

The Role of Theory in SLA Empirical Research on the Critical Period Hypothesis

Paige Yi

Teachers College, Columbia University

The construct of theory is a value-laden one, with its time-honored Ancient Greek root denoting viewing and beholding, whose derivatives allude to contemplation and speculation, as well as its modern sense, “an intelligible explanation based upon observation and reasoning” (*theory*, n.d., para. 2). In contemporary research realms, the notion of theory is also of vital significance in any scholarly discipline that reckons itself as scientific, inasmuch as contended by Suppes (1974, p. 4) well over four decades ago, “the obvious and universally recognized importance of theory in the more mature sciences is strong evidence for the universal generalization that theory is important in all sciences.”

At its simplest, the word ‘theory’ in scientific research enterprises is to provide rich insights into questions as to what, why and how—making sense of natural and social phenomena. To wit, that sense-making process per se can to a large extent be deemed as a matter of theorizing itself. According to VanPatten et al. (2020, p. 4), a theory as a heuristic ought to satisfy the following three broad duties: to explain, to predict and to unify. To be more specific, accounting for observed phenomena or perhaps even illuminating phenomena different from those the theory is originally developed to demystify; making predictions about what is likely and unlikely to transpire under specific conditions; unifying a series of phenomena that on the surface level may seem to be unrelated, yet deep down are closely intertwined with one another; and generating hypotheses which can be tested empirically are all that could be subsumed under a theory (VanPatten et al., 2020, pp. 2-4). Of note, although evidently different from each other, the distinction between the construct of theory and that of hypothesis is not consistently upheld in the real world including the research field of second language acquisition (henceforth SLA), resulting in the interchangeable usage between these two in the literature (VanPatten et al., 2020).

This paper therefore sets out to peruse the role of various theories or more precisely, hypotheses invoked in SLA research by surveying three empirical studies pertaining to the critical period hypothesis in the SLA of phonetics and phonology. In particular, the three studies which will be reviewed in the next section are titled in chronological order as (1) *A critical period for learning to pronounce foreign languages?* (Flege, 1987); (2) *Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment.* (Ioup et al., 1994); (3) *Ultimate attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction* (Moyer, 1999).

Albeit a general consensus that language-learning aptitude (Carroll, 1965, 1991; Skehan, 1986, 1993) and motor skills (e.g., Smith, 1978) continue to develop throughout early adulthood, child learners appear to find the task of language acquisition much more straightforward than their adult counterparts. This has been borne out in areas including phonetics and phonology, morphology, syntax, semantics and discourse pragmatics for both native (i.e., L1) and non-native

(i.e., L2) acquisition (e.g., Ioup et al., 1994). While it might be the case adult L2 learners initially outshine children notably in terms of their learning pace, child L2ers by and large catch up and subsequently trump their adult equivalents in ultimate attainment (Moyer, 1999). As a corollary to this observation, some may tend to endorse the *prima facie* hunch that there exists inherent difference among L2 acquirers of various ages, with adult learners possessing innately inferior language learning abilities to children by virtue of their maturational constraints, or vice versa. This postulation has in fact found its bearing on what is known as the ‘Critical Period Hypothesis’ (hereafter CPH) in the inquiry of language acquisition.

First coined by Penfield and Roberts (1959) and later refined and popularized by Lenneberg (1967), the CPH was initially meant to provide a neurobiological account and later predict difficulty of acquiring one’s L1 after the optimal biological window of opportunity closes. Probably the most empirically cogent evidence espousing the CPH was that children were able to recover their language from brain lesion owing to a transfer of language functions from the dominant to the non-dominant brain hemisphere whereas aphasic adults could not if the brain trauma was post-pubertal (Lenneberg, 1967). Other well-known evidence in support of a critical period came from dismal cases of child abuse or feral children, deprived of language use and social interaction until after the age of puberty when recused, who later failed to fully acquire their L1s, despite much instructional efforts attempted by the researchers (Ioup et al., 1994). Another kind of L1 acquirers that figured prominently in the CPH research literature were congenitally deaf children born to hearing parents, who grew up being exposed to neither spoken nor signed input until school age or much older. And when juxtaposing these late L1 acquirers of postponed L1 exposure in deaf population with those in normal contact with sign language from birth, the former group usually exhibited incomplete acquisition of their late-learned native language especially with respect to the closed class constructions of morphosyntax (e.g., Mayberry, 1993, 2007), thus bolstering the existence of maturational constraints of language acquisition (Long, 1990).

The hypothesis for a critical period, hence by extension, seemed just natural for SLA scholars who later put forth ‘Sensitive Period Hypothesis’ (i.e., SPH) instead which was a less rigid proposal in an effort to capture the variability of L2 acquisition outcomes. As such, the biologically determined optimal period of sensitivity is no longer delineated as an abrupt or absolute timeline beyond which L2 acquisition is impossible to come about, but rather a gradual process whereby ultimate attainment in L2 is highly variable across individuals (Long, 1990). Although a critical or sensitive period has largely been substantiated through both experimental and case studies in L1 acquisition and comparing child-adult L2 achievement in SLA, the precise nature of the age factor and success (i.e., ultimate attainment) relationship still very much remains an area of contention in SLA (see, e.g., Flege, 1987; Patkowski, 1990). The empirical dissonance, at least where the L2 is concerned, centers primarily around the issue of whether the attainment effects observed are in and of itself maturational in nature (viz., to be explained in terms of a critical/sensitive period), or due to the nuanced interplay between the age factor and other non-neurobiological ones. Albeit the wide currency the hypothesis has enjoyed in the field, the three studies sampled for discussion in this paper all sought to challenge the CPH/SPH in a way that factored other variables potentially conflated with age into their investigation, thereby bringing counterarguments to the fore of SLA research arena.

Ioup et al (1994) probed the linguistic competence of an exceptionally successful L2 learner of Egyptian Arabic (EA) by using a case study methodology that yielded a wealth of data shedding light on her tacit knowledge of multiple linguistic domains and learning history. More

specifically, it compared in detail the phonological and syntactic intuitions of an exceptional naturalistic adult L2 learner of EA to 1) those of a highly successful formally instructed adult L2 learner of EA and 2) to those of native speakers of EA. The two highly proficient non-native speakers of EA involved in the study were Julie (naturalistic learner) and Laura (instructed learner), both of whom were first exposed to EA after the close of their critical periods (after 21 years old).

In order to gauge the extent to which the two exceptional L2 users' linguistic repertoires match those of the native speakers of EA, respectively, five tasks designed to probe their intuitions on different areas of the target language were carried out. They consisted of a speech production task, an accent identification task, a translation task, a grammaticality judgment task and an anaphoric interpretation task, respectively. The purpose of this study is to first determine how closely an untutored yet talented adult L2 acquirer resembles native speakers of that target language in regards to her linguistic competence; and secondly to identify if formal instruction serves as a factor that sets L2 learners apart. Ioup et al. (1994) argued that this study could not only allow for a reexamination of the CPH, but might provide implications for the roles of talent in exceptional adult language learning.

In the introduction/background section leading up to the actual study, various theories or hypotheses were reviewed by Ioup and her colleagues including the CPH/SPH; 'no access to Universal Grammar (UG) vs. full access to UG'; as well as constructs of 'language learning talent' vs. 'language aptitude'. In particular, two versions of L2 access to UG (viz., no vs. full access) were alluded to in an attempt to account for, firstly why it is possible for adults to acquire some degrees of L2 with their L1(s) already in place, and secondly why L2 ultimate attainment exhibits a multitude of various learner outcomes at their end stages, unlike its L1 counterpart. Adherents of zero access to innate language faculty purport that adult language learning is primarily achieved through reliance upon L1 knowledge, facilitated by general cognitive learning mechanisms (e.g., Bley-Vroman, 1990; Clahsen, 1990). At the opposite extreme, proponents of full access to UG by the late L2 acquirers offer two strands of thinking: the first of which posits a competing problem-solving cognitive structure developed right around puberty in tandem with the language-specific mechanisms in L2 input data analysis. And it is the improper applications of the competing procedures that contrive the weakened L2 skills observed in adults (Felix, 1985). That said, scholars (Flynn, 1989; White, 1989, 1992) from the second strand maintain that adult learner difficulties largely emanate from their inability to reset the appropriate parametric values in the L2, hinging instead on their prior L1 settings for the respective L2 features.

As highlighted by Ioup et al. (1994), none of the above theories (viz., the CPH and different versions of UG theory) in fact addresses or accounts for the issue of exceptionally successful L2 adult learners who actually manage to reach native or near-native competence levels. Thus, the construct of 'language-learning talent' was introduced in juxtaposition with that of 'language-learning aptitude' in the background section. According to them, these two constructs differ from one another in that 'talent' is defined as "neurocognitive flexibility" in a sense that "talented learners do not process L2 input in terms of the rigid parameters they have set for L1" but rather, they "are flexible enough to set new parameters or neural pathways for L2", hence implying "bilateral processing" (p. 75). On the other hand, 'language aptitude', typically discussed in relation to the instructed setting of L2 acquisition (Ioup et al., 1994), traditionally has four components (Carroll, 1965, 1991): phonetic coding ability, grammatical sensitivity, inductive language learning ability and associative memory, which Skehan (1986, 1993) later collapsed into three categories: phonetic coding ability, language-analytic ability, and

associative memory. An issue with language aptitude research, as pointed out by Ioup et al. (1994), is that there exist few measures with which to tease apart exceptionally talented learners from a pool of acquirers also with high aptitude, which is why SLA research may necessitate a construct of ‘language talent’ in its own right. Nonetheless, a reader might naturally arrive at the following questions: How to operationalize the construct of ‘talent’? Is there any neurological evidence underpinning this very notion in a way that distinguishes itself from ‘language aptitude’? Because in order for a theory to be one, it ought to fulfill its duties: making predictions (i.e., generating hypotheses) that are empirically testable, and affording constructs that are operationalizable.

In the conclusion section, Ioup and her associates put much explanatory emphasis on the construct of ‘language-learning talent’ to contribute to the L2 success of the study whilst spared much of their ink on the CPH and UG theory. This appears to suggest that at least as the authors viewed it, ‘language learning talent’ was the sole one in the study which not only had played an explanatory but also predictive role in addressing the success of the exceptional L2 user. And as anticipated, the authors availed themselves of the L2 user’s self-report and anecdotal learning history as the last recourse to ‘operationalize’ language-learning talent, which perhaps could have been complemented with empirical evidence obtained through behavioral measures tied to concurrent neurological developments such as ERPs and neuroimaging techniques, so as to increase confidence in their findings as well as validity of the ‘talent’ construct.

Flege (1987) and Moyer (1999) both took a similar stance against the CPH at the very outset of their papers holding that it provides an overly simplistic perspective on an inherently complex phenomenon that is highly susceptible to both biological and sociopsychological factors. Flege (1987) even argued in his conclusion that since it is extremely difficult, if not entirely impossible, to control all the non-neurobiological variables that largely confound and covary with the age factor in purely behavioral research (perhaps at that period of time), the CPH is naturally reduced to an *a priori* assumption instead of an empirically testable hypothesis concerning human speech learning. In line with Flege’s viewpoint, the CPH, as aptly revealed in Bialystok (1997), has often in its weaker forms “lingered in discussions of both theory and practice, occasionally becoming explicit but usually lurking as a set of tacit assumptions” (p. 116). This invites me to infer that invocation of a critical/sensitive period may as a matter of fact deflect scholars away from unraveling equally important causes of heterogeneity of L2 ultimate attainment, given that proof of its existence in human beings (e.g., with a predictable and abrupt endpoint) has largely been elusive. As such, one pivotal conclusion at which Flege (1987) arrived is that presupposing the presence of a critical period may have a detrimental effect on other hypotheses that in fact *can* be tested yet seem to be unwarranted as a consequence of being overshadowed by the CPH.

One competing hypothesis put forth by Flege (1987) has to do with psycho-perceptual and phonetic causes associated with L2 learners’ massive prior experience with their L1. In other words, rather than attributing the varied ultimate attainment in L2 phonetics/phonology to the maturational constraints as the explanatory factor, Flege assigned the interpretive as well as the predictive power to learners’ entrenched motor articulatory habits and restricted perceptual targets for their L1-biased phonetic categories in the production and perception realm, respectively. As neatly delineated by Flege (1999, p. 125), one may speak a L2 with a trace of their mother tongue accent “not because one has lost the ability to learn to pronounce, but because one has learned to pronounce the L1 so well”!

In a similar vein, Moyer (1999) also aimed to contest the CPH/SPH by deliberately recruiting highly advanced late L2ers of German, the majority of whom self-reported high levels of affective disposition towards learning the L2 pronunciation as well as a concern to sound native-like. Following Flege's line of reasoning apropos the CPH, she took one step further hypothesizing the inextricability of the nature (neurobiological) and nurture (non-biological) factors. Before conducting her study, Moyer predicted that the age factor would be observed to indirectly affect L2 production through its dependent relationship with other non-biological factors, to the point where the predictive power of age alone is equivocal! In particular, the various sociopsychological factors she incorporated in her study primarily fall under two categories: (1) instructional variables including total years of instruction, immersion, types of phonological instruction and feedback received, segmental or/and suprasegmental contrastive drills received in terms of hours if any; (2) affective variables such as personal and professional motivation for studying German, self-appraisal of pronunciation, concerns for sounding native, and personal assessment for linguistic and cultural assimilation while studying abroad.

In designing the production tasks of the study comprising controlled reading-aloud task along with less-controlled narration task, Moyer specifically chose a series of phones such as /ø,y,ç,x,ʔ/ that could be a challenging stumbling block to English acquirers of German according to the Contrastive Analysis Hypothesis (CAH). In essence, the CAH is a 'theory' with a direct behaviorist tie harkening back to the 1960s, under which the extent of the typological difference between L1 and the target language is considered the primary source of learner difficulty and error (VanPatten & Williams, 2015). In Moyer's study, CAH was first invoked in the procedure section in a way to generate the author's own hypothesis of what might come about in the read-aloud tasks. Subsequently, Moyer circled back to the CAH in the discussion section trying to partially ascribe the failed endeavor to the very hypothesis that the sociopsychological factors examined did not override the effect of age (since all her participants had performed within the non-native range except for one outlier which she later excluded).

To sum up, theory has an essential status in any scientific enterprise for research advancement, due not only to a structured set of lenses offered through which aspects of a given phenomenon can be observed and analyzed, but fundamentally to its explanatory, predictive and unifying powers. The current paper combed through the roles various theories and hypotheses have played in three empirical SLA studies placed under scrutiny, all of which are concerned with the CPH and L2 ultimate attainment/success. In light of the foregoing discussion, what can be gleaned overall from the three sampled studies is that firstly there does not appear to be a long-standing unified 'theory' concerning age-effects which is supported by all SLA researchers representing different schools of thought, some of which can seem mutually incompatible, if not directly contradictory. By the same token, none of the theories or hypotheses discussed in the studies above has managed to unify or bring all of the observable phenomena in various populations (e.g., the hearing or the deaf; early L2ers or late L2ers, etc.) "under one umbrella" (VanPatten & Williams, 2020, p. 4). That being so, most of them indeed have taken on the guiding, explanatory or/and predictive roles in the respective studies to various extent as delved into earlier, thereby rendering their linkage with their corresponding study designs tight.

REFERENCES

- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13(2), 116–137.
- Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, 20(1-2), 3–49.
- Carroll, J. (1965). The prediction of success in foreign language training in R. Glaser: *Training, research, and education*.
- Carroll, J. B. (1991). Cognitive abilities in foreign language aptitude: Then and now. *Language aptitude reconsidered*, 11–29.
- Clahsen, H. (1990). The comparative study of first and second language development. *Studies in Second Language Acquisition*, 12(2), 135–153.
- Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics*, 8(2), 162–177.
- Flege, J. (1999). Age of learning and second language speech. *Second language acquisition and the critical period hypothesis*, In D. P. Birdsong (ed.), 101–131. Mahwah, NJ: Erlbaum.
- Felix, S. W. (1985). More evidence on competing cognitive systems. *Second Language Research*, 1(1), 47–72.
- Flynn, S. (1989). The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japanese speakers. *Linguistic perspectives on second language acquisition*, 89–108.
- Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16(1), 73–98.
- Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12), 59–67.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251–285.
- Mayberry, R. I. (1993). First-language acquisition after childhood differs from second-language acquisition: The case of American Sign Language. *Journal of Speech, Language, and Hearing Research*, 36(6), 1258–1270.
- Mayberry, R. I. (2007). When timing is everything: Age of first-language acquisition effects on second-language learning. *Applied Psycholinguistics*, 28(3), 537–549.
- Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 81–108.
- Patkowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11(1), 73–89.
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton University Press.
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3(2), 188–221.
- Skehan, P. (1993). Cognitive abilities and language acquisition. *Unpublished manuscript, Thames Valley University, London*.
- Smith, B. L. (1978). Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics*, 6(1), 37–67.
- Suppes, P. (1974). The place of theory in educational research. *Educational Researcher*, 3(6), 3–10.
- Theory. (n.d.). In *online etymology dictionary*. <https://www.etymonline.com/word/theory>

- VanPatten, B., Keating, G. D., & Wulff, S. (Eds.). (2020). *Theories in second language acquisition: An introduction*. New York: Routledge.
- VanPatten, B., & Williams, J. (2015). Early theories in SLA. *Theories in second language acquisition: An introduction*, 17–33.
- White, L. (1989). The principle of adjacency in second language acquisition: Do L2 learners observe the subset principle. *Linguistic perspectives on second language acquisition*, 134–158.
- White, L. (1992). On triggering data in L2 acquisition: a reply to Schwartz and Gubala-Ryzak. *Interlanguage Studies Bulletin (Utrecht)*, 8(2), 120–137.

Paige Yi is a doctoral student in the Applied Linguistics program at Teachers College, Columbia University. Her research interests primarily lie in cross-linguistic phonetics, language policy, as well as socially applied aspects of SLA and multilingualism. Correspondence should be sent to E-mail: pjy2107@tc.columbia.edu