and sociolinguistic competencies. Bachman (1995) followed this with a distinction of language competence and strategic competence, where language competence is the knowledge specific to language while strategic competence is the use of general metacognitive abilities to apply language competence, as shown in Figure 1. Language knowledge is divided into organizational knowledge and pragmatic knowledge, with organizational knowledge subdivided into grammatical knowledge of sentences, and textual knowledge of cohesive discourse. Pragmatic knowledge includes knowledge of propositional content, functional knowledge of speakers' intentions, and sociolinguistic knowledge of context. Similarly, strategic competence is divided into three sets of metacognitive strategies; assessment, goal setting, and planning, further subdivided into seven metacognitive strategies.

As Hadley (2001) explains, the ACTFL Proficiency Guidelines are an influential attempt at defining proficiency, so tests based on this, such as the Standard Speaking Test, or SST (ACTFL-ALC Press, 1996), aim to holistically assess global communicative proficiency, and attempt to address multiple components of proficiency simultaneously. However two things are given quite high priority in the SST, on the basis that they are considered relatively simple to define and assess,

and provide proxies that correlate highly with overall proficiency, so these contribute to test practicality, as emphasized by Bachman and Palmer (1996) and are typically the first features that SST raters consider. The first of these features is known as “oral text type” in SST terminology, the other “initiative”, so it’s worth quickly reviewing these, as shown in Table 1, derived from the *Standard Speaking Test Manual* (ACTFL-ALC Press, 1996).

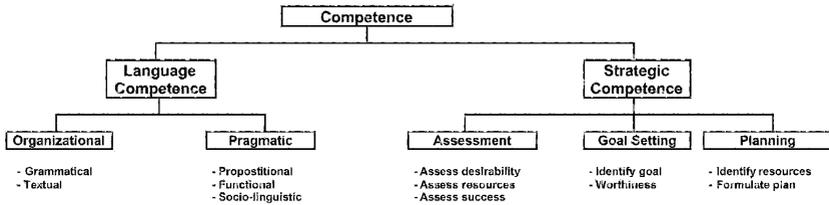


Figure 1 Bachman’s Taxonomy of Competence (from Bachman, 1995)

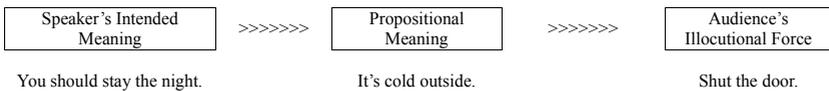
Table 1 ACTFL Proficiency Descriptors

	Oral Text Type	Initiative
Novice	Words, phrases, some sentences	Depend almost entirely on sympathetic partner, typically only able to respond to direct questions.
Intermediate	Variety of sentence patterns, strings of sentences, some use of subordination, can’t organize longer discourse effectively.	Can initiate and maintain communication with a sympathetic partner on familiar topics in informal settings.
Advanced	Can organize complex sentences into paragraph-like discourse.	Can initiate and manage communication with non-sympathetic partners in a wide range of formal and informal contexts.

Oral text type appears closely related to Bachman’s organizational competence and is heavily weighted because of its practicality. Organizationally, novice speakers can’t always sustain performance at the level of sentences, intermediate speakers use a wide variety of sentence patterns, but can’t always organize these into coherent discourse, while advanced speakers can produce coherent paragraph-like discourse. Initiative relates to who carries the burden of communication, so novice speakers depend almost entirely on a sympathetic partner to initiate and maintain communication, whereas advanced speakers have the full range of skills needed to initiate, maintain, manage, and terminate communication on their own terms. In other words, advanced level speakers can negotiate

meaning autonomously, instead of being dependent on a sympathetic partner to manage the communication.

The emphasis within the ACTFL framework on negotiation of meaning complicates our theory of proficiency however, with major implications. When we say something, our words have a literal propositional meaning, for example “It’s cold out there.”, but there could be several intended meanings, such as “Why don’t you stay the night?”, or “Shut the door!”, so the listener must infer the speaker’s intention from the context, and then make a decision about the illocutional force of the utterance. However, this might differ from the speaker’s intention, as Hatch (1992) explains, which puts subjective judgments at the heart of proficiency, as illustrated in Figure 2.



**Figure 2** Subjectivity of meaning.

Further complicating this, we have literary devices such as irony, where the intended meaning is opposite to the propositional meaning, and metaphor, where the intended meaning isn’t part of the propositional meaning at all, so before a listener can assign an illocutional meaning, they must assess the relevance of the propositional meaning to the context of the utterance, so comprehension depends heavily on contextual and cultural knowledge, as well as the linguistic knowledge used in decoding the propositional meaning. If the utterance does not match any contextual clues available to the listener, then the listener is unable to find any relevance in the utterance, and it will be pragmatically meaningless regardless of whatever propositional meaning it may carry. However, every person’s stock of knowledge and experiences are different, so what is relevant to one person may be completely irrelevant to another, and thus different listeners might make completely different judgments about the pragmatic meaningfulness of an utterance because the pragmatic meaning we assign to an utterance is so entwined with our existing schemata of knowledge and experience. Thus meaningfulness is subjective and cannot be described purely in terms of linguistic coding.

In order to ensure that the participants in communication are actually communicating, we must therefore give constant feedback to each other to check that we have a shared agreement on what our utterances mean, so negotiation of a shared meaning that is relevant to the existing knowledge and experiences of all participants is at the heart of communicative proficiency. From this it is obvious that the audience does not just receive meaning from the speaker, but must make considerable effort to actively construct meaning, and, as Hatch (1992) argues, this creative process of constructing a shared meaning is central to giving it emotional value to us, over and above the simple communication of facts.

The problem this raises, though, is that the communication of information beyond rather trivial factual data requires us to convince other people to empathize with us in order that they will share our feelings about why the communication is relevant to them, so empathy must precede communication of facts, and thus a model of empathy is central to a description of proficiency. Humphrey (1986, 2006) investigates empathy from the perspective of evolutionary psychology, based on the assumption that the human brain and intelligence evolved due to conferring a survival advantage. Given that, as Pinker (1994) explains, human language has a physical basis in the structure of our brains, the evolution of brains structured to learn and use language implies that language proficiency confers some survival advantage.

Humphrey found gorillas puzzling because their physical behavior is not particularly complicated and doesn't appear to require a lot of cognitive power, so the evolutionary advantage of large brains isn't obvious based on physical behavior. Instead, Humphrey theorizes that gorillas use most of their brain power to calculate social relationships, not physical action, and it is the increased social skills that confer a survival advantage on the large-brained individuals, and Cheney and Seyfarth make similar arguments in describing their notion of "baboon metaphysics" (2007). Similarly, if language was just used to communicate factual information, the subjectivity of meaning would be a handicap because it complicates communication far beyond the propositional meaning that codes the factual data, but if we view language as a tool used to develop empathy with other people and maintain social harmony, then this subjectivity has an evolutionary purpose

because, as Hatch (1992) points out, the negotiation of meaning serves as emotional glue. Thus, rather than describing proficiency in terms of communication of information, we need a model that emphasizes negotiation of empathy.

Humphrey further argues that empathy is based on mechanisms related to physical sensation, not perception, so we physically feel the same sensations as others when we see them undergo a stimulus. For example, if we watch someone having an accident, we have an involuntary physical sensation, the physical jolt in our stomach, that mimics the sensation they feel, and Humphrey argues that this sharing of physical sensations is where empathy comes from, which implies that physical feelings play a central role in language use, and thus that face-to-face contact is essential for developing or assessing language ability.

In addition to empathy, Humphrey also describes an “Inner Eye” (Humphrey, 1986), which allows us conscious awareness of ourselves and conscious monitoring of our interaction with the physical and social environment. Thus it is the combination of physical empathy and abstract consciousness that lets us understand how other people are feeling and construct a theory of our identity in relationship to other people, which then makes it possible to negotiate social roles and use language for communication.

This is illustrated by the problems that Norton (2000) describes in her study of immigrants, namely that before communication can take place, we have to be recognized by other people as having an identity that has the right to speak. When we dehumanize other people, we deny them empathy, so we have to be humanized in the eyes of other people or else what we say won’t be considered meaningful. The immigrants that Norton studied could not negotiate an English speaking identity that matched their own sense of identity and were excluded from meaningful participation in groups they felt they were legitimate members of. Anything they said was irrelevant because nobody would listen, regardless of any propositional meaning that their words might convey, so their inability to negotiate an identity worthy of empathy in the eyes of native speakers rendered them incapable of communicating, regardless of any formal language proficiency they might have acquired.

On this view, individuals do not have a single, unchanging identity, but construct identities to fit the norms of the social context they find themselves in (Joseph,

2004), and we typically belong to a number of different groups (Wardhaugh, 2006). The author, for example, belongs to the groups of pakeha (New Zealanders of European descent), New Zealanders, philosophy majors, language teachers, mountain bikers, and many other groups. Each group has different social and linguistic norms, so the code appropriate for talking to other mountain bikers is not appropriate for talking to other language teachers, meaning that each social identity we wish to adopt requires negotiation with other group members over mutual acceptance of social and linguistic norms. If language learners cannot access the code required to gain acceptance by members of the target group, they will not be seen as legitimate speakers and what they say will not be meaningful to the audience, leading to the problems that Norton documented.

Further to this, unlike the physical environment, our social environment is subjectively constructed through the cooperation and negotiation of people, and thus individuals who are more proficient negotiators have greater power over the social environment and can manipulate it to their own ends. The minimal social survival level of proficiency requires awareness of the norms of the social environment we are in and the ability to adopt a social identity acceptable to others, but advanced proficiency entails more active control over the social environment, so advanced users of language must be able to take initiative in defining social norms and assigning roles.

Thus, in seeking to define language proficiency, we must include negotiating an L2 identity that other language users will accept. Given that learners must learn to negotiate a wide range of social contexts, no single identity will suffice, so, if learners wish to become fully proficient users of language, they must be able to project a range of identities appropriate for the full range of social contexts found in the target language. Without this, no amount of grammar or vocabulary will allow them to communicate, so our proficiency descriptors need to reflect this. Table 2 illustrates how this might be attempted, so, rather than simply teaching students to mimic the behavior of native speakers, proficiency development requires students to develop awareness of their own identity within the social contexts available in the target language and to learn how to negotiate an identity that other speakers will accept and empathize with.

**Table 2** *Modified Proficiency Descriptors*

	Organization	Initiative	Identity Negotiation
Novice	Words, phrases, some sentences	Depend almost entirely on sympathetic partner; typically only able to respond to direct questions.	Have little grasp of the social contexts of the target language, must rely on sympathetic partners to assign social roles, identities are not fully grounded in the social reality of the target language and culture.
Intermediate	Variety of sentence patterns, strings of sentences, some use of subordination, can't organize longer discourse effectively.	Can initiate and maintain communication with a sympathetic partner on familiar topics in informal settings.	Can sometimes project an identity of their own with help from a sympathetic partner, but often only respond to a social reality defined by others, unable to negotiate preferred identity with a non-sympathetic partner.
Advanced	Can organize complex sentences into paragraph-like discourse.	Can initiate and manage communication with non-sympathetic partners in a wide range of formal and informal contexts.	No longer have to accept social roles imposed by others, can participate as equals in defining the social reality and assigning roles and identities.

Defining proficiency is an important first step, but equally important to classroom teachers is finding practical ways that students can improve their proficiency, so it's worth quickly reviewing Communicative Language Teaching, or CLT, given the current emphasis on this methodology. The goal of CLT is naturally to learn to communicate, but crucially communication is also seen as the mechanism of learning, so, by using language meaningfully, learners are claimed to acquire proficiency. (Richards, 2006). One premise of this is that language use primarily revolves around communication, which is contestable (Verity, 1995), and if we consider the tasks typical of commercial textbooks such as Helgesen, Brown, & Mandeville (2004), Martin (2003), Richards, Hull, & Proctor (1997), Soars & Soars (2001), and Wilson and Barnard (1998), we can see that they are predominantly mechanical repetition requiring little creativity or negotiation of meaning, or the rehearsal that Verity (1995) emphasizes. In short, these tasks are behavioristic drills, not meaningful language use.

Rather, meaningfulness in the classroom will rest on students actively working to relate their own identity to the context they find themselves in, which is a

classroom. Thus, following van Lier's (1996) view of authenticity, rather than trying to pretend we are not in a classroom, meaningfulness will arise from learners negotiating identities and roles with other learners and the teacher, so the teacher has an authentic role as a teacher, the learners have roles as students, and it's highly inauthentic to pretend otherwise. However, within the norms of society, there is constant negotiation of identity and roles, so each participant will need to adopt a range of identities according to the immediate context. In a test, for example, the appropriate identities will be different to those in a class party.

Thus, while good commercial textbooks can provide the starting point for a curriculum, commercial reality dictates that these texts must appeal to the widest possible audience so they will lack the subjective meaningfulness that is central to authentic language use and cannot be relevant to specific contexts in the way that teacher developed material can. Also, inexperienced teachers and unmotivated learners must be able to use commercial textbooks mechanically, so they must, by their very nature, be overwhelmingly behavioristic. Some students are simply not interested in meaningful communication in English, they may just have an instrumental need to pass tests, for example, so curriculums have to allow choices about how learners address tasks, just as in real language use we have choices of how and whether we participate in language use. The impossibility of universally meaningful language use tasks means that teachers need to supplement commercial books with customized material that is subjectively meaningful to their class. Such material will focus on negotiation of identity and meaning among a specific group of individuals and may not be meaningful in any other classroom. Thus the creation of meaningfulness falls to classroom teachers, who can access the face-to-face negotiation that underlies meaningful language use, so developing effective teaching material must be considered a key skill of classroom teachers.

Obviously, similar questions arise concerning classroom assessment, so it's instructive to contrast these with standardized tests. Standardized tests are designed to compare large numbers of students at different times or places, so they must be cheap, easy to administer, reliable, and correlate with the knowledge or proficiency under consideration (H. D. Brown, 2004; Henning, 1987). Validity is essential, of course, but a test that allows valid decisions for one purpose might not

do so for other purposes, and the validity of decisions based on standardized tests presupposes reliability. Thus, if candidates take the test twice then they should get the same result or we can't use the test to compare different groups of students. For reasons of practicality, test designers must find a manageable number of proxies for the construct under consideration that are sufficiently cheap and reliable enough to be administered on a large scale, so standardized tests can't directly measure all the components of proficiency. Rather their validity is typically established by showing that their scores correlate with other accepted measures of proficiency (Bachman, 1990; J. D. Brown, 2005).

The problem this raises for classroom teachers is that a curriculum that focuses narrowly on a specific test is unlikely to cover all the components of proficiency, because no single test can sample the full range of language features, but when tests are perceived as having high-stakes, there may be the temptation to narrow the curriculum to focus only on test content, which is known as negative washback, or backwash (Andrews, 2004). However, classroom teachers face a different context to the designers of standardized tests, needing to make different decisions with different constraints and opportunities, as Shohamy (1992) explains. Thus, rather than mimicking standardized tests' emphasis on reliability in one-off, high-stakes contexts, classroom teachers can use many small assessments that cover a wide range of content and skills and emphasize washback and formative features.

Washback is the result of how we perceive a test, which is inherently subjective (Cheng & Curtis, 2004), and classroom teachers, due to the face-to-face contact they have with students, can influence students' perceptions quite directly, whereas standardized test designers cannot. The implication of this is that classroom teachers should be wary about basing classroom tests on standardized tests, but instead need to take advantage of their extended face-to-face contact with learners and the subsequent negotiation of empathy to do things that standardized tests cannot do. Reliability of such tests will almost certainly be poor if they are used as a one-off summative assessment, but by using a series of small assessments and emphasizing the formative benefits, sufficient reliability should be possible that valid and fair decisions are possible for classroom purposes.

Such an approach complicates classroom teachers' jobs somewhat, however,

because the test and curriculum must be integrated parts of a holistic course design, so consideration of how to assess performance on classroom tasks needs to be a central concern of the entire materials writing process, not a mere afterthought. Thus, if we accept that negotiation of identity is central to language use, then our classroom assessment tasks must be designed accordingly, and our classroom teaching tasks must also be designed with assessment in mind.

Unfortunately, given that every classroom is a unique context, no single type of assessment will be universally applicable, so teachers must develop a broad repertoire of assessment tasks and select assessment types appropriate for the particular situation they find themselves. However, it is useful to consider one set of tasks that have been useful with my own university classes and which illustrate one approach to meaningful assessment. The first stage requires learners to write a diary entry every week and then work in self-selected groups to find interesting things from each other's diaries. While they are doing this, the teacher moves from group to group and holds short mini-discussions about their diaries, giving students the opportunity to work together to create a meaningful discourse on familiar topics from their everyday lives. Rather than basing assessment on linguistic performance, grades are based on making a genuine attempt at finding topics of mutual interest and negotiating empathy, on the assumption that making this is what will lead to developing an English speaking identity and making long-term progress in acquiring English. The rehearsal that Verity (1995) emphasizes is central to this, and the emphasis is not on behavioristic mimicry, but on cooperation to create shared subjective meaningfulness. Simply taking up invitations to communicate is enough for a passing grade, but refusing to take part meaningfully is not, and taking the initiative in communication warrants a superior grade. The group nature of the task allows learners to negotiate identities of their own rather than being forced into cookie cutter identities typical of mechanical pair-work activities, so there may be an informal group leader, an authority on grammar, and so on. This activity, although advocated here as an assessment task, was originally designed as a learning task aimed at preparing students to take an interview test such as the ACTFL SST (ACTFL-ALC Press, 1996), and the use of it for assessment was later added to extrinsically motivate learners by rewarding behavior

considered beneficial to language learning, and then adapted to focus more on the intrinsic motivation that van Lier (1996) advocate. This illustrates how teaching and assessment approaches can be integrated around the same tasks, based on a shared model of proficiency and model of learning.

## Conclusions

The central argument of this paper is that a model of proficiency is central to both successful teaching and language learning, but that behavioristic conceptions cannot adequately describe meaningful language use, involving generation of empathy through face-to-face contact and negotiation of identity. If CLT's claim that authentic language use is the learning mechanism, then we must emphasize negotiation of empathy and identity, but commercial textbooks and standardized tests cannot achieve this, so it falls to classroom teachers, who must focus on materials and assessments that address the needs of the individual learners rather than relying on the generic tasks found in commercial textbooks.

## References

- ACTFL-ALC Press. (1996). *Standard speaking test manual*. Tokyo: ALC Press.
- Andrews, S. (2004). Washback and curriculum innovation. In L. Cheng & Y. Watanabe (Eds.), *Washback in language testing* (pp. 37-50). Mahwah: Lawrence Erlbaum Associates.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (1995). What does language testing have to offer? In H. Brown & S. Gonzo (Eds.), *Readings on second language acquisition*. (pp. 415-447). Upper Saddle River: Prentice Hall Regents.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains: Longman.
- Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment (New Ed.)*. New York: McGraw-Hill.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.

- Cheney, D., & Seyfarth, R. (2007). *Baboon metaphysics*. Chicago: University of Chicago Press.
- Cheng, L., & Curtis, A. (2004). Washback or backwash? In L. Cheng & Y. Watanabe (Eds.), *Washback in language testing* (pp. 37-50). Mahwah: Lawrence Erlbaum Associates.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Hadley, A. O. (2001). *Teaching language in context (3rd ed.)*. Boston: Heinle and Heinle.
- Hatch, E. (1992). *Discourse and language education*. Cambridge: Cambridge University Press.
- Helgesen, M., Brown, S., & Mandeville, T. (2004). *English firsthand 1*. Hong Kong: Longman.
- Henning, G. (1987). *A guide to language testing*. Boston: Heinle & Heinle.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Humphrey, N. (1986). *The inner eye: Social intelligence in evolution*. Oxford: Oxford University Press.
- Humphrey, N. (2006). *Seeing red: A study in consciousness*. Cambridge: Belknap Press.
- Joseph, J. (2004). *Language and identity*. New York: Palgrave Macmillan.
- Martin, D. (2003). *Talk a lot*. Okegawa City: EFL Press.
- Morrow, K. (1981). Communicative language testing: evolution or revolution? In J. C. Alderson & A. Hughes (Eds.), *Issues in Language Testing* (pp. 9-25). London: The British Council.
- Norton, B. (2000). *Identity and language learning*. Harlow: Pearson.
- Pinker, S. (1994). *The language instinct*. New York: Harper Collins.
- Quine, W. (1970). On the reasons for the indeterminacy of translation. *Journal of Philosophy*, 1970, 178-183.
- Richards, J. (2006). *Communicative language teaching today*. Cambridge: Cambridge University Press.
- Richards, J., Hull, J., & Proctor, S. (1997). *New interchange*. Cambridge: Cambridge University Press.
- Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching*. (2 ed.). Cambridge: Cambridge University Press.
- Searle, J. (1998). *Mind, language, and society*. New York: Basic Books.
- Shohamy, E. (1992). New modes of assessment: the connection between testing and learning. In E. Shohamy & A. Walton (Eds.), *Language Assessment for Feedback: Testing and Other Strategies*. Dubuque: Kendall Hunt.
- Soars, L., & Soars, J. (2001). *American headway 1*. Oxford: Oxford University Press.
- van Lier, L. (1996). *Interaction in the language curriculum*. Harlow: Pearson.
- Verity, D. (1995). Rehearsal and language learning: Unburdening the method. *Speech Communication Education* 8.
- Wardhaugh, R. (2006). *An introduction to sociolinguistics (5th ed.)*. Malden: Blackwell.
- Wikipedia. (2004). Turing test. from [http://en.wikipedia.org/wiki/Turing\\_test](http://en.wikipedia.org/wiki/Turing_test)
- Wilson, W., & Barnard, R. (1998). *Fifty-fifty: Book 1*. Hong Kong: Longman.

