CAMBRIDGE
UNIVERSITY PRESS

FIRST PERSON SINGULAR

A personal bookshelf of foreign language aptitude publications

Peter Skehan

University College London, London, UK
Email: peterskehan@gmail.com

Peter Skehan taught at universities in the United Kingdom, Hong Kong, and New Zealand, before retiring several years ago. He is currently an Honorary Research Fellow at the Institute of Education, University College London. His main interests have been in language aptitude, language testing, and task-based performance, and he has published in all these areas.

Like many in the profession, from the United Kingdom certainly, the teaching of English was a second, or more accurately, third career for me. After periods as a clinical psychologist and primary school teacher, the lure of travel took me to France. The need to earn money, the discovery that the choices about how to do this were limited, led to a move into teaching English as a Foreign Language. And, again like many in the profession, there was the slow realisation that I had stumbled into a career, and that language teaching was not to be a temporary interlude in the sun.

This thought, that language teaching was for the long haul, pulled several things together. My background was in psychology, in contrast to the language and linguistics people who were typical colleagues. I had also become interested in what was then termed mixed-ability teaching, and was impressed by the insight of the unit director at the University of Besancon, where I taught, that learner differences, and aptitude particularly, were under-researched. This idea stayed with me. I started to read in the area and then, when I decided that doing a doctorate might be the best career choice, I realised that I already had my topic, and that my psychology training had, without my noticing, given me a whole set of relevant skills: statistics, familiarity with relevant literatures, and insight into research design.

I did complete that doctorate, and then had to confront the unanticipated. At the time (the late 1970s, early 1980s), I seemed to be the only person interested in language aptitude. An exaggeration, obviously, but at the same time, not that far from the truth. Since then, it has been a source of fascination to me that interest in aptitude has slowly re-emerged. From the 1990s onwards there has been a resurgence of interest, and so now aptitude has become quite a hot topic. In some ways, in those earlier years, aptitude was never central to applied linguistics – more a tool used in a limited range of situations. But the renewed interest changed that. There were connections with second language acquisition research. It also started to emerge that aptitude might be relevant to first language–second language (L1–L2) connections and a possible critical period for language development – in other words, connected with some of the central preoccupations in applied linguistics.

The 12 publications that I have included here are intended to capture these developments. They give a snapshot about how the field has developed, about what the current hot areas are, and most of all, they are the publications that have interested and impressed me most in the study of aptitude over my career. There are some from the past, but many that are recent – reflecting current interests. I have organised the 12 articles into four loose groups: aptitude batteries, the use of aptitude scores, focussed research into aptitude, and large data studies. This categorisation is far from watertight, but it does provide a framework for wider discussion, as I try to do what all bookshelf compilers do – argue for the totally central position of my subject!

Developing and researching aptitude batteries

There are three indispensable articles in this section: two (very contrasting) papers focus on aptitude battery construction, and one investigates validity in the currently most popular aptitude battery. They are central to how aptitude research can be done but they also raise some of the fundamental questions that arise regarding our understanding of aptitude.

(1) Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). University of Pittsburgh Press.

Essentially, the modern study of aptitude starts with an incredibly influential article by Jack (J. B.) Carroll (1962). Although principally a psychologist, focussing on cognitive abilities, in the 1950s he became interested in foreign language aptitude as one possibly specialised branch of his wider views on intellect. Working with Stanley Sapon, he developed the Modern Languages Aptitude Test (MLAT: Carroll & Sapon, 1959), a test that is still used, nearly 70 years after its development. The included article is the most comprehensive account of how this test was developed, how it was validated, and how it relates to underlying theory. The chapter had a huge influence upon me and inspired my interest and research in aptitude. Carroll seems to me to have been exactly the right person, at the right time, to shape how aptitude can be understood.

Carroll started by thinking about what foreign language aptitude might consist of. So, with Sapon, he developed a set of over 40 potential aptitude tests. Painstaking validation work was then carried out, involving hundreds of participants, of various ages, all L1 English but learning a wide range of L2s. The five tests that were retained all predicted success reasonably well, and distinctly from one another. They formed the MLAT, and all this research was the basis for the MLAT Manual, a model of what such a publication should be. Carroll also produced his enduring theory of foreign language aptitude, consisting of phonemic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory. If one puts the two language factors together, this gives his view of aptitude as sound – language/learning – memory, a view that has been foundational to the field ever since, and I will return to it throughout.

(2) Linck, J. A., Hughes, M. M., Campbell, S. G., et al. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*(3), 530–566. doi:10.1111/lang.12011

Next, we have a more recent account of aptitude battery construction. The Center for the Advanced Study of Language, Maryland, was established with the mission of developing an aptitude test to identify individuals with the potential to attain very high levels of language proficiency. The test that emerged, the Hi-LAB (Linck et al., 2013), named to reflect its function as a high-level language aptitude test, is a very impressive test, and its innovation, thoroughness, and contrast to the MLAT are why I have included this article.

Four features stand out. First, psychology has changed enormously since Carroll's aptitude research, particularly regarding psycholinguistic processes and learning, and Hi-LAB reflects this. Second, the focus is on those who will achieve close to native speaker proficiency. Previous attempts to predict

high-level achievement have failed, and so the Hi-LAB is a noble attempt to address a long-standing and difficult problem. Third, the procedures in the development of Hi-LAB sub-tests have been very impressive. Considerable trialling has been involved, reliability standards are very high, and validation of the different sub-tests carefully done. Fourth, a major issue, though, has been the closed nature of this test. Research was funded by the U.S. government, and the use of the test is largely confined to that context.

An evaluation of this article has to be rather mixed in nature. The development work was of extremely high standard. But because the test is not freely available, I feel that wider impact is disappointing. Subsequent publications have been interesting, and perhaps clarify the potential that Hi-LAB has, such as a basis for counselling, and for diagnostic purposes, and even use in conjunction with MLAT (Doughty, 2019). One of the most striking features of the Hi-LAB, at least in my view, is that the focus on language itself is not great, which I think is a pity.

(3) Bokander, L. (2023). Exploring the predictive validity of the LLAMA language aptitude tests: A research synthesis. In Wen, E., Skehan, P. & Sparks, R. L. (Eds.). *Language aptitude theory and practice* **(pp. 94–116). Cambridge University Press. doi:10.1017/9781009076463.007**

Most recent aptitude research has used the LLAMA battery (Meara, 2005; Rogers, Meara, & Rogers, 2023). This battery, loosely modelled on the MLAT, is computer-delivered, very accessible, and free. Yet, in contrast to the MLAT or HiLAB, it is not, at the time of writing, a validated test. Recently, though, Bokander (2023) has attempted to investigate LLAMA validity. First, he had to locate as many research studies as he could that use LLAMA, and that met his adequacy criteria. He found 36 such studies. In this way, a reasonable patchwork view of LLAMA validation is possible. This study is on my bookshelf partly because of how methodologically instructive it is, and partly because such work is vital for aptitude progress if LLAMA continues to be widely used.

The pattern of results is concerning. Bokander (2023) explored the separate LLAMA sub-tests and examined whether the correlations they generate reach significance. For example, LLAMA B (vocabulary learning) generated 82 correlations in total, of which only 16 were significant. The composite LLAMA test fares better, with 43 correlations, of which 14 were significant. Bokander also explored whether there was a better relationship between LLAMA sub-tests and different types of outcome measure – for example, LLAMA B with tests of vocabulary learning. There was not. Bokander also explored the studies that do report higher correlations, and showed that where such relationships are reported, there is a strong tendency for them to be based on small sample sizes. These findings suggest we may now need to re-evaluate the robustness of many of the findings reported in the literature, if they are based on LLAMA. Even so, it is important to stress, following Bokander, that we are not dealing here with negative evidence, but rather lack of evidence. A systematic validation study of the LLAMA tests simply has not been completed yet, but is vital.

These three bookshelf entries give a clear view of where we currently are with aptitude measurement, and are important to anyone thinking of doing research in the area. If I needed to measure aptitude, I would still choose the MLAT, since it is the only battery that is validated and available. I hope this changes. The entries, though, clearly raise some major issues in aptitude research. First is the question as to whether language aptitude needs to involve language (since the HiLAB suggests not), or whether Carroll's sound-language/learning-memory conceptualisation is still the best we have. Second, is aptitude for selection, or to make instruction more effective? Third, how will we develop a publicly-available aptitude test for high levels of accomplishment?

Using aptitude information in different contexts

The basic theme in this section is occasions when aptitude itself has not been the focus, but where aptitude information is used to illuminate some other research question. Two studies are concerned with understanding language learning itself, while the other article explores the relationship between

aptitude and age of arrival (AoA) in a country and examines whether aptitude helps our understanding of the possible existence of a critical period for language development.

(4) Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research* 9(2), 147–171. doi:10.1191/1362168805lr161oa

Erlam (2005) is a study based on 14-year-old New Zealand schoolchildren learning French – an example of research that I would like to see more of, and hence its inclusion on my bookshelf. She administered aptitude tests to three groups of learners. The focus of the instruction was on direct object pronouns. Each group received different sorts of instruction: deductive, inductive, and structured-input. In the first condition, there was rule explanation, form-focussed activities, and feedback. Inductive instruction had no rule explanation, but practice activities facilitated first hypothesis testing, and then production. In the structured instructional condition, explicit instruction and rule explanation were given, followed by input-based activities requiring interpretation of meaning.

The deductive condition produced the best performance overall, and then the structured input was slightly better than the inductive condition. Erlam also correlated the scores for each of the groups, separately, with aptitude, and this produced the most interesting results of all. Aptitude did correlate with achievement scores in the structured input and inductive groups, so it seemed that in these groups aptitude contributed to learning. In contrast, there was no such correlation in the deductive group – aptitude did not seem to make a difference here. One of Carroll's great insights was to suggest that a major reason for a greater understanding of aptitude is to know how best to negate it. The Erlam study is an excellent example of this – it suggests that aptitude is most relevant when teaching is less structured. The important conclusion to this, of considerable significance, is that quality and clarity of teaching has the potential to neutralise aptitude and equalise learning difficulties.

(5) De Graaf, R. (1997). The eXperanto Experiment: Effects of explicit instruction on second language development. *Studies in Second Language Acquisition, 19*(2), 249–276.

I have chosen De Graaf's (1997) article partly because it complements Erlam's (2005) work in interesting ways. The focus is similar, exploring different methodologies, and so both have potential relevance to pedagogic settings (although De Graaf is experimental). I have also included De Graaf both because of the interesting results, and also because, just as Erlam showed what can be done in actual classrooms, De Graaf shows that there can still be pedagogic connections even in the context of a more experimental research design.

The major contrast was between explicitly and implicitly presented self-study materials, focussing on a range of language rules. The instructional phase lasted some 15 hours in total, making the study a reasonable basis for generalising to actual learning contexts. In general, the explicit presentation produced better results, although implicit presentation was also effective, if less so. But, in an interesting contrast to Erlam, the aptitude measures correlated with achievement scores with both explicit and implicit presentation conditions.

What intrigues me most with this study is these contrasting results. Erlam reported correlations with aptitude for the inductive and structured input conditions, but not for the deductive. De Graaf reports correlations for both the explicit and implicit conditions, so it raises the question of why Erlam's lack of correlation with the deductive condition was not matched by a lack of correlation in De Graaf's explicit condition. One wonders, in that respect, whether the difference concerns the more classroom-based context with Erlam relative to the experimental context for De Graaf. It could be that classroom contexts, with teachers who mediate instruction to take account of mixed abilities, may lead to a reactivity to the learners that experimental self-study materials struggle to provide.

(6) Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. doi:10.1017/s027226310808073x

We turn next to an unusual piece of research in which aptitude information is fundamental. Following DeKeyser (2000), linking foreign language aptitude and a critical period for language development, Swedish researchers (Abrahamsson & Hyltenstam, 2008) set out to replicate and extend De Keyser's work. The participants were all immigrants to Sweden, and a central criterion for inclusion in their study was that they could pass for Swedish native speakers. The article reports on three very significant areas. First, all participants were tested for language aptitude. Second, aptitude results were related to AoA and Swedish proficiency. Third, challenging tests to probe intuitions about Swedish morphosyntax were given.

The results give strong confirmation to De Keyser. Participants who passed for native speakers had high or very high aptitude – that is, aptitude seemed to compensate for language learning beyond a critical period. But such participants did not meet a more stringent criterion – native-like performance on the more demanding Swedish language proficiency test. This result also speaks to the existence, or not, of a critical period, suggesting high-level L2 explicit learning but not implicit learning. But an additional major finding in this study is that, unlike De Keyser, there was a relationship between aptitude scores and Swedish performance for the younger AoA participants – foreign language aptitude seemed to constitute an advantage here too.

In reflecting on this section, the first two studies are unusual in their relative rarity. In other words, one striking feature of aptitude research is the relative lack of pedagogy-linked studies – using aptitude to tailor instruction to learners' strengths and weaknesses. Fortunately, there are studies such as Erlam's (2005) and De Graaf's (1997) that offer important insights (and reviews of more of such studies are found in Li (2015)). Erlam, for example, suggests that some instructional methods might have greater potential to minimise aptitude effects, and both Erlam and De Graaf take a nuanced view of when aptitude relates differently to different types of instruction

The other bookshelf application concerns a critical period. This is remarkable research and speaks to issues such as connections between L1 and L2 development (and see Sparks & Dale (2023) later on), and centrally as to whether language is special. It also connects with DeKeyser's (2000) proposal of an implicit-explicit learning contrast. The implications, if these claims are valid, go to the heart of linguistics more generally. I have, I must admit, been surprised that this study (and others by the same group) have not had greater impact throughout the linguistics world and have not stimulated a great deal of additional research.

Research into focussed aspects of aptitude

There are four rather disparate studies in this section of the bookshelf. What brings them together is that they explore aptitude itself, but do so in focussed, but different ways. If one returns to Carroll's multi-componental view of aptitude (sound – language/learning – memory), three of the studies are concerned with just one of these: Saito (2023) and sound, Li and De Keyser (2021) with language learning or at least pattern learning, and Wen (2019) with memory. Each of these strives for greater depth within one limited area, an interesting strategy for progress. Then, Turker et al. (2021) present a fascinating account of neurolinguistic correlates of aptitude, surely a forerunner of the future. I include all four because they show the vitality of contemporary aptitude research.

(7) Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing the Auditory Precision Hypothesis-L2. *Language Teaching, 56*(4), 522–538. doi:10.1017/S0261444822000453

Carroll's original research convinced me that sound, in general, is not the issue for aptitude, but that there is something about language-related sound that is vital. This next item on my bookshelf is a huge

challenge to that belief – hence its inclusion! I have chosen an article by Kazuya Saito (2023) that emphasises domain-general measures of auditory ability. He argues that there is such a thing as a 'good ear', which is vital for language learning, (and other areas too).

Saito (and colleagues) have developed a wide range of tests of different auditory abilities, including pitch, formants, duration, and intensity. These tests are largely computer-generated and computer-delivered. They have been widely researched, involving many different L2s, at different levels of proficiency. In addition, tests have been developed targeting, separately, explicit and implicit auditory processes. It is not claimed that the measured auditory abilities explain everything, by any means, but they are shown, consistently, to explain a large amount of the variance in second language learning success (as do experiential factors). They are relevant at all proficiency levels, though, interestingly, implicit auditory processing makes stronger predictions at higher levels. Such factors are less important in classroom contexts.

Personally, I am still reluctant to let go of the idea of domain-specific abilities being relevant in the auditory area but have to accept that the body of evidence for domain-general capacities is compelling. The only qualification that I have concerns research design. This has typically been to locate populations who have been in an L2 country for some period of time and then to relate performance on the auditory tests to level of achieved proficiency. This is fine, yet I would like to see such testing done with more controlled research designs, of a pre- and post-nature.

(8) Li, S. & De Keyser, R. (2021). Implicit language aptitude: Conceptualising the construct, validating the measures, and examining the evidence: Introduction to the special issue. *Studies in Second Language Acquisition*, *43*(3), 473–497. doi:10.1017/s0272263121000024

I turn next to TYPES of learning. Broadly, cognitive psychology suggests a major contrast between declarative/explicit learning and procedural/implicit learning. Recently, aptitude researchers have shown increased interest in more implicit approaches and explored implicit-oriented aptitude. This particular contribution to my bookshelf (Li & De Keyser, 2021) is a special issue of a prestigious journal, devoted exclusively to the issue. There are multiple contributions by distinguished researchers, exploring different aspects of implicit learning. It is clearly a landmark publication. I will not discuss the individual contributions much here but recommend them all for the insights and contributions they make.

Based on the special issue, several conclusions emerge. First, it is clear that implicit learning is relevant for aptitude, and that there is now evidence that it plays a role in predicting success and can complement explicit approaches. But a second conclusion has to be that more work is needed. Theoretically, we are not clear about the inter-relationships of areas such as implicit learning, statistical learning, and automatised/proceduralised learning. It can be very difficult to distinguish between all of these, and to establish what sort of learning has occurred in particular circumstances – direct implicit learning, or highly automatised learning that started as declarative? Third, there are measurement issues. Many of the authors wrestle with this, and it is central to the commentary article by Perruchet (2021). It currently is challenging to point to one unitary implicit learning process since different tests of implicit learning do not tend to inter-correlate highly, questioning any underlying construct of implicit learning. This bookshelf entry captures the excitement of this developing area perfectly, as well as the questions that need to be addressed in the future.

(9) Wen, Z. (2019). Working memory as language aptitude: The phonological/executive model. In Z. Wen et al. (Eds.). *Language aptitude: Advancing theory, testing, research and practice* (pp. 187–214). Routledge. doi:10.4324/9781315122021–10

We have seen already, in previous bookshelf entries, that working memory (WM) is important, and there are even claims that WM IS aptitude. Wen (2019) certainly argues that WM is central but does not quite equate WM with aptitude. The reason the chapter is on my bookshelf is that it provides an

up-to-date account of these issues. Along the way, it also provides an instructive historical perspective, contrasting the British-European approach, emphasising phonological aspects, and the North American, emphasising on-line executive processing.

Wen also provides an account of his P-E model, (P = phonological and E = executive), which takes the two 'schools', European and North American, and integrates them to be relevant to aptitude. The 'P' component includes processes of rehearsal and recycling, and Wen argues that this is vital for vocabulary and syntax learning. The 'E' component is concerned with general executive memory functions, including updating, switching, and inhibition. These become central for the way attentional limitations are handled, and Long Term Memory connections made (cf. Hi-LAB sub-tests).

The key issue then becomes how relevant this interpretation is for language aptitude. Wen proposes that phonological working memory (PWM) is important in acquiring new phonological forms and in chunking. In this way he suggests that PWM is important acquisitionally. He also argues that PWM is modality specific, that is, to sound. In contrast, the different executive working memory (EWM) functions are then seen as relevant for attention-allocation, supervising operations, and monitoring during language processing, and functions as a language performance device.

(10) Turker, S., Seither-Preisler, M. & Reiterer, S. M. (2021). Examining individual differences in language learning: A neurocognitive model of language aptitude. *Neurobiology of Language, 2*(3), **389–415. doi:10.1162/nol_a_00042**

A neurolinguistic perspective is a different approach to the language-is-special claim. Since brain-localisation for language is well established, this raises the question whether there is evidence linking aptitude scores to established brain-related areas. Turker et al. (2021) is focussed on exactly this issue and adds a fresh twist to this bookshelf! The article reviews evidence showing that there are indeed such correspondences, such as between Broca's areas and aptitude measures of inductive language learning ability. The neurolinguistic data also means that a re-evaluation is required of our understanding of auditory ability. The connection between aptitude and the auditory cortex is clear, particularly in relation to the structure of Herschl's gyrus (with this brain area also implicated in musical ability). We do here seem to be dealing with a biological basis for differences in aptitude that experience may mitigate, but only partially. Fascinating here is that for auditory cortex functioning, the right hemisphere may also have a role.

A second major insight from this publication is the importance of pathways that connect the different language-learning related parts of the left hemisphere, such as the left temporal pathway (connecting Broca's area and the auditory cortex/Wernicke's area). The linkage with aptitude may not have been demonstrated so clearly here, but we are learning more about how these pathways may differ between individuals, in extent, speed, synchronisation, and adaptability. In other words, the localised areas are important, but connections between them are just as vital.

Reflecting on these four bookshelf entries, focussing on one facet of aptitude can be a successful research strategy. It is also interesting that in each case, external disciplines have made major contributions. The results, though, are slightly mixed. It is clear that the auditory work (Saito, 2023; Turker et al. 2021) has made major contributions to our understanding and could have a significant impact on aptitude testing and battery construction. Implicit learning, though, is more problematic. This is an active area, and conclusions may change. But for now, it appears that the potential of linking implicit learning to aptitude has yet to be realised, and findings suggest only small effects (and see Li (2019) in the next section). Similarly, with working memory. Carroll's view of aptitude concerns a much wider conception of memory, so that while the progress linking WM to aptitude is considerable, there is a great deal more of memory ability to be probed.

The studies also bring us back to the question of whether language aptitude needs to involve language-linked material. Saito (2023) suggests general auditory abilities are sufficient, and this is a powerful challenge to LINGUISTIC auditory abilities being relevant. Generally, implicit learning research (Li & De Keyser, 2021) does not emphasise linguistic material. Then, Wen neatly suggests that one aspect of WM (the phonological) is domain specific while the other (executive functions) are not. So, we have something

of a mixed picture, which I interpret as an excellent stimulus for more research. Finally, we have Turker et al. (2021). Fascinating, tantalising, and in my view, partially supportive for the language-is-special position. But once again, things are mixed. Turker et al. (2021) agree with Saito in focussing on general auditory abilities, but their approach seems more consistent with a linguistic perspective where MORPHOSYNTAX is concerned. Undoubtedly the correct conclusion here is to await further research.

Large-scale and review publications

The (more tenuous) unity in the final bookshelf section depends on scale. One of the contributions, Li (2019), achieves scale by providing an important summary and overview of aptitude research, succinctly covering theory and a vast amount of relatively recent research. The other, by Sparks and Dale (2023), reports one study (from a set of such studies by the same authors), with significant numbers of participants and the use of a wide range of tests, including aptitude battery data. The two bookshelf contributions provide an additional perspective on how aptitude research can be carried out, and of L1–L2 relationships.

(11) Li, S. (2019). Six decades of language aptitude research: A comprehensive and critical review. In Z. Wen et al. (Eds.), *Language aptitude: Advancing theory, testing, research and practice* (pp. 78–96). Routledge. doi:10.4324/9781315122021-5

Any busy readers looking for one up-to-date but wide-ranging account of aptitude should start with Li (2019). The two major foci are theory and empirical research. Regarding the first, Li provides admirable and even-handed coverage of Carroll's views, of the Linguistic Coding Difference Hypothesis, of aptitude-treatment interactional work, of the Fundamental Differences hypothesis, and of my own Stages analysis. Regarding empirical research, Li draws upon his own recent meta-analytic work and provides an authoritative overview of the evidence. On general prediction, he reports an overall, aggregated correlation of 0.50 – substantial, but clearly leaving a lot of variance to be accounted for by other variables, such as motivation, context, and quality of instruction. At a more precise level, Li suggests that aptitude plays a stronger role with earlier rather than later stages of language learning, and with school-age students compared to college age. He also argues that an aptitude influence is clearer with explicit rather than implicit learning. Li also investigated the relationship between aptitude and other ID variables. He reports on a clear relationship between aptitude and intelligence, but lower with WM. Li reports a low correlation between aptitude and motivation (0.16), and a negative correlation with anxiety (−0.35).

Li also reviews micro studies, often of much shorter duration, and often linked to wider second language acquisition (SLA) theorising. With feedback, Li reports aggregated correlations of 0.59 and 0.32 for the relationship between aptitude and explicit and implicit feedback respectively. Aptitude also correlates more with achievement after explicit instruction rather than implicit, and with deductive more than inductive. The scope of this chapter means that it provides indispensable coverage within an aptitude library.

(12) Sparks, R. & Dale, P. (2023). The prediction from MLAT to L2 achievement is largely due to MLAT assessment of underlying L1 abilities. *Studies in Second Language Acquisition, 45*(5), 1345–1369. doi:10.1017/S0272263123000037

Richard Sparks has been associated for some years with the Linguistic Coding Difference Hypothesis (LCDH), which proposes that first and foreign language learning are based on a similar group of abilities. I have included a recent article on my bookshelf (Sparks & Dale, 2023), since this relationship has been a long-standing interest of mine. The article explores, with a sizable database, L1 connections with both foreign language aptitude and foreign language achievement. U.S. high school students were measured in terms of L1 skills, foreign language aptitude (with the MLAT), and L2 (Spanish) achievement. Then, the researchers explored two major ideas, UNIQUENESS (the additional contribution

aptitude information makes over and above that of first language skills) and EFFICIENCY *(*the extent to which aptitude scores transmit the effectiveness of first language skills).

Sparks and Dale (2023) report that L1 skills, particularly L1 word decoding and L1 reading comprehension, generate consistently significant correlations with foreign language aptitude and with foreign language achievement. L1 word decoding was also impressive in predicting aptitude. Regarding uniqueness, aptitude's 'added value' appears to be somewhere in the range 2–7% of the variance with foreign language achievement scores. Sparks and Dale (2023) report an efficiency index for MLAT, of moderate strength, typically around 50%.

With these two bookshelf entries, we focus on large-scale studies with complete batteries, exploring whether aptitude is associated with degree of language learning success. The data support, anew, the case for the relevance of aptitude as one important influence (among others) on development. There is less of an emphasis on more focussed issues (cf. the last section). Rather, they support the general view that aptitude, as a whole, has stood the test of time, and also of context variation, relative to other predictors of success. In addition, Sparks and Dale develop the very important argument that (a) language is special, and (b) there are L1–L2 connections.

Final thoughts

The different entries on the bookshelf provide, in my view, an interesting and fair overview of aptitude research. Compiling these bookshelf entries has been an interesting experience, and captured, among other things, the 'doldrums' phase of aptitude research and its more recent resurgence and vitality. I hope this more recent promise is fulfilled, but by way of conclusion, I will offer some thoughts of what I would like to see happen in future research. There is huge potential to build on micro studies of aptitude linked to education, such as those covered here by Erlam (2005) and De Graaf (1997). Not only for more such research to be done (and see Li (2015) for review), but also for there to be more systematicity in this research, so that there is more organised coverage of variables such as specific language focus, difficulty, proficiency level of participants, and so on. There is also scope to develop what has been achieved in the 'facets' section. I would particularly like to see more research into the areas of language structure and memory, such as new tests of inductive language learning, and broader aspects of memory. There is great potential in these areas.

Finally, there is a pressing need for a validated, easily-available and comprehensive aptitude battery. Of course, possible starting points for this would be the HiLAB entering the public domain (supplemented, perhaps, by the MLAT), or alternatively LLAMA being validated successfully. But I suspect such an aptitude test will need to be a new construction, not least to incorporate the sorts of exciting findings that have emerged from the research into aptitude facets. Such a development would be the most relevant vehicle for deeper understanding of aptitude, its connections with language learning, and a deeper understanding of language itself.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30(4), 481–509. doi:10.1017/s027226310808073x

Bokander, L. (2023). Exploring the predictive validity of the LLAMA language aptitude tests: A research synthesis. In E. Wen, P. Skehan & R. L. Sparks (Eds.), *Language aptitude theory and practice* (pp. 94–116). Cambridge University Press. doi:10.1017/9781009076463.007

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). University of Pittsburgh Press.

Carroll, J. B., & Sapon, S. (1959). *Modern language aptitude test (MLAT)*. The Psychological Corporation.

De Graaf, R. (1997). The eXperanto Experiment: Effects of explicit instruction on second language development. *Studies in Second Language Acquisition*, 19(2), 249–276. doi:10.1017/s0272263197002064

DeKeyser, R. (2000). The robustness of critical period effects. *Studies in Second Language Acquisition*, 22(4), 499–533. doi:10.1017/s0272263100004022

Doughty, C. J. (2019). *Cognitive language aptitude. Language Learning*, 69, 101–116. doi:10.1111/lang12322

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*(2), 147–171. doi:10.1191/1362168805lr161oa

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, *36*(3), 385–408. doi:10.1093/applin/amu054

Li, S. (2019). Six decades of language aptitude research: A comprehensive and critical review. In Z. Wen, P. Skehan, A. Biedroń, S. Li & R. L. Sparks (Eds.), *Language aptitude theory and practice* (pp. 78–96). Routledge. doi:10.4324/9781315122021-5

Li, S., & De Keyser, R. (2021). Implicit language aptitude: Conceptualising the construct, validating the measures, and examining the evidence: Introduction to the special issue. *Studies in Second Language Acquisition*, *43*(3), 473–497. doi:10.1017/s0272263121000024

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*(3), 530–566. doi:10.1111/lang.12011

Meara, P. (2005). *LLAMA language aptitude tests*. Lognostics.

Perruchet, P. (2021). Why is the componential construct of implicit language aptitude so difficult to capture? *Studies in Second Language Acquisition*, *43*(3), 677–691. doi:10.1017/s027226312100019x

Rogers, V., Meara, P., & Rogers, B. (2023). Testing language aptitude: LLAMA evolution and refinement. In E. Wen, P. Skehan & R. L. Sparks (Eds.), *Language aptitude theory and practice* (pp. 47–72). Cambridge University Press. doi:10.1017/9781009076463.005

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing the auditory precision hypothesis-L2. *Language Teaching*, *56*(4), 522–538.

Sparks, R., & Dale, P. (2023). The prediction from MLAT to L2 achievement is largely due to MLAT assessment of underlying L1 abilities. *Studies in Second Language Acquisition*, *45*(5), 1345–1369.

Turker, S., Seither-Preisler, M., & Reiterer, S. M. (2021). Examining individual differences in language learning: A neurocognitive model of language aptitude. *Neurobiology of Language*, *2*(3), 389–415. doi:10.1162/nol_a_00042

Wen, Z. (2019). Working memory as language aptitude: The phonological/executive model. In Z. Wen, P. Skehan, A. Biedroń, S. Li & R. L. Sparks (Eds.), *Language aptitude theory and practice* (pp. 187–214). Routledge. doi:10.4324/9781315122021-10