CONTINUING COMMENTARY

Apples and Oranges Are Both Fruit, But They Don't Taste the Same: A Response to Wynne Wong and Bill VanPatten

Betty Lou Leaver
Coalition of Distinguished Language Centers

Benjamin Rifkin
University of Wisconsin–Madison

Boris Shekhtman
Specialized Language Training Center
Washington Branch CDLC at Howard University

with David Branner, Patricia Chaput, Dan E. Davidson, Gerard L. Ervin, Karen Evans-Romaine, Thomas J. Garza, Olga Kagan, Maria Lekic, Cynthia Martin, Frank Miller, George Morris, Richard Robin, and Andrei Zaitsev[1]

We understand the desire of Professors Wong and VanPatten ("The Evidence is IN: Drills are OUT," *FLA*, 2003, 36[3], 403–23) to share with teachers of the less commonly taught languages the ideas and philosophies they have developed for the more commonly taught ones. With all due respect, we must point out that the application of their experiences was, quite simply, fraught with errors bound to occur when specialists in one foreign language attempt to generalize their experiences to other languages in which they do not have adequate background for making appropriate and accurate comparisons.[2]

There has been considerable empirical research on long-term successful acquisition of Russian, especially in intensive and study-abroad programs. Despite this abundance of information, Wong and VanPatten referred to only one study—one whose general applicability is highly questionable because it had such a small number of subjects (22 Russian students). Furthermore, they extrapolated from that study conclusions that were not examined in the study and used them in a way that even the study's authors stated would be inappropriate and likely inaccurate.

In this essay, we will discuss Wong and VanPatten's claims about learning Russian (i.e., it is just like learning any other language, except that it takes longer) and the questionable place of drill in the study of Russian. There is a large body of evidence that refutes both these claims. We will then turn to the broader question of the place of drill in foreign language study more generally.

Setting the Record Straight About Russian

Wong and VanPatten argued:

> One of the objections we sometimes hear from colleagues in language teaching is that Russian is "more difficult" or that Japanese is not Indo-European and thus learning these languages requires special or different instructional approaches than learning Spanish or French or even English as a second language. It is true that Russian involves a different alphabet and has little Latinate basis on which to rely for teaching and learning in a classroom. It has a complex morphological system for verbs and nouns when compared to English, Spanish and French. Japanese, too, has a different writing system, no cognates with English, a complicated system of honorifics, and so on. Similar arguments can be made for Amerindian languages like Quechua and Inuit. Nonetheless, learners do acquire these languages without instruction, as is evidenced by the diaries and record of Marco Polo and missionaries who came to the new world and by the multilingualism that must have existed ever since the first time two different cultures came

> into contact . . . Imagine if a scientist suggested that the laws of physics developed on Earth are not applicable to the Moon, Mars, or another galaxy.
> . . .
> Our point here is not that instruction cannot help the learning of Russian, Japanese, or any other language. It is that the role of drills cannot change depending on language. Drills are no more necessary for Russian than they are for Spanish or English. What is plausible, of course, is that learners might need extra help in getting linguistic data from the input. This is precisely the aim of PI as described and discussed in the previous section as well as all the other work on focus on form (Doughty & Williams, 1998). In one study, Kempe and MacWhinney (1998) showed that learners of Russian could acquire case marking without explicit instruction or drills. This was not the point of their study, but we cite it here as an example of research on one of the "more difficult" languages, to illustrate that drills are not necessary: Learners can learn Russian from input just like anyone else. They just may need to have it more structured and may need more time.

There are many points here with which we take issue. Wong and VanPatten made several serious errors in arguing their case, including (1) mistakes about the level of difficulty of Russian, (2) misunderstanding where the difficulties lie in learning and teaching Russian, (3) the erroneous concept that "laws" of linguistics (or even physics) are fully known, immutable, and universally applicable, and (4) asserting that the universality of laws of physics provides a good analogy for laws of language learning.

Level of Difficulty of Russian

It is unclear to us why the fact that Russian is a more difficult language than Spanish or French for native speakers of English was conveyed by means of quotation marks intended to make this statement suspect. In fact, research at the Foreign Service Institute long ago established that Russian (and most other Slavic languages) are significantly more challenging than Romance and Germanic languages for native speakers of English and that Japanese (and other East and Southeast Asian languages) are significantly more challenging than the Slavic languages (see Omaggio Hadley, 2003, p. 26).

In fact, languages like Chinese and Russian require higher language-learning aptitude scores for enrollment in U.S. government language training institutes than do languages such as French or Spanish. In addition, Gass and Selinker (2002, pp. 75–77, 130–32) described several studies of language distance; work in this area is quite well established. Scholars, teachers, and students recognize the truth that some languages are more challenging for native speakers of English than others. (If readers doubt this idea, they need only ask university students seeking the "easiest" way to fulfill a foreign language requirement if they will consider taking Japanese!)

Difficulties in Teaching and Learning Russian

Now, let us turn to our area of expertise, the teaching and learning of Russian. Wong and VanPatten suggested difficulty where little to none exists and missed the truly difficult moments in the acquisition of Russian. Alphabet and vocabulary are really some of the "easier" aspects of Russian, with case markings and verbal morphology presenting moderate difficulties. The complicated features, such as those contained in the aspectual/semantic nature of the verbal system, the flexibility of Russian word order, and the difference between formal and informal use of the language, which surveys and the experience of experts show to be acquired more readily through direct instruction and repetition, were completely ignored by Wong and VanPatten.

The use of the Cyrillic alphabet presents relatively little or no challenge for most learners. This topic is dispensed with in the very beginning of most Russian language programs, sometimes in as little as one class period, depending on the method of instruction being used (B. L. Leaver, 1984).

To say that Russian has no Latinate base is true in terms of linguistic relationships. However, historically, there has been a strong influence of French on the Russian language, and there are a large number of Latinate roots that can lead students to acquire at least some vocabulary without great difficulty.

The morphology of Russian, on the other hand, is complicated, with nouns *and modifiers* marked for three genders, two numbers, and six cases (and for a few words, a seventh case), with phonological variations and numerous exceptions. This means that students must learn to use accurately hundreds of endings that carry significant meaning; to misuse the endings, in many cases, results in the conveyance of incorrect information, disabling communication. More important, however, meaning in the Russian language is expressed principally morphologically, rather than syntactically as in English—a major psychological hurdle for students to cross in acquiring a "feel" for the language and locating the source of information in any given utterance.

Nonetheless, the morphology of Russian nouns, adjectives, and verbs presents only moderate difficulty compared with the verbal system itself. The verbal system of Russian and other Slavic languages features an aspectual system unlike anything in Romance or Germanic languages. There is often no match between vestigial aspect in English and Slavic aspect, and in some instances, the "rules" change, depending on the verb itself; aspect is both semantic and

grammatical. The amount of comprehensible input needed to acquire this system is far beyond that which one can get in a classroom through authentic materials alone and in the time available. Since there is no comparable paradigm that English speakers possess, this means, in Piagetan terms, that a brand new category of information must be developed. Without direct help, this usually does not happen, and even with help, aspect is acquired very late (usually at a strong superior, even distinguished, level of proficiency).[3]

There is also the uniquely challenging system of verbs of motion in Slavic languages. There are at least 36 choices of verb (plus their various conjugations)—all of them meaningful—to say "to go." New ways of looking at motion have to be developed and the accompanying forms internalized. The number of repetitions (a form of natural drilling) that Russian children get in internalizing these verb choices is simply not possible in the classroom. In our experience, these forms are generally acquired through a combination of direct instruction, classroom practice (meaningful and communicative drills), and subsequent study abroad.

Acquiring Russian is not, then, a matter of simply more and more input, that is, more time on task. There are individuals, many of them, who have spent 20, 30, and 40 years in a culture without acquiring superior-level fluency in its language. Since so many features of Russian grammar do not exist in the English grammatical system, Americans with only naturalistic acquisition of Russian may come close enough to be understood (or are misunderstood and do not realize it), but their language remains at a very low level of sophistication, marking them as semiliterate at worst and foreign at best. In a Vygotskian sense, they never reach the Zone of Proximal Development for acquiring the most critical (meaningful) features of Russian language. This may be, to a large degree, because they are not children and have a grammatical system already in place; therefore, the logical next step in learning is something that matches that system or is something entirely new to add to it. Where there is a semimatch, as in the case of aspect, there is great confusion that, in our collective experience, is rarely resolved without direct instruction, including explication and controlled practice.

Overgeneralizations

Wong and VanPatten argued a logical fallacy: If laws of physics apply everywhere in the universe, then the principles proposed for the teaching of Spanish must also be applicable to the teaching of Russian or Japanese. First, we would note that few physicists would claim that they have all the laws for all the physical phenomena in the universe; in fact, the laws of physics have been updated several times in our lifetime alone to account for newly discovered phemomena that previous systems could not explain. Therefore, the hypothesis itself—if laws of physics apply everywhere—is suspect. Second, if such a proposal is true, we would like to see the evidence of successful implementation of such an approach to the teaching of Russian. We know of none, and Wong and VanPatten failed to point out any. Third, the research that *is* available indicates that teaching and learning Russian is different from teaching and learning English or the more commonly taught foreign languages (B. L. Leaver, 2001).

A survey of 55 languages users with tested proficiency at the near-native level (ILR Level 4, ACTFL Distinguished Level) or native-language equivalence, showed responses from the 33 Russian learners (estimated to be approximately one sixth of all identified American speakers of Russian at Level 4 and higher)[4] that differed from the responses of learners of English, Spanish, and the more commonly taught languages. Specifically, 67% of the Russian learners (compared with 50% of foreign-language learners in general) considered grammar explanation crucial to their success in language acquisition, 58% (compared with 21% of the overall group) wanted grammar drills even at high levels, and 58% (compared with 21%) also wanted opportunities for the deliberate practice of grammar. All students who had received direct instruction at high levels considered it instrumental to their success and 75% of those who had not had it felt that they would have reached the higher levels faster had they been instructed. By instruction, these respondents were alluding to deductive, overt classroom situations.[5]

Wong and VanPatten argued their case on the basis of a study by Vera Kempe and Brian MacWhinney, "The Acquisition of Case Marking by Adult Learners of Russian and German," which appeared in *Studies in Second Language Acquisition* (1998, 20, 543–87). We argue that this paper cannot be used to justify an approach to teaching Russian without grammar drills. Kempe and MacWhinney's study was based on a sample of merely 22 learners of German and 22 learners of Russian. Although this sample size was more than adequate for a case study, it is not acceptable for broader generalization to the populations of learners of German and Russian as a whole. Indeed, detailed information about these learners was not provided, so we cannot say whether they were male or female, how old they were, how much prior instruction they had in the language of question or any other foreign language, what their high school or college GPAs or SAT, ACT, or IQ scores were, or how they scored on any other measures of academic achievement or intelligence. It is therefore impossible to know whether or not the two groups of learners were, in fact, *comparable*.

Kempe and MacWhinney's point was that the Russian learners did better than the German learners in learning case markings despite the fact that Russian case markings are more complicated than German case markings. It is possible, however, that a group of students that elects to

study a more difficult language (Russian) might, on the whole, have greater language-learning aptitude or skill than a group of students taking a less difficult language (German). Indeed, many of the German learners may have started their study of German in high school or junior high school and may have had a study program that did not train them to study case endings, whereas learners of Russian are more likely to have begun their study of Russian in college (since there are fewer high school programs in Russian) and much of the college Russian curriculum focuses on teaching case endings. In fact, the grammatical success of college Russian programs could be the very source of the greater success of the Russian learners in the Kempe and MacWhinney experiment.

Most important, however, Kempe and MacWhinney focused on *reception* in their paper and they explicitly stated that they were making *no claims about production* skills by the two groups of language learners:

> [This] research has focused **exclusively** on comprehension …. the present study makes no claims about the **active use** of case marking in the **production** of Russian and German by L2 learners.... [boldface emphasis is ours] (Kempe & MacWhinney, 1998, p. 581)

Thus, we argue that the Kempe and MacWhinney study has been misused by Wong and VanPatten, because the first team of researchers investigated a phenomenon in the *comprehension* of case markings, but the latter team of researchers applied these findings (questionable as they are) to the *production* of case markings.

Unpublished research, using recall protocols, on 102 Russian students at the Foreign Service Institute between 1983 to 1989 showed that students, on average, did not *acquire* case at novice and intermediate levels but rather frequently *recognized* it. That is, they did not interpret case in meaningful ways, indicated by students' misinterpreting case in Russian OVS and OSV sentences, which differ syntactically from SVO sentences that parallel English structure. In other words, when the syntax was the same as English syntax, students recognized case accurately; when it differed, they did not do so until they approached the superior level of proficiency.[6] It could be argued, then, that although they seemed to understand case, they were actually processing meaning through syntax, just as in English. The same might have been true of the Kempe and MacWhinney students —they may not have been *processing* case at all, but simply recognizing it in SVO sentences while actually processing the given information through lexicon.

Examining the Full Scope of Language Acquisition (and Learning)

Unfortunately, Wong and VanPatten focused their entire argument on the acquisition of language at the novice and intermediate levels. Such a focus can lead to grossly incorrect assumptions about the overall language learning/acquisition process and, worse, to teaching practices that deny or delay the achievement of upper levels of proficiency.

The Difference between Lower and Upper Levels of Language Proficiency

Wong and VanPatten presented no research showing that the learning of grammar in drills cannot help students build a foundation necessary for the ultimate attainment of upper-level proficiency (advanced, superior, distinguished). We will not argue that students in a traditional college Russian curriculum achieve such proficiency levels in the context of four years of classroom instruction, but a fair number of students of Russian who study abroad for a semester or year do make the breakthrough into the advanced level proficiency. Brecht, Davidson, and Ginsberg (1993) argued that success on the preprogram grammar test was the best indicator of language gain in all areas by students participating in study abroad programs in Russia:

> [Q]ualifying grammar and reading achievement scores show significant predictive value for speaking proficiency, reading proficiency, and listening proficiency alike … [H]igher levels of control of basic grammar and reading skills, as measured by [pre-program grammar and reading tests] are positively related to gains in all the skills. Of particular importance here is the strength of the relationships for gains in OPI across different levels and combinations of levels. While significant for all OPI gains, grammar/reading achievement proved to be most significant (t-statistic + 2.6) for the group of learners at the 1+/2 speaking threshold, precisely the critical level in speaking proficiency for the greatest numbers of Americans studying abroad. …The data in the current study provide the first empirical evidence that investment in grammar instruction in the early years of instruction may result in advances in speaking and listening skills at the upper-intermediate and advanced levels. In particular, if one views formal instruction as only one component in a student's language learning career, formal instruction in grammar can be seen as one key element in producing expert language learners who will develop the independent capacity to gather and assimilate information and skills on their own through contact with native speakers. (Brecht, Davidson, & Ginsberg, 1993, pp. 20-21).

Neither we nor Brecht, Davidson, and Ginsberg argue for an abandonment of the communicative approach to language teaching. Rather, we question Wong and VanPatten's views about the role of grammar instruction and, specifically, grammar drill, in the foreign language curriculum. Our goal, and the goal of many foreign language instructors, is to lay a foundation upon which learners can build and ultimately attain the highest levels of communicative

language performance. The quality of Marco Polo's Chinese is questionable: He clearly was able to communicate, but could he communicate at the level we now call "distinguished" or even "superior"? It has even been argued that he probably knew no Chinese and may never have gone to China. In fact, in his era, all government business was transacted in literary Chinese, a formal language extremely difficult even for native speakers of Chinese dialects and one that required native speakers to invest years of study before being qualified for a civil service appointment. It is inconceivable that Marco Polo could have picked this language up without instruction; he cannot, therefore, serve as a model for naturalistic acquisition. As for the missionaries, even they worked assiduously on grammatical accuracy, as evidenced by Francisco Varo's *Arte de la lengua Mandarina* (1703, translated by Coblin and Levy in 2000). Indeed, Varo advised missionaries to master structure before striving for fluency and to study set pieces and practice them at great length, until fluency and automaticity began to assert themselves:

> It is preferable to speak less and well than to say a great deal in the wrong order. And if in the beginning there are mistakes, these can become a habit, whereupon it will not be easy to correct them (Varo, trans. Coblin & Levy, 2000, p. 21).

Our collective experience teaching Russian (and learning it) tells us that grammar drills play an essential role in the Russian-language curriculum precisely in the preparation for communicative language performance at levels beyond intermediate-mid. At the intermediate-mid level, students have *no need* to use case markings (or verbal aspect) accurately; they can be understood in the context of basic needs virtually without any correct grammar in Russian. However, if our students are to attain more sophisticated control of Russian and to use it in more challenging contexts, they need to gain control of Russian grammar. Our experience teaches us that this control is forged in the classroom and at early levels—otherwise, the duration needed to acquire high-level language skills is considerably greater.

The Role of Direct Instruction in the Acquisition of Foreign Languages

Byrnes (2002) suggested that the Proficiency Movement inappropriately "bypassed" the nature, function, and significance of the formal features of foreign languages in its haste to "implement proficiency." The result has been an accuracy or fluency split that is unnatural, since structure is part of meaning-making. The issue, as many have said, is not *whether* to teach grammar but rather *how* to teach it.

Similarly, in the earlier-mentioned study of successful language learners who had achieved Levels 4 and 5 (Leaver & Atwell, 2002), one surprise was the significance that all of the learners placed on direct instruction. In fact, in listing the factors that contributed to their success, 100% of the respondents described direct instruction as very important and 80% considered the improvement in their understanding and control of grammar to have been critical to their achieving high levels of proficiency.

Teachers of Russian who have successfully brought students to Levels 3 to 4+ believe very strongly in the role of direct instruction and the value of drills for creating automaticity where none existed and correcting incorrect-but-automatic patterns where these did exist (Shekhtman et al., 2002; Shekhtman, Lord, & Kuznetsova, 2003). An informal survey of Slavic language teachers working at high levels of proficiency showed that nearly all of the teachers taught grammar overtly and very often used drills.[7]

Lack of Grammar and Fossilization

Learners who learn only through comprehensible input, particularly outside the classroom (i.e., in country), often fossilize at lower levels of foreign language learning (Higgs & Clifford, 1982). This phenomenon is common among American soldiers stationed in Germany; their German is "fixed" so that speaking in literate ways can be extremely difficult.

Ehrman (2002) identified five different kinds of fossilization that can keep students from reaching the highest levels of foreign-language proficiency. One of these—strategic fossilization—is the use of inappropriate strategies that are associated with lower levels of proficiency, particularly as taught in input-driven classrooms that cause students to ignore some important information and focus only on the gist. "Without developing attention to the complex relations among language form, meaning, and context, learners find it difficult to cross the SD [Superior-Distinguished] threshold" (Ehrman, 2002).

Soudakoff (2001), who teaches students of Russians to read at Level 4, found fossilization to be a strong hindrance to success at higher levels, one that requires a great amount of repetitive translation (a form of drilling) to overcome. He attributed the lack of preparation for higher-level study to teaching methods such as those proposed by Wong and VanPatten at the lower levels.

Comprehensible (?) Input: Variability in Comprehensibility among Languages

Wong and VanPatten promoted the ultimate fallacy when they assumed, without evidence or attempted research, that contextualized, authentic input will be comprehensible to students of Russian. Some will be; in fact, much will be. However, critical elements often are not because speakers of English do not and cannot have the requisite schemata for comprehending the nature of the input, other than perhaps superficially, and even superficial understanding is not always possible. The deeper processing needed for acquisition and (re)production is unattainable

much of the time for those features that do not parallel English and that are not marked by overt meaning (see note 2 for an example). This is the fallacy that many teachers of English and "easier" languages such as Spanish and French make when it comes to languages that are more structurally complex or more culturally opaque for speakers of English.

Drills that Work

The problem with studies on the effectiveness of drills is that none, to the best of our knowledge, have taken into consideration the content of the drills themselves or the quality/skill of the driller. Shekhtman (2003) contended that if a student does not like drills or does not learn from them, the problem may lie with the teacher: Either the drill has been inappropriately structured or ineptly delivered. Shekhtman defined well-designed drills as those that

- have clear goals;
- put the onus of the work on the students;
- accommodate students with different learning styles;
- are of manageable length and complexity;
- incorporate feedback and judicious, learner-centered error correction;
- move at an appropriate tempo;
- require emotional investment;
- vary in nature (keeping the learning interesting); and
- are fun.

Among the studies that have been done on the effectiveness of appropriate drilling have been those conducted by DeKeyser (1997, 1998). He found empirical evidence that drilling, contrary to popular belief and current theory, was critical for the development of the automaticity that is characteristic of speech that is both fluent and accurate (and, we would add, characteristic of native speech). His explanation is that intensive practice turned declarative knowledge into more habitual and qualitatively different procedural knowledge.

What DeKeyser demonstrated has been intuitively known at government institutions for a very long time; in fact, the Foreign Service Institute incorporates meaningful, communicately-oriented, learner-centered drills as an integral part of its 44-week Russian program designed to bring learners to superior-level proficiency. The Defense Language Institute, which has larger classes and learners who are less expert, does the same in its 47-week program designed to bring learners to advanced-level proficiency. Both institutions have found that the right kinds of drills, used at the right times, modified to meet specific student learning profiles, can be highly effective in an environment where acquisition of language to high levels is not only a goal but also a necessity.

Knee-jerk reactions against drilling in any form are not helpful for the profession. We would argue that more research is needed about the kinds of drills that are needed, at what stages they are needed, and for what languages they are most effective. We would also point out the fact that natural drilling occurs in all language learning activities, including first language, especially in the case of comprehensible output. One of two ways of getting information into long-term memory is repetition, the other being associative memory (Reiser, 1991; E. Leaver, 1999). Associative memory depends on current, related schemata being present; this is not always the case. So, even in one's own language, repetition plays an important role. Children's songs, nursery rhymes, and games teach first language through what is essentially drill. Classroom instruction in the native language for those parts of grammar that are not part of daily living are typically overt. Those who would claim that direct instruction is "unnatural" are discounting a good portion of the "natural" learning spectrum.

Drills do not have to be (and should not be) the choral repetitions that are reminiscent of ALM classrooms. They can be highly individualized and highly isolated—thrown at an individual student for multiple, rapid repetition meant to correct a fossilized error that has cropped up in a communicative act. As such, they are both training and awareness activities.

A Better Course for the Future: Individualization, Not Generalization

The Case of Polyglots

Wong and VanPatten made the same mistake that many nonpolyglots make: They erroneously assume that all languages are learned in the same way. The evidence from the study of polyglottic situations is that the same individual will learn two unrelated languages in sometimes very different ways. The Level 4 study described above examined the way in which polyglots—individuals who spoke at least two foreign languages at native-like levels—had acquired their languages and found that different languages were acquired differently by the same individual. Moreover, the patterns among languages were more evident than the patterns among individuals. Most of those acquiring Slavic languages reported learning the language through a combination of direct instruction in a cognitive mode, study abroad or its equivalent, and self-drilling for comprehensible output through structured interactions with native speakers. Those learning English and Romance languages reported learning them through a more inductive approach, with less direct study (or the desire for it) than was reported for languages like Chinese and Russian (B. L. Leaver, 2001).

Variability Among Students

Wong and VanPatten, in arguing for one approach (theirs) to fit all, excluded a key component in making language

study accessible to all students: adaptation for learning style. Some learners learn much more readily in a more structured approach than do others. In fact, Wong and VanPatten revealed their own styles in their article, and their style is diametrically opposed to that of the ectenic learner, who needs correction, structure, hypothesis confirmation, direct instruction, and the inclusion in their language programs of some elements of more traditional approaches to language study, along with being deliberately provided the tools for survival in communicative classrooms, if they are to succeed (Ehrman & Leaver, 2003).[8] Although no student likes a poorly constructed and/or poorly delivered drill, the ectenic student appreciates drills. On feedback forms, such students are apt to praise the teacher who drills and scorn the teacher who does not (B. L. Leaver, 1999b). It is not a matter of the student not understanding what is and is not good teaching. It is a matter of the student subconsciously knowing what does and does not work for him or her.

Most of today's approaches are synoptic in nature. As such, they benefit the synoptic learner more than the ectenic learner. Clearly, such styles as induction, globality, synthesis, randomness (i.e., stochasm), and leveling (key elements of synoptic learning) assist the learner who is presented with authentic input in divining meaning from it. Equally clearly, such ectenic styles as deduction, particularity, analysis, sequentiality, and sharpening hinder the learner from comprehending authentic input without assistance with the input and with learning strategies to cope with such input. Unfortunately, most teachers are not yet skilled enough in learner-centered instruction to fine-tune tasks and activities, so as to make input truly comprehensible.

Summary

Wong and VanPatten's proposed methods disenfranchise some kinds of learners. Learners are different: Diversity of learning styles is well established in the research literature. Moreover, Wong and VanPatten's approach is not, contrary to their claim, universally applicable to all languages, since even the same learner will approach the study of different languages differently, precisely because languages are different. Wong and VanPatten did not consider the reports of learners themselves, reports in which learners identify grammar instruction as critically important to their learning success.

Failing to consider the different needs for short-term and long-term learning goals, Wong and VanPatten did not recognize the importance of building a foundation for the attainment of superior- or distinguished-level proficiency. Although we recognize that the attainment of such a high level of language performance is extraordinarily unlikely in the context of a high school or college curriculum, it is precisely the foundation built in formal instruction that can help learners achieve this level of function in life-long learning. The foreign language profession adopted a monolithic approach to language teaching in the audiolingual era; it is encouraging now that many language teachers see the value of varying approaches and learning activities in order to meet the needs of all the students in the classroom and help all the students build towards the longest-term goals. The authors of this article therefore endorse the balanced use of different learning activities in the classroom.

Notes

1. The authors represent current and former elected officers of the American Council of Teachers of Russian, the American Association of Teachers of Slavic & East European Languages, textbook authors for first-, second-, third-year Russian and Russian for heritage speakers, current and former directors of the Middlebury Russian School, and members of the ACTR Advanced Placement and Russnet Editorial Boards. Together they have hundreds of years' experience teaching Russian at the college and precollege levels and include professors from some of the United States' most prestigious institutions of higher education.

2. Tracy Terrell, who did have professional expertise in teaching Russian and who subscribed to many of the philosophies and approaches to teaching foreign language that Wong and Van Patten advocate, used a purely Natural Approach far more extensively in teaching Spanish and French than he did when personally teaching Greek or Russian. He once commented to one of the authors of this article that in the case of these highly inflected languages, he relied on comprehensible input only for the first six weeks of teaching, after which he "used learning to prompt acquisition [our emphasis], especially as students gained greater proficiency in the language and started meeting structures for which they had no relevant schemata" (B. L. Leaver, 1999a, p. 103).

3. We will give an example. The verb, to read, has two forms: читать *chitat'* (imperfective) and прочитать *prochitat'*(perfective). There is also another perfective form, прочесть *prochest'*, which is in free variation with прочитать *prochitat'*. The imperfective verb has present, past, and future tenses; the perfective has past and future tenses. The perfective form is telic in nature. That is, it would be used in the case in which someone would say, "I read the book," meaning that he or she had finished reading it. The imperfective, with this verb, is unmarked, and is used where the book was not finished or the finishing does not matter. It is unlikely that simply through interactions with native speakers students would ever be able to figure out what was going on because the listener often has no idea whether the aspectual usage is correct or not, so simply accepts the input. Only in such rare cases in which a student might say, "I read (perfective) *War and Peace* last night," is it likely that a native speaker would realize that the student had made an aspectual error and react with surprise, "How is that possible?" Otherwise, the student gets repetition (natural drilling) of incorrect forms.

4. These figures are based on U.S. government figures and augmented by an informal review of the teaching and interpreting profession. There is no way to know for sure what the actual number of speakers at this level is, but it is nonetheless clear that the number of subjects in the study is a significant portion

of the whole and their aggregate responses can be considered valid indicators of the whole.

5. To be fair, we must note that the cause–effect relationship is not clear here. All students who reached Level 4 in this study had been instructed through cognitive approaches. No products of communicative classrooms were found for inclusion in the study, although they were sought. Since it takes an average of 17 years to reach near-native levels, it could be that the communicative classroom products have not yet reached the level because they have not had enough time to do so. Alternatively, it could be that communicative classrooms need to be adjusted, as Soudakoff (2001) argues, to deal with accuracy issues earlier on, avoiding the kinds of fossilization of "approximation" in language output that can keep these students from reaching the highest levels.

6. This study, which examined the nature of language processing by students as they progressed from novice to superior levels, was repeated with several hundred Russian and German students at the Defense Language Institute (also unpublished research), with very similar results.

7. This survey was conducted as part of unpublished and continuing research on the issue of reaching advanced and higher levels of foreign-language proficiency, in which some of the authors of this article have participated. The study was an outgrowth of a conference, "Reading to the Four," that was held at the Foreign Service Institute and hosted by the National Foreign Language Center in spring 2001; it was initially undertaken by the National Foreign Language Center, with the assistance of the Defense Language Institute and is currently a project of the Coalition of Distinguished Language Centers, with the continuing assistance of the Defense Language Institute.

8. Ehrman and Leaver (1997, 2003) defined ectenic learners as those who are atomistic in their approaches; they "extend" (from the Greek, *ektasis*) or disassemble information, making conscious all its parts. The styles that form their profiles include abstraction, sharpening, field dependence, lack of field sensitivity, analysis, reflectivity, particularity (as opposed to global learning), sequentiality, digitality (seeing things in linear ways), and deduction. Synoptic learners (from the Greek, *synopsis*) process information as an interrelated whole, with the parts perceived at the subconscious level. The styles that form their profiles include concreteness, leveling, field independence, field sensitivity, synthesis, impulsivity, globality, randomness, analogue (metaphoric learning), and induction.

References

Brecht, R.; Davidson, D.; & Ginsberg, R. B. (1993). *Predictors of Foreign Language Gain During Study Abroad*. Washington, DC: National Foreign Language Center and the American Council of Teachers of Russian.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*(2), 195–222.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.

Ehrman, M. (2002). Understanding the learning at the superior–distinguished threshold. In B. L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 245–59). Cambridge, UK: Cambridge University Press.

Ehrman, M. & Leaver, B. L. (1997). Sorting out global and analytic functions in second language learning. Paper presented at the American Association of Applied Linguistics (AAAL), Orlando, FL.

Ehrman, M.; & Leaver, B. L. (2003). Cognitive styles in the service of language learning. *System, 31*(3), 393–415.

Gass, S. M.; & Selinker, L. (2002). *Second language acquisition: An introductory course*. Mahwah, NJ: Lawrence Erlbaum.

Higgs, T., & Clifford, R. (1982). The push toward communication. In T. Higgs (Ed.), *Curriculum, competence and the foreign language teacher* (pp. 57–59). Skokie, IL: National Textbook Co.

Kempe, V., & MacWhinney, B. The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition, 20*, 543–87.

Leaver, B. L. (1984). Twenty minutes to mastery of the Cyrillic alphabet. *Foreign Language Annals, 17*(3), 215–20.

Leaver, B. L. (1999a). Editor's note: Reevaluating proficiency issues. In B. L. Leaver (Ed.), *Twelve years of dialogue on teaching Russian* (pp. 103–5). Washington, DC: ACTR/ACCELS Publications.

Leaver, B. L. (1999b). Understanding student complaints. In B. L. Leaver (Ed.), *Twelve years of dialogue on teaching Russian* (pp. 261–70). Washington, DC: ACTR/ACCELS Publications.

Leaver, B. L. (2001). Is teaching Russian different from teaching other foreign languages? *ACTR Letter, 28*(2), 1–5.

Leaver, B. L., with Atwell, S. (2002). Preliminary qualitative findings from a study of the processes leading to the advanced professional proficiency level (ILR 4). In B. L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 260–79). Cambridge, UK: Cambridge University Press.

Leaver, E. (1999). Making the most of memory. In B.L. Leaver, I. Dubinsky, & M. Champine (Eds.), *Passport to the world: Learning to communicate in a foreign language* (pp. 87–110). San Diego, CA: San Diego State University/LARC Press.

Omaggio Hadley, A. (2003). *Teaching language in context*. Boston: Heinle & Heinle.

Shekhtman, B. (2003). *Working with advanced students*. Salinas, CA: MSI Press.

Shekhtman, B., Leaver, B. L., with Lord, N. A., Kuznetsova, E., & Ovtcharenko, E. (2002). Developing professional-level oral proficiency: The Shekhtman method of communicative teaching. In B. L. Leaver & B. Shekhtman (Ed.), *Developing professional-level language proficiency* (pp. 119–40). Cambridge, UK: Cambridge University Press.

Soudakoff, S. (2001). Developing 3+/4 reading and translation skills for presidential needs: A summary. *ACTR Letter, 28*(1), 1–3.

Coblin, W. S., & Levy, J. (2000). *Francisco Varo's grammar of the Mandarin language* (1703). Amsterdam: John Benjamins.

Wong, W., & VanPatten, B. (2003). The evidence is IN: Drills are OUT. *Foreign Language Annals, 36*(3), 403–24.

AUTHOR'S RESPONSE

Beyond Experience and Belief (or, Waiting for the Evidence): A Reply to Leaver et al.'s "Apples and Oranges"

Wynne Wong
The Ohio State University

Bill VanPatten
The University of Illinois at Chicago

Facts are stubborn things; and whatever may be our wishes, our inclinations, or the dictates of our passions, they cannot alter the state of facts and evidence.
John Adams, December 1770

We appreciate the opportunity to respond to the various authors and voices of "Apples and Oranges Are Both Fruit" (Leaver, Rifkin, & Shekhtman, 2004) for their comments on our article ("The Evidence is IN, Drills are OUT," *FLA*, 2003, 36[3], 403–23). Furthermore, we are glad that our article has sparked discussion within foreign language teaching circles, especially among those who teach what are often referred to as the less-commonly-taught languages (in this case, Russian). Our response to Leaver et al. is grounded in a critical reading of their arguments; it is not an attack on the authors or the teaching of Russian. To this end, we ask the following question: What is Leaver et al.'s empirical evidence to refute our basic claim that drills are not necessary for acquisition of any language?

In this response, we will use the terms "evidence" and "empirical evidence" interchangeably with the standard meaning in the sciences and social sciences, thus excluding anecdote, experience, and personal belief as the bases for argument. This is important because Leaver et al. erroneously state that "the application of [Wong and VanPatten's] *experiences* was fraught with error" (emphasis added). Nowhere do we rely on our experiences or beliefs to conclude that drills are not necessary. Our argument is based on the empirical evidence from a variety of sources (i.e., our research and that of many others). We will examine the empirical evidence offered by Leaver et al. as we discuss (1) the difficulty of Russian, (2) the overgeneralization of teaching one type of language to teaching other types, (3) the scope of acquisition, (4) individualization and individual differences, (5) and other points raised in Leaver et al.'s essay.

As a preamble to our response, we begin with excerpts from our article. The first excerpt summarizes our major point and conclusion:

> What we are stating in this article comes down to this: As far as acquisition is concerned, drills are simply unnecessary and at best a waste of time for the development of communicative ability. (Wong & VanPatten, p. 418)

Any counterarguments, then, must offer evidence against this particular point; that is, any claim that our conclusion is wrong must show evidence that drills are *necessary* and in particular, the drills that we discuss in our paper. Alternatively or in addition, such arguments must show that the extensive evidence we cite against drills is flawed, irrelevant, or some combination of the two. In the particular case of the nonnecessity of drills, Leaver et al. fail to do any of these requisite tasks.

To be clear about what we mean by drills, we again quote from our own article:

> We would also like to point out that our concern is in the belief that *mechanical* drills are essential for acquisition. We are less concerned about meaningful and communicative drills (à la Paulston) being viewed as essential or beneficial to acquisition, as will become clear later. Thus, in the rest of this paper, we will use the cover terms *drill* and *drills* to mean mechanical drills (p. 407, emphasis original)

This quotation is critical because any argument against our conclusion can only be an argument about mechanical drills and not the other drill types that we detailed in our review (i.e., meaningful and communicative drills). As is clear from our review of Paulston's work, mechanical drills focus on form only; meaning or any type of communication is absent. As will emerge in our discussion here, Leaver et al. appear to use the term "drill" differently and attribute conclusions to us that are unfounded. We, however, will continue to use the cover term *drill(s)* to mean *mechanical drill(s)* as in our original article.

A third excerpt from our article is also essential. Here, we made a distinction that is important for any discussion of language teaching:

> There are two aspects to learning a language for oral communicative purposes: the creation of an underlying implicit linguistic system that consists of rules, forms, and lexical items; and the development of the ability to use that system to express meaning….In short, learners must simultaneously do two things during the course of acquisition: develop a linguistic system and develop mechanisms for language production. This distinction is important for the present discussion because the utility of drills in language instruction needs to be examined from two different perspectives, namely, whether drills help to develop the underlying system and whether they are useful in promoting accuracy and fluency. (p. 404)

Here we were clearly suggesting that acquisition is more than just reaching a certain level (e.g., on a proficiency scale) and that in both SLA research and theory, the conceptualization of acquisition as multiple processes and products is a given (e.g., Archibald, 2000; R. Ellis, 1994; Gass & Selinker, 2001; Towell & Hawkins, 1994; VanPatten, 2003). In this response, it will become clear that Leaver et al. do not make these distinctions. Thus they conflate the various dimensions of acquisition into a unitary phenomenon and ignore the vast literature on second language acquisition in and out of the classroom with different languages. The result is a failure to offer a convincing argument against our position.

One more quotation from our article is crucial to our discussion here:

> By claiming that drills are not necessary and in some cases can hinder acquisition, we are not saying that no focus on form is necessary or that we are against instruction of any kind. Our position is clearly different from the position taken by Krashen (1982) and others. We are obviously advocating some kind of focus on form, given the research we have been involved in regarding [processing instruction]. (p. 417–18)

Any comments regarding our argument, then, must be careful not to confuse an argument against drills as an argument against instruction, focus on form,[1] or anything else. Our argument is carefully and purposefully centered on mechanical drills and mechanical drills alone. We believe that Leaver at al. have conflated drills, focus on form, and instruction in general in their response. Because they do not understand the scope of our position, they imbue our claims with the promotion of certain approaches to language teaching that we do not espouse.

Language Difficulty

Leaver et al. are undoubtedly concerned about the learning and teaching of Russian and argue that what applies to Spanish and French, for example, does not (or may not) apply to Russian. We state in our article that some languages may be more difficult than others to acquire for speakers of a particular first language (Wong & VanPatten, p. 416). In fact, the lengthy quotation from our article used by Leaver et al. underscores this, (e.g., we say "our point is not that instruction cannot help the learning of Russian, Japanese, or any other language. It is that the role of drills cannot change depending on language" p. 416), so we are unsure why the authors need to defend the difficulty of Russian, as though this were in dispute. More to the point, in our reading of their discussion of language difficulty, we find no argument or citation of empirical evidence that drills are necessary to learn these more difficult languages. Consider the following points that Leaver et al. make:

(1) The FSI has data that it takes longer to learn Russian than, say, French;
(2) Greater aptitude is required for classroom learning of some languages;
(3) Russian morphology is complex;
(4) The interface between semantics and syntax may pose problems in Russian.

None of these points is disputed. However, none of these points is evidence that drills are necessary. In other words, "harder/more difficult" does not equate with "drills are necessary." It is not even clear that "harder/more difficult" equates with "drills are useful." The question that needs to be researched is the following: What happens in Russian language teaching if drills are replaced by something like processing instruction?[2] That such research has not been done ought not to lead anyone to the conclusion that drills are necessary for Russian. It ought to lead to the desire for research on this topic (assuming that we all understand what processing instruction is, a point we address later). As we see it, the correct conclusion from our discussion of Russian and other more difficult languages is that drills are not necessary, but *maybe* some other kind of pedagogical intervention is.[3]

Overgeneralizations

Leaver et al. argue that the learning and teaching of different languages may require different instructional approaches and techniques for each, claiming that we overgeneralize work on Spanish and French to other languages. In particular, they cite research in which higher-level-proficiency speakers of Russian claim that instruction is useful:

> . . . 67% of the Russian learners (compared with 50% of foreign language learners in general) considered grammar explanation crucial to their success in language acquisition, 58% (compared with 21% of the overall group) wanted grammar drills even at high levels, and 58% (compared with 21%) also wanted opportunities for the deliberate practice of grammar.

As we look at these reported percentages, we search for the evidence against our fundamental claim that drills are not necessary. We do not find it. "Considering grammar explanation crucial" does not equate with "drills are necessary"; "wanting grammar drills" does not equate with "drills are necessary"; and "wanting opportunities to practice grammar" does not equate with "drills are necessary." The only possible conclusion from the study that Leaver et al. cite is this: *Students have beliefs about language learning, specifically about explicit instruction and drills*. But we cannot use learner beliefs about language learning to construct a theory of language learning and we certainly cannot use those beliefs as empirical evidence to argue for or against anything (other than that students have those beliefs). In other words, "student beliefs" do not equate with "drills are necessary."[4]

Our take on the use of these beliefs by Leaver et al. is that these authors have conflated the necessity of drills with the utility of focus on form, a conflation and confusion we clearly and explicitly set out to avoid (Wong & VanPatten, p. 417, and elsewhere in our article). The question, then, of overgeneralization of ideas or constructs may lie not with our article but with their reading of it. That is, overgeneralization in the form of extrapolation comes from them and not us.[5]

Leaver et al. also take to task the research by Kempe and MacWhinney (1998) that we cited in our article. The reason we cited this study is because learners of Russian demonstrated acquisition without drills, thus demonstrating that drills are not necessary. Leaver et al. suggest that the research was flawed because the participants may not have been representative of a larger and more varied pool and because there were no measures of production. Leaver et al. also argue that Russian students in general may have greater language aptitude than learners of German. So, do these observations contribute to countering our position that drills are not necessary? No. We believe that Leaver et al. argue these points to show why the Russian learners did better than the German learners in the study. However, the fact remains that the learners demonstrated acquisition of case marking without drills (in both language groups). This point stands independently of any other point anyone would like to make about the results of the Kempe and MacWhinney study.[6]

As for the criticism that the Kempe and MacWhinney study used no formal measure of production, we see this as evidence of Leaver et al.'s failure to view acquisition as a multifaceted phenomenon. In terms of the acquisition of an underlying mental representation of language, assessment of production is not needed (see White, 2003, for examples). If the argument is that drills are not necessary for the acquisition of an underlying system, comprehension measures are perfectly fine.

In addition to the arguments above, Leaver et al. offer an explanation about the problem of learning case markings in Russian, saying that lower-level learners demonstrate they actually do not understand case when they misinterpret OVS and OSV sentences and that possibly learners in the Kempe and MacWhinney study were processing meaning via word order (i.e., all sentences in the study must have been SVO). Two points deserve mention here. First, Kempe and MacWhinney did *not* test learners on SVO sentences, but on OVS and OSV sentences, so if the learners were correctly interpreting non-SVO sentences in Russian, they were demonstrating acquisition of case marking. Second, we find it ironic that Leaver et al. choose this particular argument to offer as an alternative explanation for the results of the study and thus to bolster their argument that drills are necessary. It is precisely OVS and OV sentences in Spanish that sparked the original research in processing instruction (VanPatten & Cadierno, 1993). We *know* that learners rely on lexicon and canonical word order to get meaning. The aim of processing instruction in such instances is to push learners to rely on form to retrieve meaning. Leaver et al., then, indirectly offer the possibility (and we think it's more than just possibility) that processing instruction would be quite useful for the acquisition of case marking—in any language.

Our Take on the Full Scope of Language Acquisition

There are a number of points that Leaver et al. make in discussing the "scope of language acquisition." It is difficult to deal with any one underlying theme here, so we will take the various points one by one.

First, Leaver et al. argue that there is a difference between what is required at lower levels and what is needed at upper levels of language ability. We do not deny this, but our conclusions are not the same as those arrived at by Leaver et al. They state:

> …if our students are to attain any more sophisticated control of Russian and to use it in any more challenging contexts, they need to gain control of

> Russian grammar. Our experience teaches us that this control is forged in the classroom and at early levels—otherwise, the duration needed to acquire high-level language skills is considerably greater.

Examining this point, it appears once again that they are confusing drills with focus on form. Once again, for the record, we do not argue against focus on form in the classroom. What we argue is that there is no evidence that drills are necessary. Interestingly, Leaver et al. appear to agree with us and imply that drills are not necessary when they suggest that higher levels would be reached anyway, albeit more slowly. In short, "taking longer" does not equate with "drills are necessary."

Leaver et al. state "the issue, as many have said, is not *whether* to teach grammar but rather *how* to teach it" We were pleased to see this point made because this is exactly the point made by VanPatten and Cadierno in the initial processing instruction study in 1993 and it is precisely the point that underlies our argument; that is, we are concerned about the *how* of focus on form. We are particularly concerned about the relationship between instruction and underlying processes in acquisition (see Doughty, 2003, as well as the recent volume edited by VanPatten, 2004). But the issue of *how* is dashed later when Leaver et al. refer to an informal study of Slavic language teachers working at high levels of proficiency in which "nearly all of the teachers taught grammar overtly and very often used drills" We do not question the results of this study (i.e., that teachers used drills) but we do not see how this is evidence that drills are necessary—or what it has to do with the issue of "how" instruction in grammar takes place. "Using drills often" does not equate with "drills are necessary," nor does it equate with "research on the *how* of grammar teaching" (again, assuming our use of the term drill[s]).

Leaver et al. also discuss the problem of learning only through comprehensible input and state that "Wong and VanPatten promoted the ultimate fallacy when they assumed, without evidence or attempted research, that contextualized, authentic input will be comprehensible to students of Russian."[7] We scrutinized our paper for evidence of this assumption; that is, where did we imply that this is what we believe? We cannot find any reference, statement, or implication leading to this assumption. We can only assume (now) that Leaver et al. must think that because we see no evidence for the necessity of drills that we must believe that comprehensible input takes care of everything. We remind readers that we have spent the last 12 years researching the effects of processing instruction, which—to be sure—is a focus on form and not some kind of "contextualized, authentic input."[8] In fact, it was research on input processing and problems that learners have with authentic and contextualized input that led to processing instruction. "The nonnecessity of drills" does not equate with "learners only need comprehensible input." We remind the reader of our very explicit statement, repeated here:

> By claiming that drills are not necessary and in some cases can hinder acquisition, we are not saying that no focus on form is necessary or that we are against instruction of any kind. Our position is clearly different from the position taken by Krashen (1982) and others. We are obviously advocating some kind of focus on form, given the research we have been involved in regarding [processing instruction]. (p. 417–18)

Leaver et al. discuss "drills that work." There is no evidence in this discussion that drills are necessary—whether they "work" or not. Moreover, they do not offer any definition (operational for research) of what it means for something "to work." The point here is that "describing a good drill" does not equate with "drills are necessary."

To bolster their argument further, Leaver et al. refer to research in first language acquisition and associative memory as well as the research of Robert DeKeyser. Unfortunately for their claim, there is no evidence that drilling exists in first language acquisition as they state. In fact, the opposite exists: Attempting to get young children to repeat something correctly almost always results in failure (see, for example, McNeill, 1966; Pinker, 1994, pp. 279–83). Their claim that "Children's songs, nursery rhymes, and games, teach first language through what is essentially drill" is not true. First, there is not a single first language acquisition researcher that we know of who would suggest that language is "taught" to children, let alone through some steady diet of songs and rhymes that act as drills (see, for example, Crain & Lillo-Martin, 1999; Pinker, 1994; Slobin, 1978, 1985; and all the research since 1960). Just as important, however, is that Leaver et al. get it backwards; children don't engage in songs and rhymes to get language, *they engage in songs and rhymes because they have the necessary language to do so*.[9]

In terms of memory research, current cognitive psychology does not argue that repetition or drills are necessary for learning either anything or everything; it argues that repeated exposures or behaviors tied to deliberate intentions lead to learning. In the case of language, this means communication. Thus, if meaning is absent from any kind of language practice, it would—in theory—be useless. This is part of our argument against drills. Moreover, research on drilling and automaticity (DeKeyser's research) does not mean that drills are necessary. Any low-level language function can be drilled to a point where people may look good at what they do. But again the crucial element here is necessity. Is it possible that there are other routes to automaticity? Cognitive psychologists would surely think so; otherwise no human would learn anything except through drilling—and this clearly is

not the case. In our article, we were very careful to discuss this matter:

> First, in a classic skill theory scenario such as learning chess, the novice chess player does not sit around drilling moves. *All chess moves occur in the context of playing a game.* Thus, whatever chess players learn about playing and however they automatize their moves happens as a result of playing chess, not practicing chess. (p. 416)
> . . .
> [Skill theory] would suggest that language learning is like learning tennis, typing, or chess, and would ignore language learners bringing to the task of acquisition certain mechanisms that are specific to language processing (see Schwartz, 1998, for an excellent discussion of the special nature of language in SLA)….Currently, skill theory has no means of explaining how it is that all successful learners of a second language come to know more than they were ever taught or practiced and that this knowledge is of a particular kind and is not based on what skill theorists call "analogy."(p. 416)

Segalowitz (2003) offered support for our position in his review of automaticity in second language acquisition. He stated:

> The challenge then is to incorporate activities that promote automaticity into the language learning situation in a manner that respects *transfer-appropriate processing* and other positive features of communicative practices…[others] have shown that in complex skill-learning situations the transfer of automatized skills depends on *the psychological similarity of the learning and transfer contexts*. (p. 402, emphasis added)

What Segalowitz meant, essentially, is that language practice that does not tie into communication or some kind of meaning is suspect. In our terms, we would say that drills are questionable as a technique because they are devoid of meaning and communicative intent. Even DeKeyser (whom Leaver et al. cite) asserted that drills (as we define them) are not the appropriate means by which automaticity can develop. He stated:

> …what are mechanical drills for? *They are not what is needed to proceduralize [grammatical] knowledge… because they do not engage the learner in the target behavior of conveying meaning through language. Instead, they provide practice in a very peculiar behavior, a "language-like behavior*," which consists of linking forms with other forms, of shuffling forms around, according to a pattern held in working memory, without ever linking those forms with meaning, that is, without the student ever engaging in the target behavior of using language. (DeKeyser, 1998, p. 53, emphasis added)

In another publication, he stated:

> [A]mong commonly used classroom techniques, communicative drills stand out as an activity that is likely to lead to production compilation, because of the repeated execution of the target routine while the relevant declarative knowledge is highly active. (DeKeyser, 2001, p. 150)

DeKeyser's position is consistent with the idea that if practice is necessary, it must be practice that is transfer-appropriate, that is, learners must be engaged in the expression of meaning. Mechanical drills—the only drills that are the object of our argument—do not do this.

Although they mention associative memory, Leaver et al. do not discuss just what associative memory has to do with language acquisition. Casual references to cognitive research (without citation from that research) do not constitute an argument against the points we make above; nor do they constitute an argument against our conclusion that drills are not necessary.[10]

A final point to consider is related to Leaver et al's statements about fossilization (in the section about scope of acquisition). Leaver et al. appeal to the argument that learning only through comprehensible input causes many learners to fossilize at lower levels; they cite one source: Higgs and Clifford (1982). Aside from the problems in citing this particular source as their only source (see, for example, the discussion in VanPatten, 1988, and Long, 2003), there is a problem with relying on fossilization and American soldiers abroad as evidence for the necessity of drills. The first is, of course, that not everyone fossilizes in untutored contexts. Second, the construct of fossilization is chock full of problems. As Long (2003) correctly observed, the evidence for fossilization is largely sketchy. He stated that the common problems in demonstrating fossilization include but are not limited to "(i) assuming, not demonstrating fossilization (or stabilization); (ii) selecting inappropriate learners for study; (iii) basing findings on insufficient data; and (iv) using inadequate analyses" (p. 492). In his subsequent review of the research, Long found little clear support that fossilization is a viable theoretical construct. Even if fossilization does exist, Long argues, the explanations for it that have been offered must meet the criterion of testability under empirical means. This is not the case; all explanations of fossilization are speculative, including the one offered by Leaver et al.[11]

Even if we granted that non-classroom learners fossilize at lower levels, the question is this: What does this have to do with the nonnecessity of drills? *Our argument is not about the outcomes of classroom learners versus non-classroom learners but about the outcomes of classroom learn-*

ers who get drills compared with other classroom learners who don't get drills. Invoking the image of American soldiers in Germany eking out a pidginized variety of the language that surrounds them is irrelevant to the claim that when it comes to classroom practices, drills are not necessary. The correct comparison would involve empirical studies in which the outcomes of classrooms where drills are removed and/or are replaced by something else are compared with classrooms where drills are used. This is precisely the evidence we offer in the studies we cite on processing instruction (as well as Savignon's study), and the evidence is irrefutable: Drills are not necessary.

On Individualization

Leaver et al. state the following: "Wong and VanPatten made the same mistake that many non-polyglots make: They erroneously assumed that all languages are learned in the same way" Our assumption was actually this: *All internal mechanisms for language acquisition are the same for everyone; they do not change from context to context or from language to language*. This assumption is not unique to us; it exists at the core of SLA theory and research and is clear after 40 years of empirical research in and out of classrooms (see, for example, the various articles in Doughty & Long, 2003, as well as R. Ellis, 1994; Hawkins, 2001; Pienemann, 1998; Schwartz, 1998, White, 2003, and many others). Leaver et al. seem to equate individual preferences for drills with "drills are necessary" for some people. What people do externally or what they like to do cannot be equated with language processing, the role of Universal Grammar in acquisition, or any other language specific processing (linguistic or cognitive) that lies at the core of acquisition. In short, we do believe that everyone learns all languages the same way as far as processing and mental representation are concerned.

> Leaver et al. also make the following statement:
> Wong and VanPatten, in arguing for one approach (theirs) to fit all, excluded a key component to making language study accessible to all students: adaptation for learning style. Some learners learn much more readily in a more structured approach than do others.

What is revealed by Leaver et al.'s statement? Two conclusions: "drills aren't necessary" equates with "unstructured classrooms"; "drills aren't necessary" equates with "some learners are left out." Both of these conclusions are erroneous. First, taking drills away and replacing them with something else does not result in an unstructured or even less structured approach. A careful examination of the sample French lesson with processing instruction that we provided at the end of our article (Appendix B, p. 421–23) would show this. A typical processing instruction lesson is almost identical in structure to a traditional drill-oriented lesson in that: (1) there is explicit teaching about how something works; and (2) there is subsequent practice using individual activities. To suggest, then, that we are leaving learners to their own devices or that we eschew structured presentations and activities is incorrect.

> Later Leaver et al. state:
> [S]uch ectenic styles as deduction, particularity, analysis, sequentiality, and sharpening hinder the learner from comprehending authentic input without assistance with the input and with learning strategies that cope with such input.

We agree that there could be a problem, which is why we offer processing instruction as an alternative to drills and have been researching its effects for over a decade. Processing instruction offers exactly what Leaver et al. want (as do other focus on form techniques): assistance with the input.

Some Remaining Issues

As we reflect on what Leaver et al. have written, we find that we do not disagree with the following:

- focus on form can be useful, if not necessary, in some cases for some structures for some learners

What we disagree with is that the focus on form needs to be a drill. We are concerned that perhaps there is a misunderstanding of what it would mean to not have drills. Does not having drills mean "throwing lots of input at learners"? This seems to be the conclusion of Leaver et al., and we are struck by their repeated references to input and comprehensible input.

> The amount of comprehensible input to acquire this system is far beyond that which one can get in a classroom.
> . . .
> Acquiring Russian is not, then, a matter of simply more and more input.
> . . .
> Wong and VanPatten promote the ultimate fallacy when they assume, without evidence or attempted research, that contextualized, authentic input will be comprehensible to students of Russian.
> . . .
> [Soudakoff] attributed the lack of preparation for higher-level study on teaching methods such as those proposed by Wong and VanPatten at lower levels.

We can only speculate about these references to comprehensible input and methods. We believe that Leaver et al. have misread or misunderstood the role of input in language acquisition as we laid it out in our article. Our point in discussing input was that input is necessary for acquisition. *At no time did we say it is sufficient*. This position is the standard of instructed SLA theory and research regardless

of the particulars of one's theory (e.g., Universal Grammar [White, 2003 and scores of others]; connectionism and other frequency-based cognitive theories [N. Ellis, 2003 and many others]; general nativism [O'Grady, 2003]).[12] It appears that Leaver et al. have confused necessity with sufficiency and assume we mean the latter. This is simply wrong.

Our speculations seem justified when one notices the various references and allusions to our "proposed methods." Just what are these proposed methods? Nowhere in our paper did we offer any advice or suggestions about methods or approaches. We did one thing and one thing only: argue that drills are not necessary. "Arguing that drills are not necessary" does not equate with "proposing methods." Again, we can only speculate on Leaver et al.'s motivations for this particular claim. In addition to their misreading of our position (and, again, that of the entire field of SLA theory and research) on the role of input, we are guessing that they must be misreading the nature of processing instruction. We say this because research on processing instruction forms a large part of our evidential base for the nonnecessity of drills. What do Leaver et al. believe processing instruction to be?

We know that processing instruction has been misinterpreted by a number of professionals and we have repeatedly attempted to point out the misinterpretations and the assumptions that underlie them (e.g., Sanz &VanPatten, 1998; VanPatten, 2002a, 2002b, 2002c; Wong, 2004, as well as the description in Wong & VanPatten, p. 410–11). For the reader's benefit, we will succinctly state what processing instruction is: Processing instruction is an interventionist focus on form that attempts to correct processing problems so that learners end up with better intake data for their developing linguistic systems (see, also, Lee & VanPatten, 1995, 2003; VanPatten, 1996). There is no way that a careful reading of what we argue can be construed as simply throwing more input at learners, for any language—even the "easy" ones. Nor should a careful reading of our article result in the conclusion that we advocate extreme or even moderate synoptic approaches (to borrow Leaver et al.'s term).[13]

We are also troubled by an off-handed remark made by Leaver et al.: "Knee-jerk reactions against drilling in any form are not helpful for the profession" We take issue with any implication that our position is not a carefully considered one. The long list of theoretical and empirical sources that we cited shows that we were not engaging in "knee-jerk reactions." Four decades of empirical research on second languages learned around the world in various contexts, coupled with five decades of research on first language acquisition in languages around the world in various contexts, ought not to be so neatly swept under the rug because people wish to cling to beliefs about teaching practices and classroom learning behaviors. Our position is that instruction ought to be informed by both theory and research and it is these foundations of scholarship—rather than our own experiences, beliefs, and opinions—upon which we drew.

Conclusions

So, what are the points we have made in this response to Leaver et al.'s "Apples and Oranges"? First, we have shown how Leaver et al. miss the point of our article and, by most standards, have set up a straw man they then set out to destroy. Second, we have shown that Leaver et al. are short on empirical evidence that drills are necessary (for Russian or any language) and they are short on empirical evidence that drills are even useful. Third, we have shown that Leaver et al. confuse and conflate constructs, thus arriving at wrong conclusions. Fourth, we have shown that Leaver et al. misconstrue and distort some of the most clearly documented research on first and second language acquisition to bolster their argument.

We have also underscored that our conclusion ought *not* to lead to incorrect extrapolations such as:

- the nonnecessity of drills means abandoning any focus on grammar;
- the nonnecessity of drills results in an unstructured classroom;
- the nonnecessity of drills means an exclusive focus on comprehensible input and the necessity of input means that a focus on grammar is not useful or necessary;
- the nonnecessity of drills means that different languages may offer different problems;
- the nonnecessity of drills means that all learners are created equal in terms of intelligence and learning styles.

We do not conclude, imply, or assume any of the above points. The nonnecessity of drills—and again, to be absolutely clear on the matter, *mechanical drills*—simply means that drills aren't necessary. We cannot put it any clearer than that. As we have repeatedly said, although we are advocates for processing instruction as a focus on form, this is by no means the only choice left to instructors if they choose to drop drills from their repertoire of techniques (see, e.g., Doughty & Williams, 1998, as well as Wong, in press). As Lightbown (2004) says in a commentary on processing instruction, "VanPatten and Wong have been very clear that they do not see PI as the best or only approach to teaching all language features. They have also left no doubt that PI is not proposed as the basis for taking learners all the way to spontaneous, accurate, automatized production of any language features" (p. 73).

What we see in Leaver et al.'s response to our article is a fervent *belief* that drills must be *useful* because "that is what we do," "students like them," "teachers like them," "there are good drills," "street learners fossilize," "there are

different kinds of learners," "our experience is," and because "Russian is different/harder." To be sure, our intention was and is not to take any teaching tools away from anyone; after all, drilling never killed a student as far as we know—and drills have not caused students such irreparable harm that an educational institution was taken to civil court. Moreover, our intention was not to deny that some languages pose unique learning problems to some populations compared with other languages. Our intention was and is only to show that drills devoid of meaning are not necessary and this we have done. Leaver et al., in turn, have failed to supply any convincing argument or evidence to refute this claim. Much more than this failure, though, they have made extrapolations that associate us with contexts and belief about language and language teaching to which we do not adhere.

Acknowledgments

We would like to thank the following scholars and researchers who carefully read our response and offered feedback: Janice Aski, Joe Barcroft, Gregory B. Keating, James F. Lee, Michael J. Leeser, Frank Morris, Cristina Sanz, Mary Ellen Scullen, and Jessica Williams. The usual caveat of responsibility (ours alone) applies.

Notes

1. Focus on form, in case the term is unfamiliar to the reader, is any pedagogical device that draws learner attention to formal properties of language during interpretation, expression or negotiation of meaning (see the various articles in Doughty & Williams, 1998).

2. One could replace "processing instruction" with any other focus on form and the question remains. However, as we state here and in "Drills are OUT," processing instruction is not the only available alternative to drills.

3. We note that Leaver et al. make repeated reference to "considerable empirical research" on the acquisition of Russian and a "large body that refutes [Wong and VanPatten's] claims." No such empirical evidence is cited in their response. In addition, they make reference to unpublished research. It is hardly fair to hold the professional community accountable for unavailable research. They do cite tangential evidence that suggests (somewhat) that a focus on form is useful for acquisition, but they do not cite research on the actual acquisition of Russian or research showing drills are either necessary or are the causative factor in acquisition. What is more, we cite here from Leaver and Shekhtman (2002) as they discuss teaching and learning at advanced levels of language ability: "The teaching of Superior-level skills is virgin territory…Therefore, there are more questions than answers, more theory than practice, and *more anecdote than research in this area*" (p. 33, emphasis added).

4. We do not discount that learners may have *insights* about acquisition, but we distinguish insights and well thought out reflection from beliefs. In any event, even insights and reflections are not evidence; they provide fodder for the development of hypotheses that can subsequently be tested empirically. One "empirical" study (Leaver & Atwell, 2002) provided preliminary findings of a study that used interviews of superior-level students about their experiences and perceptions about learning at that level. The use of interviews is a respectable practice, to be sure, but cannot be the basis of evidence for the present arguments. More important, in that study learners did not report that drills were necessary or useful. In the list of specific activities that the authors suggested were useful, all were clearly meaning or communicatively-oriented activities (see Leaver & Atwell, 2002, p. 270).

5. We add here that Leaver et al. repeatedly refer to study abroad as a component of achieving superior-level proficiency in Russian. This is true for all languages, however, and is hardly evidence that drills are necessary (see Carroll, 1967).

6. Leaver et al. unfairly claim that we uses this article to support our position when Kempe and MacWhinney's purpose was to demonstrate something else. To cite ourselves: "*This is not the point of [Kempe and MacWhinney's] study*, but we cite it here as an example of research on one of the 'more difficult' languages to illustrate that drills are not necessary" (Wong & VanPatten, p. 416, emphasis added). In short, regardless of the purpose of the Kempe and MacWhinney study, learners showed evidence of acquisition without drills. This is a fact.

7. We found the paragraph that begins with this sentence to be particularly difficult to understand, as Leaver et al. also claim that "speakers of English do not and cannot have the requisite schemata for comprehending the nature of the input, other than perhaps superficially, and even superficial understanding is not always possible." How is this an argument in support of drills in language instruction? Wouldn't the implication be a meaning-based approach to dealing with comprehension?

8. In other research, Wong has also investigated what is called text enhancement (see Wong, 2003).

9. Leaver et al. also state, "Classroom instruction in the native language for those parts of grammar that are not part of daily living are typically overt." This statement is difficult to evaluate for the following reasons. First, by the time children engage in any overt instruction about their native language, they already have language. Acquisitionists concur that for most children, language is in place by the age of five—for all languages, even the "difficult ones." Second, instruction is normally about prescriptive issues (e.g., "don't dangle your participles", "*that* versus *which*") and not about what is grammatical/ungrammatical in a linguistic sense (e.g., dangling participles are grammatical or we wouldn't be able to dangle them, *that* and *which* are often interchangeable or we wouldn't be able to do so). So, native speakers' underlying mental representations of language are largely unaffected by attempts to alter normative matters related to "good language use." Thus, what any of this has to do with second language learning is not clear unless one wants to argue that all instruction should wait until second language learners already have a grammar and a linguistic system, a point we think is not what the authors intend.

10. We add for the record here that Anderson's skill theory (e.g., Anderson, 2000) has little to say about language acquisition. His final chapter on applications to education focused on "skill" related to high level conceptual learning, such as math and skills in which a language is already in place, namely, reading in the first language.

11. What is more, Leaver et al. undermine their own take on fossilization in citing Soudakoff (2001), when they say that he

"has found fossilization to be a hindrance to success at higher levels, one that requires a great amount of repetitive translation (a form of drilling) to overcome" How can a system be fossilized if it can be overcome? (And for the record, translation is a meaning-based activity and not a mechanical one.)

12. It is also the standard in Gass and Selinker (2001), which Leaver et al. cite at one point to refute a different claim.

13. To be sure, we do advocate exposure to comprehensible input from the first day of language learning. What we do not advocate is that comprehensible input is enough.

References

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. 2nd ed. New York: John Wiley and Sons.

Archibald, J. (Ed.). (2000). *Second language acquisition and linguistic theory*. Oxford, UK: Blackwell.

Carroll, J. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals, 1*, 131–51.

Crain, S. & Lillo-Martin, D. (1999). *An introduction to linguistic theory and language acquisition*. Oxford, UK: Blackwell.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 195–221.

DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. J. Doughty & J. Williams, Eds., *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 124–51). Cambridge, UK: Cambridge University Press.

Doughty, C. J. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256–310). Oxford, UK: Blackwell.

Doughty, C. J. & Long, M. H. (Eds.) (2003). *The handbook of second language acquisition*. Oxford, UK: Blackwell.

Doughty, C. J. and Williams, J. (Eds.) (1998). *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.

Ellis, N. (2003). Constructions, chunking, and connectionism: the emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63-103). Oxford, UK: Blackwell.

Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.

Gass, S.M. & Selinker, L. (2001). *Second language acquisition: An introductory course*. Mahwah, NJ: Lawrence Erlbaum.

Hawkins, R. (2001). *Second language syntax: A generative introduction*. Oxford, UK: Blackwell.

Higgs, T. V. & Clifford, R. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence and the foreign language teacher* (pp. 57–79). Skokie, IL: National Textbook Co.

Kempe, V. & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition, 20*, 543–87.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon.

Leaver, B.L., & Atwell, S. (2002). Preliminary qualitative findings from a study on the processes leading to the Advanced Professional Proficiency Level (ILR 4). In B.L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 260–79). Cambridge, UK: Cambridge University Press.

Leaver, B. L, Rifkin, B., & Shekhtman, B. (2004). Apples and oranges are both fruit, but they don't taste the same: A response to Wynne Wong and Bill VanPatten. *Foreign Language Annals, 37, 1, 125–132*.

Leaver, B. L., & Shekhtman, B. (2002). Principles and practice in teaching Superior-level language skills: Not just more of the same. In B.L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 3–33). Cambridge, UK: Cambridge University Press.

Lee, J. F., & VanPatten, B. (1995). *Making communicative language teaching happen*. 1st ed. New York: McGraw-Hill.

Lee, J. F., & VanPatten, B. (2003). *Making communicative language teaching happen*. 2nd ed. New York: McGraw-Hill.

Lightbown, P. M. (2004). Commentary: What to teach? How to teach? In B. VanPatten (pp. 65–78). Mahwah, NJ: Lawrence Erlbaum.

Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford, UK: Blackwell.

McNeill, D. (1966). Developmental psycholinguistics. In F. Smith & G.A. Miller (Eds.), *The genesis of language: a psycholinguistic approach* (pp. 15–84). Cambridge, MA: MIT Press.

O'Grady, W. (2003). The radical middle: Bativism without Universal Grammar. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language* (pp. 43–62).Oxford, UK: Blackwell.

Pienemann, M. (1998). *Language processing and second language development: Processability theory.* Philadelphia: John Benjamins.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: Harper Collins.

Sanz, C., & VanPatten, B. (1998). On input processing, processing instruction, and the nature of replication tasks: a response to M. Rafael Salaberry. *The Canadian Modern Language Review, 54*, 263–73.

Schwartz, B. (1998). The second language instinct. *Lingua, 106*, 133–60.

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Oxford, UK: Blackwell.

Slobin, D. I. 1978. *Psycholinguistics*. Glenview, IL: Scott Foresman and Co.

Slobin, D. I. (Ed.) (1985). *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Soudakoff, S. (2001). Developing 3+/4 reading and translation skills for presidential needs: a summary. *ACTR Letter, 28*, 1–3.

Towell, R., & Hawkins, R. (1994) *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.

VanPatten, B. (1988). How juries get hung: Problems with the evidence for the effects of a focus on form. *Language Learning, 38*, 243–60.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Hillsdale, NJ: Ablex.

VanPatten B. (2002a). Processing instruction: an update. *Language Learning, 52*, 805–23.

VanPatten, B. (2002b). Processing the content of input processing and processing instruction research: A response to DeKeyser, Salaberry, Robinson and Harrington. *Language Learning, 52*, 825–31.

VanPatten, B. (2002c). Processing instruction, prior awareness, and the nature of second language acquisition: A (partial) response to Batstone. *Language Awareness, 11*, 240–58.

VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. New York: McGraw-Hill.

VanPatten, B. (Ed.) (2004). *Processing instruction: Theory, research, and commentary*. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225–43.

Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*, 109–32.

Wong, W. (2004). The nature of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 33–63). Mahwah, NJ: Lawrence Erlbaum.

Wong, W. (in press). *Input enhancement: From theory and research to the classroom*. New York: McGraw-Hill.

Wong, W. & VanPatten, B. (2003). The evidence is IN: drills are OUT. *Foreign Language Annals, 36*, 403–23.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge, UK: Cambridge University Press.