Lang. Teach. (2013), **46.4**, 511–517 © Cambridge University Press 2012
doi:10.1017/S0261444811000620 First published online 25 January 2012

Plenary speech

Is the Second Language Acquisition discipline disintegrating?

Jan H. Hulstijn Amsterdam Center for Language and Communication (ACLC), University of Amsterdam
j.h.hulstijn@uva.nl

After characterizing the study of second language acquisition (SLA) from three viewpoints, I try to answer the question, raised by DeKeyser (2010), of whether the SLA field is disintegrating. In answering this question, I first propose a distinction between SLA as the relatively fundamental academic discipline and SLA as the relatively applied field of language education. Instead of portraying the field in terms of quantitative or laboratory studies on the one hand, and qualitative or anthropological studies on the other, I will look at SLA in terms of theories that differ in their empirical basis. All scientific disciplines must create room for ideas or theories that do not yet lend themselves to empirical testing, but for a discipline to develop fruitfully it is crucial that nonempirical ideas do not outnumber the empirical. The fact that the number of empirical SLA theories is large is not in itself a problem: through the practices of rational 'normal science' (Kuhn 1962), the best theories (in terms of coherence, testability and scope) will rightfully come out on top.

1. Three ways of characterizing SLA as an academic discipline

There are numerous ways of characterizing an academic discipline. In this presentation I make three attempts to portray the study of SLA, that is, SLA as a scientific discipline. I will therefore not be talking about applied linguistics as the study of language teaching or language assessment.

One way of characterizing SLA is to look at the major themes covered in textbooks. The first cohort of SLA textbooks appeared in the 1990s. If we take Corder (1967) as the first major SLA publication, then Rod Ellis's (1994) monumental textbook covered roughly the first 25 years of SLA. In Ellis (1994) the major topics are learner language (including error analysis, acquisition orders and variability), external factors (including social factors and input interaction), internal factors (including transfer, implicit and explicit learning, Monitor Theory and the Competition Model), individual differences (including age, aptitude and learning strategies) and classroom SLA (including interaction and formal instruction).

Revised version of part of an invited keynote presentation given at the 30th annual *Second Language Research Forum* (SLRF), held at Iowa State University, 15 October, 2011.

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

When we look at a more recent SLA textbook, published 15 years later (Ortega 2009), we note some shifts (see also Mitchell & Myles 1998, Chapter 9; Spada 2010). Ortega gives central place to the age question, Monitor Theory is less saliently present, and prominence is given to bilingualism and cognitive neuroscience, as well as to various social perspectives on SLA (e.g. socio-cultural theory, conversational analysis, language socialization theory and identity theory). In general, the study of SLA has developed from a mainly linguistic discipline to one with a wider scope, including socio-psychological, neuro-cognitive and social perspectives.

Another way of characterizing a field of inquiry is to examine how and to what extent its major theories address and explain certain important observations (listed below). This is the approach adopted by Towell & Hawkins (1994) and later by VanPatten & Williams (2007a). VanPatten & Williams invited nine scholars to present their SLA theory and demonstrate to what extent their theory can account for these observations. In the final chapter of this edited volume, Ortega magnificently accomplishes the difficult task of evaluating and comparing the theories on a number of dimensions. The ten observations listed by Van Patten & Williams (2007b: 9–12) are the following:

1. Exposure to input is necessary for SLA.
2. A good deal of SLA happens incidentally.
3. Learners come to know more than they have been exposed to in the input.
4. Learners' output (speech) often follows predictable paths with predictable stages in the acquisition of a given structure.
5. SLA is variable in its outcome.
6. SLA is variable across linguistic subsystems.
7. There are limits on the effects of frequency on SLA.
8. There are limits on the effects of a learner's first language on SLA.
9. There are limits on the effects of instruction on SLA.
10. There are limits on the effects of output (learner production) on language acquisition.

Now let us look at this young discipline, SLA, from a third viewpoint, namely critical rationalism (Popper 1959; see also Jordan 2004). According to Popper, empiricists (in the tradition of Bacon) and rationalists (in the tradition of Descartes) are both far too preoccupied with how to arrive at certainty, that is, too concerned with attaining valid and reliable knowledge about phenomena. Popper advises us to start with 'problems' or 'puzzling phenomena', a good example of which in the field of SLA is the following: it appears that an L2 learner who starts to acquire an L2 as a young child (early onset), will ultimately attain a much higher level of L2 proficiency than one who does so as an adolescent or adult (late onset) (note that this puzzle does not appear in the VanPatten & Williams list!). The next step in the scientific process is to try to construct a theory that can explain the puzzling phenomenon. Briefly stated, a theory is a coherent explanation of a problematic phenomenon. The third step is to derive hypotheses from the theory, and the fourth is to empirically test these hypotheses. Hypotheses make predictions about phenomena and are formulated in a way that allows empirical testing, that is, they allow the operationalization of abstract theoretical constructs. Ideally, the findings of hypothesis-driven empirical research will lead to (i) support for the theory, (ii) no support for the theory and thus to its rejection or alteration, (iii) a revision of our idea of what the phenomena are, or a combination of these

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

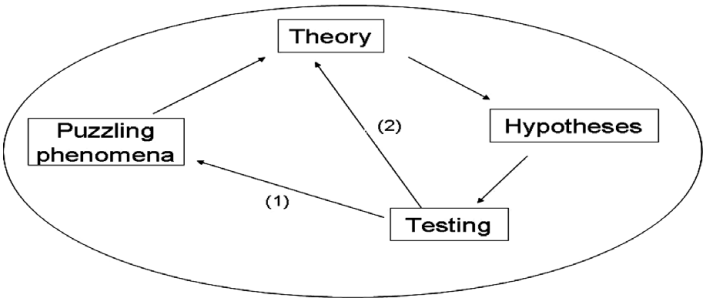


Figure 1 The scientific cycle

options, such as both (ii) and (iii). All this should be seen cyclically: after a reformulation of what the puzzling phenomena are and after a revision or replacement of the theory (arrows 1 and 2 in Figure 1), hypotheses should again be formulated and tested, which implies and includes replication. Note that, under this view, there are no rock-solid factual phenomena; there are only theory-sensitive phenomena. We improve on the approach of VanPatten & Williams, outlined above, by not stating 'observations' as established facts, but as 'provisional or assumed facts'.

When we look again at the ten observations in VanPatten & Williams (2007a), we can see that many of them are formulated from a certain theoretical stance. For instance, with respect to the first observation ('Exposure to input is necessary for SLA'), a lay person, not familiar with the generative school in linguistics, would regard this observation as trivially true rather than puzzling. And indeed, a closer look at the ten observations shows that many refer to phenomena that a generative theory of SLA aims to explain, demonstrate or verify. In other words, they are NOT trivial within this theoretical framework. Thus, to do justice to the PRESUMED factual nature of the observations, it would perhaps be better to reformulate them, using the following pattern: 'If such-and-such were facts, how would we explain them?'. This would make clearer the fact that it is almost impossible to make observations without the use of theoretical lenses.

2. Puzzling phenomena in SLA

Of course, each SLA researcher has her or his list of the most important puzzling phenomena to be explained. Here are my current favorites, most of which are interrelated.

1. *Age*. If, in SLA, it is indeed the case that young beginners ultimately attain higher levels of L2 proficiency than older beginners, how can we explain this phenomenon?
2. *Differences in learning outcomes*. If there are non-trivial differences in SLA outcomes (e.g. L2 learners of roughly the same age who have been exposed to roughly the same amount and type of L2 input over roughly the same length of time who attain different levels of L2 proficiency), how can we explain these differences?

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

3. *Acquisition orders*. If all L2 learners proceed through certain stages of L2 development, how can we explain this? What possible acquisition paths are there?
4. *Output variability within L2 learners*. If it is the case that most L2 learners, at any given point in the acquisition of an L2, deliver variable output, producing correct utterances one moment and incorrect utterances the next (with respect to the same feature of the L2), how can we explain this variability?
5. *Implicit and explicit instruction and learning*. If we assume that implicit and/or explicit instruction does not positively affect the acquisition of some L2 elements while they do affect the acquisition of other elements, how can we explain this phenomenon?
6. *Bilingualism*. If it is indeed (im)possible to control two or more languages to the same extent, how can we account for this (im)possibility?

However, SLA puzzles, in my view, are best seen as falling under the umbrella of bigger puzzles in the study of language, such as the following:

1. *Form-meaning mappings and ambiguity*. It appears that there is no one-to-one mapping between form and meaning in human languages and that utterance ambiguity forms an inherent (and necessary) feature of verbal communication. How do we best account for this inherent ambiguity?
2. *Universals*. To what extent are there universals in languages, and how do we account for the (non-)existence of universals? What are (im)possible languages?
3. *Evolution*. How do we explain the evolution of language in humans? What are (im)possible language changes?
4. *Learnability*. How can we account for the learnability of language, if it is the case that children receive input too insufficient, too superficial and too noisy to acquire the grammar of the language (known in the literature as the 'poverty of the stimulus' issue, Chomsky 1980)? Can one best account for the learnability of language by postulating a language faculty, that is, a language-specific module in human cognition, as the generative school does? Or should the best explanation be found mainly in terms of associative mechanisms, as is claimed by the usage-based schools of emergentism, connectionism, competition model, and construction grammar?
5. *Language and the brain*. How is knowledge and the use of language represented and processed in the brain? What sorts of representation are possible?

In other words, SLA, in my view, is not a separate, independent scientific discipline, or is so only in a limited sense.

3. Is the SLA discipline disintegrating?

As the preceding sections illustrate, the SLA field is characterized by a wide variety of issues and theoretical perspectives. Is this a bad thing? Are there signs of disintegration? Recently, DeKeyser (2010: 247) expressed his concern about the possible disintegration of the field of what he calls 'applied linguistics' as follows:

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

> . . . it seems to me that a bifurcation is taking place in the field between, on the one hand, ever more tightly controlled psycholinguistic experiments and ever more sophisticated statistical analyses and, on the other hand, qualitative research that uses neither experimental treatments nor inferential statistics. Regardless of one's individual preferences for one or the other, one cannot fail to observe that if this trend continues, second language acquisition (SLA) research will be absorbed completely into psycholinguistics or cognitive psychology, on the one hand, and anthropology or sociology, on the other hand. Not only would this be worrisome from the point of view of those who have a vested interest in the field of SLA, it would also be regrettable for those who take the term 'applied linguistics' seriously, as neither the extreme laboratory nor the extreme qualitative approaches have much to say that is both generalizable and of direct interest to language teachers and learners. Another consequence of this development is that more and more specialized journals have developed over the years . . . [T]his development also contributes to a narrowing of our field of vision, rather paradoxical in these times of much-vaunted interdisciplinarity.

Although I agree with DeKeyser to a large extent, I propose to look at the multitude of perspectives in the fields of SLA and applied linguistics in the following way. First of all, one should make a distinction between, on the one hand, disciplines that address fundamental issues (such as those listed above) and, on the other hand, disciplines that address issues of an applied, mainly educational nature (such as language teaching and assessment, the diagnosis and treatment of language disorders, bilingual education or language policy with respect to minority languages).

Within the field of fundamental research, I believe it is a good thing to distinguish, not so much between quantitative and qualitative approaches, but rather between theories and approaches formulated in a way that allows empirical testing and those that are not, or not yet, empirical in this sense. As mentioned earlier, nine theories of SLA are presented in VanPatten & Willams (2007a) – or actually ten theories, if we include Monitor Theory, presented by the editors in Chapter 2. In my view, of those ten theories, six fall into the empirical category: Universal Grammar (White), the Concept-Oriented Approach (Bardovi-Harlig), Skill Acquisition Theory (DeKeyser), Input Processing (VanPatten), Processability Theory (Pienemann) and Input, Interaction and Output (Gass & Mackey). That is to say, some or most of the ideas or claims made in these theories can be tested. However, a disturbingly large number of theories (or models, approaches or frameworks) are not yet empirical, or only partially so, or the scholars who propose them do not state explicitly to what extent their theory can be made empirical. Of the ten theories presented in VanPatten & Williams (2007a), this applies, I believe, to the following (and here I stick my neck out and you can shoot me down in the discussion after this presentation!): Monitor Theory (Krashen), the Associative-Cognitive CREED (N. Ellis), Autonomous Induction Theory (Carroll) and Sociocultural Theory (Lantolf & Thorne), presented in Chapters 2, 5, 9 and 11 of VanPatten & Williams (2007a).

In my view, what is potentially more threatening to the field than a split between quantitative and qualitative subfields is the PROPORTION OF NONEMPIRICAL THEORIES. If an academic discipline is characterized by too many nonempirical ideas and too few empirical ideas, it runs the risk of losing credit in the scientific community at large (and in society).

Now please do not misinterpret what I am saying here. I am not against nonempirical ideas. On the contrary, in any healthy discipline there are scholars who express new ideas that are not yet ready for empirical investigation, and the theories I listed as nonempirical or not-yet empirical all express fascinating and potentially important ideas. A genuinely

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

academic discipline explicitly and emphatically creates room for new ideas that scholars can play with, even if these ideas are not yet ready for empirical testing. However, it would be good if scholars proposing ideas (whether they be called 'theories', 'models', 'approaches' or 'frameworks') explicitly acknowledged to what extent they are not yet ready for empirical testing. (Indeed, some of the authors who present their theory in VanPatten & Williams (2007a) do acknowledge this.)

In fact it can take many years until a theory or hypothesis can be empirically tested. For instance, Krashen launched his Monitor Hypothesis, based on the distinction between acquisition and learning, in the late 70s (Krashen 1977). Several researchers (e.g. McLaughlin 1978) criticized this hypothesis because Krashen had not shown how acquired and learned knowledge could be empirically operationalized, and hence distinguished. However, almost twenty years later, cognitive neuro-scientists reached a consensus that declarative, factual knowledge (Krashen's 'learned knowledge') is stored in the medial temporal lobe (in particular in the hippocampus), whereas procedural, relatively unconscious knowledge (Krashen's 'acquired knowledge') is stored and processed in various (mainly frontal) regions of the cortex (Squire & Knowlton 2000). Thus one could argue that, in contrast to thirty years ago, Krashen's Monitor Hypothesis can now be regarded as an empirical claim.

In other words, I am not advocating banishing all nonempirical ideas from the SLA field, but what would really make the field more TRANSPARENT for both SLA-ers and outsiders is if scholars who propose theories were to indicate to what extent their theory is ready for empirical scrutiny. This would also help us gain a better view of the agenda of our discipline.

According to Long (1993) there are (or were, in 1993) between 40 and 60 theories of SLA. Long found this worrisome. I agree with Jordan (2004: 13) that, in principle, it does not matter whether the SLA field is inhabited by many theories. However, what would help us all, senior and junior researchers and our students alike, is that we see the structure of the SLA landscape not only in terms of the 'issues', as do most of the SLA textbooks mentioned at the beginning of this talk, but also in terms of their empirical or nonempirical status. In almost all of my undergraduate and graduate classes I recommend that my students, when they try to make sense of what they read, follow Jordan's unsurpassed guidelines (Jordan 2004: 115–117):

1. How coherent is the idea or theory?
2. Is the theory testable?
3. What is the scope of the theory, i.e. which and how many phenomena can the theory account for?
4. How (potentially) fruitful is the theory?
5. How simple is the theory?

In addition to these extremely helpful guidelines, I would like to add the following, with reference to the scientific cycle, shown in Figure 1.

1. Does the author explicitly state the problem or puzzle that (s)he wants to explain?
2. If the proposed ideas are not ready for empirical investigation, does the author explicitly acknowledge this?

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press

4. Conclusion

I agree with Ortega (2009, Chapter 10) that it is unwise to exclude socio-cultural approaches from the SLA discipline. SLA concerns human beings in a multitude of aspects: biological, cognitive, socio-psychological and socio-cultural dimensions. Virtually all the issues raised in the relatively short life of SLA (1967 to the present day) are relevant to L2 learners and their environment. However, not all issues that are important to us, or issues that are dear to our hearts, can (yet) be scientifically studied (think of issues in the realms of religion or art, for example). With respect to SLA, it would help us all to obtain a good view of the SLA theoretical landscape if all strived for more TRANSPARENCY along the dimensions mentioned in this talk, and this applies to ALL theories (models, approaches, frameworks) and whether we are a senior, junior or even a non-SLAer. The fact that the number of empirical SLA theories is large is not in itself a problem. Eventually, through the practices of rational 'normal science', including replication (Kuhn 1962), the best theories (in terms of coherence, testability and scope) will rightfully come out on top.

References

Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.
Corder, S. P. (1967). The significance of learner errors. *International Review of Applied Linguistics* 5, 161–169.
DeKeyser, R. M. (2010). Where is our field going? Comments from the outgoing editor of *Language Learning. The Modern Language Journal* 94, 646–647.
Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
Jordan, G. (2004). *Theory construction in second language acquisition*. Amsterdam: Benjamins.
Krashen, S. D. (1977). The Monitor Model for second language performance. In M. Burt, H. Dulay & M. Finocchiaro (eds.), *Viewpoints on English as a second language*. New York: Regents, 152–161.
Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
Long, M. H. (1993). Assessment strategies for SLA theories. *Applied Linguistics* 14, 225–249.
McLaughlin, B. (1978). The Monitor Model: Some methodological considerations. *Language Learning* 28, 309–332.
Mitchell, R. & F. Myles (1998). *Second language learning theories*. London: Arnold.
Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
Spada, N. (2010). Review of Ortega (2009). *Studies in Second Language Acquisition* 32, 651–652.
Squire, L. R. & B. J. Knowlton (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. In M. S. Gazzaniga (ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press, 765–779.
Towell, R. & R. Hawkins (1994). *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.
VanPatten, B. & J. Williams (eds.) (2007a). *Theories in second language acquisition: An introduction*. Mahwah, NJ: Erlbaum.
VanPatten, B. & J. Williams (2007b). Introduction: The nature of theories. In B. VanPatten & J. Williams (eds.), 1–16.

JAN HULSTIJN is Professor of Second Language Acquisition at the University of Amsterdam. His research interests are implicit/explicit and incidental/intentional learning, skill acquisition and components of language proficiency. He recently proposed the constructs of basic and higher language cognition, contributing to the explanation of individual differences in L1 and L2 attainment. His webpage is http://home.medewerker.uva.nl/j.h.hulstijn/.

https://doi.org/10.1017/S0261444811000620 Published online by Cambridge University Press