Lang. Teach. (2017), **50.1**, 1–31
doi:10.1017/S0261444816000276

State-of-the-Art Article

Foreign language aptitude theory: Yesterday, today and tomorrow

Zhisheng (Edward) Wen Macao Polytechnic Institute, Macau, China
edwardwen@ipm.edu.mo

Adriana Biedroń Pomeranian Academy in Słupsk, Poland
adriana.biedroń@apsl.edu.pl

Peter Skehan St Mary's University College, Twickenham, London, UK
peterskehan@gmail.com

Foreign language (FL) aptitude generally refers to a specific talent for learning a foreign or second language (L2). After experiencing a long period of marginalized interest, FL aptitude research in recent years has witnessed renewed enthusiasm across the disciplines of educational psychology, second language acquisition (SLA) and cognitive neuroscience. This paper sets out to offer a historical and an updated account of this recent progress in FL aptitude theory development and research. As its subtitle indicates, the paper centres on three major issues: following the introduction and clarification of basic concepts, Section 1 traces the early conceptions of FL aptitude dominated by John Carroll's pioneering work. Section 2 summarizes and examines more recent theoretical perspectives and FL aptitude models proposed by researchers from multiple disciplines that have significantly broadened the conventional research traditions associated with Carroll's original conception. Based on the research synthesis of current FL aptitude models, Section 3 suggests the directions FL aptitude theory and research might take in coming years. We conclude that a working memory perspective on FL aptitude presents one promising avenue for advance, as does the development of new aptitude tests to predict speed of automatization, implicit learning and greater control over an emerging language system. In addition, it is argued that issues of domain-specificity versus domain-generality for aptitude tests may lead to aptitude theory and research becoming more central in applied linguistics.

1. Introduction

Foreign language (FL) aptitude generally refers to the specific talent for learning a foreign or second language (L2) (Carroll 1981; Skehan 2002). It is a concept deeply rooted in educational psychology and its interpretation in applied linguistics is unavoidably affected by

developments in the neighbouring fields of education and psychology. In recent years, after decades of rapid developments within the cognitive revolution (Carroll 1993; Miller 2003), knowledge of human cognitive abilities has greatly expanded owing to new discoveries in the multiple disciplines of the cognitive sciences, and in particular cognitive psychology and cognitive neuroscience. As a result, it can be argued that a modern-day discussion of the role of FL aptitude in applied linguistics would be seriously impoverished if research progress from neighbouring disciplines were not taken into account (see Biedroń 2015, for example, for a preliminary discussion of the potential contribution of neurology to FL aptitude). Driven by these concerted efforts from educational psychology, applied linguistics and cognitive science, the concept of FL aptitude has undergone significant modifications since its inception, and is continually evolving (Wen 2012a; Granena 2013; Skehan 2015a). To reflect on these advances in FL aptitude theory construction and testing, the present paper sets out to trace these developments in the first section, followed, in the second section, by a discussion of more current research efforts to re-conceptualize the construct, often of a more focused nature. Based on these findings, the final section will offer suggestions for the future of FL aptitude research, both theoretical and methodological.

1.1 Clarification of basic concepts: Aptitude vs ability

Before exploring the construct of FL aptitude, it is necessary to define the basic concepts connected with it. To begin with, 'FL aptitude' should be considered an umbrella term consisting of a set of COGNITIVE abilities, thus making it a COMPONENTIAL concept (Carroll 1981, 1993; Skehan 1998; Sparks et al. 2011). As Dörnyei (2005: 33) pointed out, FL aptitude has increasingly become something of a hybrid construct related to a number of cognitive factors creating a composite measure regarded as the general capacity to master an L2 (cf. DeKeyser & Koeth 2011). At the outset, a distinction should be made between aptitude and ability. In this respect, the 'founding father' of aptitude research, John Carroll (1993), clearly differentiated the terms 'ability', 'aptitude' and 'achievement'. He identified ability with performance or potential for performance. As used to describe an attribute of individuals, 'ability' refers to 'the possible variations over individuals in the liminal levels of task difficulty [. . .] at which, on any given occasion in which all conditions appear favorable, individuals perform successfully on a defined class of tasks' (1993: 8). Carroll emphasized stability of abilities: 'An ability can be regarded as a trait to the extent that it exhibits some degree of stability or permanence even over relatively long periods of time' (1993: 7).

In this sense, 'aptitude' is a notion that is synonymous with 'ability'. To use the words of Dörnyei: 'Although some scholars distinguish between ability and aptitude, in typical practice the two are used synonymously' (2005: 32). The difference between the terms is that they are used in different contexts rather than with a different meaning. Carroll defined FL aptitude as 'an individual's initial state of readiness and capacity for learning a foreign language, and probable facility in doing so [given the presence of motivation and opportunity]' (1981: 86). As such, he regarded aptitude as a sort of ability, namely a latent trait that is relatively stable and relatively resistant to training, and which refers to the potential for achievement provided instruction is optimal. Therefore, an ability is an aptitude if it predicts the rate and speed of

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

learning. Although Carroll distinguished between 'achievement' and 'aptitude', he admitted that in some cases measures of achievement can be viewed as measures of aptitude to the extent that they may predict future learning progress.

1.2 Clarification of basic concepts: Aptitude vs intelligence

Another notion that needs clarification is the relationship between FL aptitude and general intelligence (*g*). In this respect, results from Sasaki's (1996) study are relevant. The objective of her study was to investigate to what extent, and in what aspects, language aptitude is related closely to general intelligence: further, in what aspects language aptitude is language-specific. The relationships between three measures were explored: namely, second language proficiency, FL aptitude and two types of intelligence (verbal intelligence and reasoning). It was found that first-order factor analysis of the aptitude and intelligence scores confirmed they were separate. However, a second-order factor analysis corroborated the existence of a common factor, namely analytic ability, which accounted for the variance in some of the aptitude variables as well as that in the intelligence quotients (IQ). Other aptitude factors in Carroll's original model (1962, 1981), such as phonetic coding ability and memory, did not correlate with intelligence. Drawing on these results, Sasaki concluded that intelligence and analytic ability are interrelated, whereas phonetic coding ability and memory factors are separate components of FL aptitude, independent of the *g* factor (cf. Grigorenko, Sternberg & Ehrman 2000). Even so, there is still a marked lack of research on the relationship between different FL aptitude components and other primary and higher level cognitive abilities. These issues will be discussed further in later sections.

2. Yesterday: Early conceptions of and research on FL aptitude

No review of FL aptitude research can afford to bypass the contributions from the American educational psychologist John Carroll and his early work conducted during the 1950s and 1960s (Wen & Skehan 2011; Skehan 2012). First and foremost, the Modern Language Aptitude Test (MLAT) battery developed by Carroll and his colleague Stanley Sapon (Carroll & Sapon 1959/2002) set a benchmark for all subsequent aptitude measures (e.g. Pimsleur's PLAB; Green's York Language Aptitude Test; Petersen & Al-Haik's DLAB; Parry & Child's VORD). Additionally, based on factor analyses of these MLAT sub-tests, Carroll (1962, 1981, 1990, 1993) conceptualized the FL aptitude construct as comprising four distinct and measurable abilities: PHONETIC CODING ABILITY, GRAMMATICAL SENSITIVITY, INDUCTIVE LANGUAGE LEARNING ABILITY and ASSOCIATIVE MEMORY (see Table 1). This conceptual framework of FL aptitude, as Skehan (2002) cogently puts it, has proved to be more enduring than even the MLAT test battery itself and to date it is still the most influential portrayal of FL aptitude available (Skehan 2012).

The majority of research on FL aptitude following the publication of the MLAT has been influenced substantially by the two contributions from Carroll: theorizing the structure of

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

Table 1 Carroll's four-factor aptitude model (based on Dörnyei & Skehan 2003)

Aptitude components	Definitions
Phonetic coding ability	Capacity to code unfamiliar sound so that it can be retained
Grammatical sensitivity	Capacity to identify the functions that words fulfil in sentences
Inductive language learning ability	Capacity to extrapolate from a given corpus to create new sentences (not measured in the MLAT)
Associative memory	Capacity to form associative links in memory

aptitude and developing a comprehensive aptitude test battery. More specifically, these studies can be summarized and re-grouped into five broad categories as indicated in Table 2.

When these studies are put into perspective, two conclusions emerge. First, although there have been some criticisms and skepticism, the concept of FL aptitude has stood the test of time and remained a valid concept despite the big changes that have happened in classroom practice from structure-based methods, such as audiolingualism, to the more current communicative language teaching (Erlam 2005). More importantly, this body of research has shown that the role of language aptitude is not just confined to traditional formal/instructed settings (as originally postulated by Carroll 1962), but is also viable under different learning conditions and in different learning contexts (Wesche 1981; Reves 1982), as well as in today's communicative classrooms (Robinson 1996, 2005; Erlam 2005; Vatz et al. 2013). Indeed, a recent meta-analysis conducted by Li (2015) of 34 sampled FL aptitude studies (N = 3239) in the past 50 years (1963–2013) has produced a statistically significantly positive correlation (r = 0.34) between MLAT scores and ultimate L2 learning outcomes among children and adults.

Second, despite such positive interpretations, the actual number of empirical studies (Table 2) is relatively limited, which indicates a clear lack of interest among researchers in this topic until recent years (Li 2015). Such marginalized interest in what is still accepted as the single best predictor of final language learning outcome stands in sharp contrast with the enormous research efforts in other areas of language learning and acquisition (Dörnyei & Skehan 2003). A case in point is the intensive research enthusiasm directed towards its major counterpart as an individual difference (ID) variable, i.e. L2 motivation. Indeed, previous empirical studies have repeatedly pointed to the fact that among all ID factors (except for the indisputable age factor; see Skehan 1989, 1998; Dörnyei 2005; Granena & Long 2013 for detailed reviews) that are purported to influence L2 learning, only L2 motivation can manage to perform on a par with FL aptitude in predicting ultimate learning outcomes (Ehrman & Oxford 1995; Ehrman 1998; Skehan 1998, 2002; Masgoret & Gardner 2003). Curiously, therefore, these two equally important ID factors have met with vastly different fates in SLA.

The consequence of such imbalanced research interest (i.e. between language aptitude and L2 motivation) is not difficult to discern. Over the recent few decades (roughly from the 1980s until now), conceptualization of L2 motivation theory has undergone several paradigm shifts (Dörnyei 2010), starting from Robert Gardner's classical Instrumental/Integrative motivation model (Gardner 1985), to Dörnyei's more recent Process Motivation model (Dörnyei 2001)

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

Table 2 Summary of post-MLAT aptitude research (based on Wen & Skehan 2011)

Orientations	Representative studies	Characterizations and implications
Assessment-oriented	Pimsleur's PLAB 1966; Petersen & Al-Haik's DLAB 1976; Parry and Child's VORD 1990; Grigorenko et al.'s CANAL-FT 2000; Meara's LLAMA language aptitude test 2005; Doughty et al.'s Hi-LAB 2010	Psychometric in nature; mostly MLAT alternatives or complementary tests; sometimes associated with military or government initiatives and funding
Component-targeted	Skehan 1982, 1986b; Sparks & Ganschow 1991; Sasaki 1996; Miyake & Friedman 1998; Ranta 2002; Sparks et al. 2011; Winke 2013	Concept of factor-components is still viable; relatively little research has been conducted; much room for development (particularly with memory)
Aptitude and age	Harley & Hart 1997, 2002; Bongaerts, Mennen & van der Silk 2000; DeKeyser 2000; Birdsong 2007; Abrahamsson & Hyltenstam 2008; DeKeyser, Alfi-Shabtay & Ravid 2010; Granena & Long 2013	Younger learners tend to show higher correlations with memory components and older learners with analytical components; aptitude mitigates putative critical period effects and is associated with high achievement levels
Aptitude-treatment interaction: macro studies	Wesche 1981; Reves 1982; Robinson 1995, 2002; Erlam 2005; Sawyer & Ranta 2001; Hwu & Sun 2012	Aptitude profile information based on different aptitude components has wide-ranging pedagogical implications for different L2 learning conditions
Aptitude, instruction, and feedback: micro studies	Robinson 1995; De Graaff 1997; Van Patten & Borst 2012a, 2012b; Trofimovich, Ammar & Gatbonton 2007; Yilmaz 2013; Sheen 2007a, 2007b; Li 2011; Hwu & Sun 2012	Aptitude is associated with greater success in short-term studies with focused grammar points; more strongly with explicit than implicit instruction and feedback

and to the newly proposed L2 Self and Dynamic Systems theory (DST) model (Dörnyei 2009, 2010, 2014; Dörnyei, MacIntyre & Henry 2015) and the latest theory of directed motivational currents (DMCs) (Dörnyei, Henry & Muir 2016). In sharp contrast, FL aptitude has had a less dynamic history, as Skehan suggested, back in 2002, that relatively 'little empirical work' and 'little theorizing' has taken place, with this perhaps linked to the perceived immutability of aptitude, coupled with views that it is confined in relevance to outdated methodologies in language education (also see Skehan 1998).

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

As a result of this marginalized interest, knowledge about FL aptitude has not developed very much since it started some 50 years ago. The concept has remained intact – a relatively fixed trait that is not subject to malleability by later learning experience – and until the twenty-first century the research conventions barely changed (pre-course administration of an aptitude test such as the MLAT, then correlated with post-course achievement scores). This design, termed 'macro' by Skehan (2015a), hopefully could predict the likely learning rate and outcomes (also see Aguado 2012). Viewed this way, it seems fair to claim that FL aptitude theory has lagged far behind L2 motivation theory. Therefore, the need for theoretical and methodological advances in FL aptitude research is imperative indeed (Skehan 2002, 2012, 2015a; Aguado 2012; Kormos 2013; Singleton 2014; Li 2015), and attempts to use the Hi-LAB, in conjunction with the MLAT, for more diagnostic and remedial purposes are encouraging (Jackson & Doughty 2015). Even if aptitude is not particularly malleable, it should still be possible to modify instruction to achieve greater learning by responding to different aptitude levels, as well as strengths and weaknesses (Jackson et al. 2015).

3. Today: Current theoretical perspectives and models of FL aptitude

Fortunately, the bleak scenario for FL aptitude research began to improve slightly at the start of the twenty-first century. In many ways, the anthology edited by Peter Robinson (2002) can be regarded as a turning point in the re-conceptualization of the construct of FL aptitude. In this new wave of research and theorizing, several important lines of enquiry have begun to emerge, giving rise to a number of innovative perspectives on FL aptitude that go beyond the early conception by Carroll (Ellis 2004; Wen 2012a). In this section we will summarize several major theoretical accounts of FL aptitude from the disciplines of educational psychology, applied linguistics, cognitive science and cognitive neuroscience, with a view to demonstrating how these multiple perspectives and theoretical models have broadened and expanded the research paradigm set forth by Carroll.

3.1 The learning difficulties perspective and the Lingusitic Coding Differences Hypothesis (LCDH) model

Inspired by their own earlier studies of L1 literacy research investigating learning difficulties, the two educational psychologists Richard Sparks and Leonore Ganschow, together with their colleagues (1991, 2001; see also Sparks, Ganschow & Patton 1995, 2008), proposed the LCDH. Supported by a series of empirical studies, the basic premise of the LCDH model lies in the argument that native language (L1) literacy skills are essential for predicting L2 learning. For example, if a student experiences difficulties in L1 phonology/orthography, his/her subsequent L2 learning will likely suffer as well (Sparks & Ganschow 2001: 97). As was suggested by a recent factor-analytical study (Sparks et al. 2011), four basic components of L2 aptitude, including students' L1 and L2 phonology/orthography skills (subsuming phonemic coding and phonological processing ability); their L1 and L2 language analysis

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

skills (comprising comprehension, grammar, vocabulary, and inductive language learning); their IQ/memory skills (including L1 intelligence and L2 paired-associate learning measures); and self-perceptions of L2 motivation and anxiety, combined to explain 76% of the variance in ultimate oral and written L2 proficiency.

Therefore, in developing the LCDH, Sparks & Ganschow further argue that it is necessary to examine the similarities and the differences between the two languages in question (participants' L1 and L2). The issue is particularly important regarding potential negative transfer effects from morphosyllabic languages such as Chinese to alphabetic languages such as English and vice versa. The researchers propose to include a certain phonological measure of L1 and L2 in FL aptitude tests (Sparks & Ganschow 2001: 100). In this sense, the LCDH aptitude model as advocated by Sparks and colleagues further complements Carroll's original four-factor view of FL aptitude by adding the extra (sub-)component of 'L1 and L2 phonology/orthographic decoding skill' as an important factor of aptitude (see Chan, Skehan & Gong 2011 for related suggestions). Arising from this model there is a call for greater depth in cross-linguistic analyses between the relevant L1 and L2 in future research of FL aptitude.

3.2 The successful intelligence perspective and the CANAL-F model

Grigorenko, Sternberg & Ehrman (2000), drawing on Sternberg's (1997, 2002) triadic conception of human intelligence (i.e. his so-called 'successful intelligence' perspective in which intelligence is conceived as comprising three distinct levels: analytical, creative and practical), proposed a new interpretation of FL aptitude: the Cognitive Ability for Novelty in Language Acquisition-Foreign (CANAL-F) theory. This theory emphasizes the ability to handle novelty and ambiguity when learning an L2. In line with this conceptualization, the authors also devised a new approach to assessment procedures for measuring FL aptitude, i.e. the CANAL-F test.

Specifically, the CANAL-F test focuses on measuring test takers' recall and inferencing ability to process and acquire new linguistic materials under both immediate and delayed conditions (as cited in Dörnyei & Skehan 2003: 595). The test is administered to the participants in a simulated, naturalistic language learning environment where they are gradually introduced to an artificial language (Ursulu) and then instructed to complete several small learning tasks. The test is postulated by its authors to tap into five knowledge acquisition processes, which include selective encoding, accidental encoding, selective comparison, selective transfer, and selective combination (Grigorenko et al. 2000: 392). These cognitive processes are operationalized at the lexical, morphological, semantic and syntactic levels of language and include both visual and oral input and output materials.

The CANAL-F test contains nine sections. The first section involves learning the meanings of neologisms from context, where 24 short paragraphs are presented to participants (orally and visually) and the participants have to guess which of five English alternatives corresponds to each unknown neologism in the paragraph. The second section involves comprehension of the whole passage. The third section deals with continuous paired-associate learning that involves the learning of 60 word-pairs (presented visually and orally). Participants need to produce the correct paired-associate in English (one half) and in Ursulu (the other half). In

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

the fourth section, 20 sets of three to five Ursulu sentences are presented to participants (half visually and half orally) together with their English translations. After this, the participants have to select the best translation out of five for a completely new sentence (in both directions of translation). The final, fifth, section involves learning language rules. Participants are provided with vocabulary, grammar and examples of the workings of Ursulu and are expected to generalize the most salient rules of the language. This learning is then tested with 12 items measuring understanding of Ursulu.

As claimed by the authors (Grigorenko et al. 2000), the CANAL-F test distinguishes itself from previous aptitude tests in five significant ways: (a) it is cognitively based and theoretically oriented; (b) it is contextualized; (c) it is situated and dynamic since different sections interrelate and build cumulatively; (d) it is multi-functional; and (e) it is adaptive. Furthermore, in the CANAL-F test, language learning is believed to involve encoding knowledge in working memory, and storing in and retrieving from long-term memory, all of which are assessed by the immediate and delayed recall tasks. Above all, its emphasis on cumulative learning and assessment should be able to shed new light on designing aptitude tests beyond the traditions of MLAT (Skehan 2015a). Unfortunately, despite the bold move of the authors to offer a theory-driven view on FL aptitude, validation results of the CANAL-F test, as obtained in their empirical studies, did not significantly outperform the MLAT in predictive validity (Sternberg & Grigorenko 2002).

3.3 The information processing perspective and the Macro-SLA aptitude model

In contrast to previous attempts to theorize FL aptitude which have treated the concept as self-contained and unrelated to broader issues in SLA, Skehan (2002, 2012, 2015a; also see Dörnyei & Skehan 2003; Chan et al. 2011) proposed an aptitude model that builds on developments from accumulating SLA research. In this SLA-compatible aptitude model (as shown in Table 3), Skehan stipulated that different putative components of aptitude should be effectively linked to various SLA developmental stages and their associated cognitive processes. For example, it can be argued that the aptitude components of 'phonetic coding ability' and working memory are most likely to be related to the initial stage of input processing and to noticing. As for language analytical ability (a combination of Carroll's original 'grammar sensitivity' and 'inductive language learning' as well, again, as working memory), Skehan believes they should be related to the stages of pattern identification and restructuring, and extending.

Recently, Skehan (2016) has related these stages to the burgeoning 'micro' literature on language aptitude which avoids the pre-post correlational design typical of the past and instead focuses, usually with relatively brief quasi-experimental designs, on different types of instruction or of feedback (Skehan 2015a). Such studies show more clearly, and at close to a process level, the impact of aptitude variables such as working memory, or the capacity to benefit from different types of feedback, for example. Skehan (2015a) argues that focal grammar points which have salience, for example, with inversion, or redundancy, with the use of articles, benefit differentially from feedback for higher aptitude learners. In a similar vein, as Ellis (2012: 313) cogently points out, it can be also hypothesized that L2 learners

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

Table 3 Skehan's Macro-SLA aptitude model (based on Skehan 2016)

SLA stages	L2 cognitive processes	Aptitude constructs
Language input Central processing Language output	Input processing (segmentation)	*Attentional control* *Working memory*
	Noticing	Phonetic coding ability *Working memory*
	Pattern recognition	Phonetic coding ability *Working memory* Language analysis ability
	Complexification	Language analysis ability *Working memory*
	Handling feedback	Language analysis ability *Working memory*
	Error avoidance	*Working memory* *Retrieval memory*
	Automatization	*Retrieval memory*
	Creating a repertoire	*Retrieval memory* *Chunking*
	Lexicalization	*Chunking*

Note: Aptitude constructs in italics are new components as opposed to John Carroll's classic model.

strong in phonetic coding ability will likely benefit from input-based forms of instruction, while those strong in language analytical ability will fare better with consciousness-raising instructions.

In its more recent development, Skehan (2016) further argues that there is something of a split between the first five stages, which are, broadly, concerned with knowledge acquisition, and the remaining stages, which are concerned with control over that knowledge in actual performance, as well as the capacity to sound more native-like. These bring out more clearly the issue of whether greater control is the result of implicit learning taking place, or the automatization to enable very rapid use of explicit learning from the earlier stages (DeKeyser 2007).

Skehan's conception of FL aptitude can have significant implications for theoretical advancement. First, instead of solely relying on an aptitude score to predict L2 learning outcomes (typical in Carroll's time), Skehan opts for an approach that serves to EXPLAIN the underlying causes for the more superficial IDs in the final learning outcome of an L2. This change in approach represents a significant move in that it triggers a revolution in research paradigms of FL aptitude (i.e. from 'predicting' to 'explaining'). Then, besides this enhancement in EXPLANATORY power, Skehan also proposes several additional aptitude components that are presumably functioning within various L2 cognitive processes (but were lacking in Carroll's original model), which can in turn impact L2 learners' performance in corresponding developmental stages of SLA. Despite its theoretical appeal, Skehan's

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

conception of FL aptitude remains at a speculative stage and, therefore, further empirical support is necessary before this model can be fully operationalized and put into practice (but see Skehan 2016, for relevant discussion). For example, it still lacks a comprehensive account of the nature of the putative aptitude components (e.g. attentional control, working memory and memory retrieval) as well as their fine-grained functioning within SLA processes and domains. In addition, the model may need to provide more articulate specifications of the measures and procedures for assessing these putative aptitude components in practical research.

3.4 Aptitude-treatment interaction and the Aptitude Complexes framework

Another recent proposal to re-conceptualize FL aptitude in applied linguistics is Robinson's (2005, 2007, 2012) 'Aptitude Complexes/Ability Differential' framework. As its name clearly denotes, there are two closely interlocking hypotheses incorporated in the framework. The first, the Aptitude Complexes Hypothesis (based on the ideas of 'aptitude complexes' by Snow 1987, 1994) claims that a set of basic cognitive abilities (such as 'processing speed', 'pattern recognition', 'phonological working memory capacity,') combine to form higher-order aptitude complexes (such as 'noticing the gap' and 'memory for contingent speed') that are being drawn on in learning a certain task. The second, the Ability Differentiation Hypothesis claims that L2 learners demonstrate variations within the set of cognitive abilities (with strengths in some of them and weaknesses in others), thus leading to differentiated profiles in corresponding aptitude complexes. The overall schematic representation of Robinson's framework resembles a wheel-shape that consists of several layers of circles embedded within each other (Robinson 2005: 52), where the inner circle consists of 'core' aptitudes or abilities, such as grammatical sensitivity, processing speed, and so on, and the next circle is concerned with aptitude complexes, which relate the core abilities to various learning processes. Then, moving outwards, the remaining two circles are concerned with task aptitudes and finally pragmatic/interactional abilities/traits. These are concerned with more directly instructional factors and contexts (task aptitudes) and broader communicational influences and capacities; these will not be developed here.

The heart of the model is in the two inner circles. The first, abilities, represents a view of fundamental aptitudes, similar in conception to the work of other researchers covered here, although slightly different in detail and emphasis. The next circle, aptitude complexes, represents the most innovative contribution to aptitude theorizing. More specifically, there are four 'aptitude complexes' in the framework. They are namely, an aptitude for a focus on form; an aptitude for incidental learning, via oral content; an aptitude for incidental learning via written content; and an aptitude for explicit rule learning. We will describe the first in greater detail, to clarify how the different variables interact. As demonstrated in Figure 1 (see also Ellis & Shintani 2013), an aptitude for focus on form is influenced by (a) capacity to notice and (b) by memory for contingent speech, and then each of these can have high or low values, potentially generating four combinations: high noticing and high memory for contingent speech, low noticing but high memory for contingent speech, and so on. The other aptitude complexes similarly are influenced by two ability factors: incidental learning via oral content (memory for contingent speech and deep semantic processing) and via written content (deep

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

Aptitude complex	Aptitude for focus on form		Aptitude for incidental learning via written content	
	↗	↖	↗	↖
Ability factors	Noticing the gap	Memory for contingent speech	Deep semantic processing	Memory for contingent text
	↑	↑	↑	↑
Primary cognitive ability	Perceptual speed	Phonol.WM capacity	Analogies	WM for text
	Pattern recognition	Speed of PWM	Inferring word meaning	Speed of WM for text

Figure 1 Primary and second order abilities and aptitude complexes in Robinson's model (based on Robinson 2007)

Note: Phonol.WM, PWM = Phonological working memory; WM = Working memory.

semantic processing and memory for contingent text); and complex rule learning (memory for contingent text and metalinguistic rule rehearsal).

The first and third complexes (aptitude for focus on form; aptitude for incidental learning via written content) are represented visually in Figure 1. In this way Robinson is proposing that relevant 'aptitude' is not always the same, and that content of learning, or instructional conditions interact with the aspects of aptitude that are most relevant. In effect therefore Robinson is offering hypotheses about the range of different contexts that we need to consider, the basic range of aptitude sub-components, and then how different sub-components may interact with one another in different learning contexts. There is, as yet, little data on how these hypotheses have been tested.

In a sense, Robinson's conception of FL aptitude is similar to Skehan's model in that it also seeks to provide an explanation that aims to account for observable IDs in the final learning outcome among L2 learners. However, it slightly differs from Skehan's approach in that it tries to provide a finer-grained representation of the complex structure of human cognitive abilities (Carroll 1993) and also makes clearer reference to specific educational contexts. In particular, Robinson presents the framework as involving several embedded levels of analysis. For example, basic cognitive abilities belong to the most deeply embedded and fundamental level, while aptitude complexes represent less core, embedded abilities, and finally task aptitudes, as well as actual performance, are the least embedded and most real-world aspects of the model. There is a progression, in other words, from fundamental and general abilities to more specific and variable educational contexts.

Besides depicting a more intricate picture of human cognitive abilities, Robinson's theoretical conception of FL aptitude differs from previous models with its most articulate emphasis on the DYNAMIC interactions between learner ability factors (either of lower-level core cognitive abilities or higher-level aptitude complexes) and the inherent task characteristics (or 'task complexity' in his own terms, Robinson 2005, 2007, 2012). Such aptitude-treatment interactions (ATIs), as Robinson proposes, take place at all levels, with their ultimate influence readily permeating learners' task performance and their real-world performance.

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

3.5 Cognitive science and the High Level Language Aptitude Battery (Hi-LAB) model

The most significant contribution to aptitude theory construction and testing in the last few years has been the development of the 'High-Level Language Aptitude Battery (Hi-LAB) by researchers at the University of Maryland Center for Advanced Study of Language (see Doughty et al. 2010; Doughty 2013, 2014; Linck et al. 2013). The motivation for this development has been a perceived need to develop an aptitude battery suitable for the prediction and explanation of high levels of L2 proficiency, targeted particularly at talented post-critical-period language learners. The authors raise the possibility that the components of aptitude suitable for high-level learning may be different from those at lower levels, defining high-level aptitude 'as a composite of domain-general cognitive abilities and specific perceptual abilities' (Linck et al. 2013: 535) that are relevant to high-level achievement. The Hi-LAB model is based on the constructs, with associate aptitude measures, shown in Table 4.

The rationale for this factor-based aptitude model and test is clearly rooted in contemporary cognitive science. In particular, there is a major role for working memory, and indeed a very richly interpreted conception of working memory (Doughty et al. 2010; Bunting & Engle 2015). The different functions of its central executive are probed with care, and then phonological short-term memory has two measures associated with it. Several measures of longer-term memory are also involved, along with the more traditional associative memory as well as indicators of retrieval, implicit learning, and processing speed (with the latter two measured by different facets of the same serial reaction test). Finally, there are two measures of perceptual acuity, for phonemic discrimination and categorization.

At a construct validity level this is a very impressive test and it is likely to be a milestone for high-level aptitude testing for some time to come. At present, however, many projects related to the Hi-LAB model are still in progress (Doughty et al. 2010), so there is only limited

Table 4 Constructs and measures in Hi-LAB (based on Doughty et al. 2010)

Constructs	Measures
Working memory	
Executive functioning	
Updating	Running memory span
Inhibitory control	Antisaccade
Task switching	Task-switching numbers
Phonological short-term memory	Letter span
	Non-word span
Associative memory	Paired associates
Long-term memory retrieval (priming)	ALTM synonym
Implicit learning	Serial reaction time
Processing speed	Serial reaction time
Auditory perceptual acuity	Phonemic discrimination: Hindi, English
	Phonemic categorization: Russian

Note: ALTM = Available long-term memory.

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

empirical evidence available to the public. Notwithstanding this, Linck et al. (2013) do report an important study where they used Hi-LAB to try to separate a group of 'successful' from a group of 'very successful' learners. The different sub-tests from Hi-LAB were then examined to see how effective they were in distinguishing between the two groups. The most effective sub-tests in this regard were the Paired Associates test, the Letter Span of phonological short-term memory test, (the Non-word Span was marginal) and the Implicit Learning test. To put it another way, the tests of executive working memory did not distinguish between the groups, nor did processing speed or the auditory tests. This evidence is intriguing and encouraging, but not conclusive in terms of validation for the Hi-LAB. What is needed most, of course, is more evidence from additional groups of language learners. For now, we will simply observe that the tests which seemed more effective here tended to probe long-term memory (consistent with Skehan's 1998 speculation that memory is more important at higher levels of achievement) and also tended to be tests with more prominent language-specific material (for example, the letter span test was more effective at separating the groups than the non-word span test). Overall, with its strong construct validity rooted in contemporary cognitive psychology, the Hi-LAB model will have enormous implications for FL aptitude theory building and testing development in the future.

3.6 Cognitive neuroscience and brain networks as FL aptitude

Recent years have also witnessed cognitive neuroscience research paradigms and techniques foraging into the territory of SLA (Gullberg & Indefrey 2006; Dörnyei 2009; Dien et al. 2011). Findings from these brain-based neurological studies have increasingly become an important source of evidence that complements the earlier dependence on behavioural research conventions (Schumann 2004; Li & Grant 2015). Among the ID factors outlined in most textbooks (e.g. Skehan 1998; Dörnyei 2005, 2009), the number of empirical studies on the neurological substrates of FL aptitude is growing, producing accumulating evidence that is consistent and replicable across studies (Biedroń 2015; also see Li, Legault & Litcofsky 2014). Broadly speaking, these brain-based neurological explorations can be divided into two types: those that focus on the STRUCTURAL aspects of brain anatomy and those connected with FUNCTIONAL aspects of brain activation. The anatomical or activation differences are then associated with distinct aspects and levels of L2 learning and processing. In terms of research methodology, most of these neurolinguistic explorations have used a series of hemodynamic methods such as positron emission tomography (PET; measures radioactive tracers), event related potential (ERP; measures temporal aspects of neural events), functional magnetic resonance imaging (fMRI; measures the magnetic resonance signals) and magnetoencephalography (MEG; measures the magnetic resonance signals as a result of electrical activity), and finally diffusion tensor imaging (DTI; maps white matter fiber tracks of the brain) (see Biedroń 2015: 18, Table 1 for a summary).

In terms of brain anatomy and FL aptitude, the phonological aspect (i.e. pronunciation) of an L2, and in particular, phonetic perception is believed to be the most thoroughly investigated area within these neurolinguistic perspectives (Biedroń 2015). Amid these strands of research, the most noticeable findings came from Golestani and colleagues (Golestani & Zatorre 2009;

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

Golestani, Price & Scott 2011; Golestani 2014; see also Hu et al. 2013), who investigated brain structure among talented phoneticians and found that they tend to have greater amounts of grey matter and white matter in the parietal regions of the left hemisphere of their brains. This area (left parietal cortex), generally assumed to be pertinent to auditory and articulatory aspects, is believed to sub-serve phonological working memory as well. Indeed, these strands of studies suggest that the brain anatomy feature of phonological working memory in the left auditory cortex can be a reliable predictor of phonological aptitude to a certain extent, which is compatible with Baddeley, Gathercole & Papagno's (1998) postulation of the phonological loop of working memory as a 'language learning device'.

In terms of brain activity and FL aptitude, the study by Tan et al. (2011) is revealing. In this study, the researchers demonstrated through fMRI that bilingual learners' later skills in L2 reading can be predicted by the activation level of the fusiform–caudate circuit in the participants' left hemisphere over a one-year span. The left caudate and fusiform regions are usually thought to mediate language control functions and likely are involved in resolving competition arising from L1 during L2 use. This finding thus suggests that the activity level of these brain regions serves as an important neurobiological marker for predicting FL aptitude in the area of L2 reading.

Finally, a recent trend in cognitive neuroscience research in the last decade has been a shift from focusing on tapping functions of individual brain regions to pinning down the spatial and temporal dynamics of brain connectivity networks (Li & Grant 2015). A series of studies conducted by Li and colleagues in their lab at Pennsylvania State University have amply demonstrated that brain connectivity networks can indeed serve as a reliable predictor for learning both L2 novel words and (artificial) grammar/sequence rules in different learning contexts (natural, virtual etc.). For example, Yang et al. (2015) have found that L2 learning success was not only associated with better auditory perception ability (as already found in some of Golestani's brain-imaging studies) but also with more efficient and more flexible brain connectivity detected among more successful learners as opposed to their less successful counterparts. Most interesting of all, even within the learner group that received training, distinct brain connectivity patterns were also observed as a function of explicit vs implicit instruction of the learned grammar/sequence rules. These results echo an earlier study (Yang & Li 2012) which found that the explicit learners engaged more of a network that relies on the insula as a relay station, whereas their implicit counterparts evoked a more direct frontal-striatal network in language learning and processing. As we will argue later, it is this latter finding of the distinctive brain networks underlying explicit-implicit learning that represents a promising research paradigm for advancing FL aptitude theory construction and testing in the future (also see Xiang et al. 2012).

3.7 Summary of current theoretical perspectives on FL aptitude

To sum up this section, compared with the earlier conception of FL aptitude (in Carroll's time) and its conventional research paradigm (pre-post correlations), recent theoretical perspectives and models of language aptitude as reviewed here have contributed to a deeper understanding of the concept, each with their unique contributions. To begin with, the

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

LCDH model as proposed by educational psychologists Sparks & Ganschow (2001) has complemented Carroll's four-factor model by emphasizing the importance of L1-L2 language analysis skills (which echoes an earlier call by Skehan 1986a) and the role of cross-linguistic phonology/orthography analysis (Sparks et al. 2011). These clarify the limitations of Carroll's original conception. In addition, Sternberg & Grigorenko's CANAL-F theory and test have further advanced Carroll's MLAT by advocating the construction of more adaptive and situated aptitude tests (i.e. 'more dynamic' and 'contextualized', to use their own words). In a similar vein, the Hi-LAB model has significantly expanded the construct validity of FL aptitude in previous conceptions by relying on more current cognitive psychology research in theory construction and test development. It has also been the basis for ATI research (Vatz et al. 2013).

Next we have discussed two acquisition-based approaches from applied linguistics in which Skehan's aptitude model has underscored the importance of integrating language aptitude theories with established developments in SLA, while Robinson's Aptitude Complexes framework successfully pins down the dynamic interactions between aptitude profiles and task features and thus teases out their possible ramifications for L2 instruction and real-life performance. On the other hand, the brain-network-based cognitive neuro-scientific perspectives on FL aptitude have utilized the state-of-the-art technology that offers unique insights not affordable by conventionally psychometric and behaviour-oriented methods in previous aptitude research paradigms.

Most important of all, an in-depth analysis of these more contemporary aptitude models has given further support to a significant role for working memory in FL aptitude, either implicitly or directly. For example, the proponents of the LCDH and CANAL-F models have alluded to the concept of working memory. So, for both the two SLA-oriented proposals and the Hi-LAB model, the construct of working memory occupies a central place in their conception of FL aptitude. Last but not the least, the neurobiological perspectives (such as those studies conducted by Golestani and Yang) have produced the most compelling evidence for the critical role of phonological working memory in essential aspects of L2 learning, including phonology, vocabulary, and grammar/sequence rules (also see Xiang et al. 2012; Li & Grant 2015). For these reasons, in the next section we will argue for incorporating working memory as a core component of FL aptitude in the future.

4. Tomorrow: Re-conceptualizing FL aptitude and re-orienting future research

This section has four main parts. First, we explore the relevance of working memory for any comprehensive account of FL aptitude. The second section assesses the contribution and potential of 'micro' approaches to aptitude research, and what challenges are likely to be faced by such research in the future. Then we briefly consider the contrast between aptitude in relation to knowledge and control. Finally, we discuss two major theoretical issues, domain-generality vs domain-specificity and explicit vs implicit processing and learning. We conclude this section by arguing that future FL aptitude research exploring these two contrasts should have great potential to bring the concept back into the mainstream of SLA and applied linguistics research.

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

4.1 FL aptitude re-conceptualized from a working memory perspective

4.1.1 Rationale and empirical evidence

Recall that Carroll (1962) conceptualized FL aptitude as consisting of four ability factors (Table 1). Despite the importance of the four factors characterized in this early model, research studies exploring each of them in a focused way (the component-targeted section of Table 2) have been relatively scarce until quite recently (Li 2015). This is particularly so in the case of the memory component. It is small wonder that Carroll (1990), in his later review of aptitude research, cogently observed that there might be 'other types of memory' that could also be predictive of language achievement and merited further research. In retrospect, such a comment by Carroll was forward-looking indeed, particularly considering that Carroll was researching FL aptitude in the 1950s and 1960s when people's understanding of memory was rather limited.

We now have a flourishing research tradition exploring the role of working memory in SLA and aptitude (Wen & Skehan 2011; Biedroń 2012; Biedroń & Szczepaniak 2012; Wen 2014; Wen et al. 2015; Wen 2016). Central to this is the fractionated view of working memory, of a system containing a central executive, and various buffer systems (phonological, visuo-spatial, episodic) and the existence of IDs in each of these multiple components (Baddeley 2003, 2012, 2015; cf. Cowan 2015). For example, L2 learners have displayed individual variation in their phonological short-term memory as measured by the simple version of a memory span task (e.g. the digit span, the non-word repetition span task (Gathercole 2006)) and their central executive indexed by the complex version of a memory span task (e.g. the reading span, the operation span task; see Conway et al. 2005; Linck et al. 2014). Furthermore, numerous empirical studies in cognitive psychology and SLA have demonstrated that both the phonological and the executive aspects of working memory tend to exercise consistent and distinctive influences on various aspects of L2 acquisition and processing (e.g. Juffs & Harrington 2011; Williams 2012; Wen 2015, 2016). For instance, phonological short-term memory has been shown to be most relevant for the acquisition and development of vocabulary, formulaic sequences and grammar (e.g. Ellis 1996, 2012a; Martin & Ellis 2012). The central executive, in contrast, has been demonstrated to be more important in selective and resource-demanding language processes as well as the real-time performance areas of L2 comprehension, L2 interaction and L2 speech production (e.g. Miyake & Friedman 1998; Wen, Mota & McNeill 2013, 2015; Skehan 2015b).

With regard to the respective roles of working memory and FL aptitude in SLA, the results from two recent meta-analytic studies are particularly revealing. The first was conducted by Li (2015) and investigated the correlation results between aptitude (mostly MLAT-based; excluding working memory), and L2 outcomes with a total of 33 empirical studies (17 predictive and 16 interactional) that involved 3,106 L2 learners. In addition to corroborating the overall positive link between aptitude and L2 learning outcomes ($r = .31$, 95%, CI = .25-.36), a particularly noticeable finding is the generally larger effect sizes for aptitude among younger L2 learners (e.g. high school students) as opposed to older counterparts (e.g. university students). In contrast, the other meta-analytical study was conducted by Linck et al. (2014; also reported in Bunting & Engle 2015), which looked at the relationship between

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

working memory and outcomes of L2 proficiency and processing from 79 samples (involving 748 effect sizes, totaling 3,707 participants). Again, the results were positive, as indicated by an estimated population effect size (ρ) of 0.255. A more relevant finding here was the greater effect sizes for the executive control component of working memory (as measured by the running memory span task (Bunting, Cowan & Saults 2006) as opposed to the storage component (e.g. simple span measures of phonological short-term memory).

Given these emerging patterns of theoretical links between working memory and aspects of L2 learning (Juffs & Harrington 2011), and combined with the more articulate proposal of integrating working memory theories with SLA theories (e.g. the Phonological/Executive model (Wen 2012b, 2015, 2016)), it is fair to claim that phonological working memory is more relevant for younger L2 learners (at primary or secondary schools) and/or early stages of SLA (e.g. French 2006; French & O'Brien 2008) while the executive aspects of working memory should be more substantially implicated in college-level students or post-critical L2 learners, i.e. at LATER SLA stages (also see Doughty et al. 2010). Indeed, the majority of empirical studies in the two recent edited volumes (Wen et al. 2013, 2015) have lent further support to this latter hypothesis. Taken as a whole, then, it is clear that the body of evidence suggesting an important role for working memory within FL aptitude is now considerable.

4.1.2 Implications for aptitude theory construction and test development

More importantly, we believe that incorporating working memory into aptitude is compatible with the developmental process in constructing an aptitude theory, as outlined by Snow. In a 1992 review paper he proposed his conception of an aptitude theory that contains a range of constituent components. As has been argued elsewhere (Wen 2012b, 2016), these various constituents can be re-grouped into three developmental phases in the process of constructing an aptitude theory for tomorrow, namely, the 'testing phase', the 'theory-construction phase' and the 'pedagogical execution phase'.

The first 'testing' phase has been represented by much previous aptitude work where the emphasis was on the development of tests and their use in prediction. This was typical of the 'macro' tradition of aptitude research (see Skehan 2016). The third, execution (and implementation) phase, with its implications for ATI studies, will be considered later. It is interesting here to examine the second, theory-construction phase, where the focus moves from prediction to explanation, and to locate working memory studies within this framework. In earlier sections relevant proposals have already been touched on. For example, Skehan (2002, 2012) has proposed linking SLA stages to components of aptitude. In a similar vein, Robinson (2007) has suggested a number of aptitude complexes in which different combinations of aptitude components assume particular significance in different learning contexts. Linck et al. (2013) report on an aptitude battery (which relates to Snow's 'testing phase') but one which is solidly founded in contemporary cognitive psychology, and hence relates to Snow's theory phase.

What all three proposals have in common is a central place for working memory. Skehan (2016) argues for its relevance, both phonological short-term memory and central executive in all the knowledge-construction stages he describes. It could also be considered relevant

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

to the 'more control' and 'performance' stages (Skehan 2015b) in that, for production, it is likely that speech will be assembled within working memory (Levelt's Conceptualization and Formulation stages (Levelt 1999)) before execution (Levelt's Articulation stage). Robinson (2007), in his hierarchical model of aptitude complexes and abilities, proposes a level of primary cognitive abilities, of which half are concerned with working memory (in relation to phonological short-term memory or for text), and these then feed in, with various combinations, to the four aptitude complexes he outlines, e.g. aptitude for focus on form via recasts (Long & Robinson 1998); aptitude for incidental learning. Linck et al. (2013), within the Hi-LAB model, sample working memory extensively, with distinct contributions from executive and storage systems and an underlying role in their attempt to distinguish 'very successful' from 'successful' learners. So what unites these three accounts of aptitude, each with a different starting point, is that they see working memory as fundamental to the capacity to learn L2s.

Snow's third stage is execution, or intervention, and this too has connected with recent work exploring whether working memory is malleable and can become the focus for educational intervention, i.e. 'the educational execution phase', which involves such criteria as 'selection decision rules', 'classification decision rules' and 'education decision rules'. Working memory profiles based on scores in cognitive assessment scales (Dehn 2008, 2015; Doughty et al. 2010; Tare et al. 2013) should provide ample guidance for 'selection decision rules' and 'classification decision rules' (cf. Gathercole & Alloway 2008; Hummel & French 2010). In terms of working memory and education, two lines of research efforts can be identified (see Williams 2012). First, classroom instruction can be modified in such a way as to suit L2 learners' cognitive working memory profile (Skehan 2012). Second, intervention measures can be taken to modify L2 learners' working memory capacity (though here, a caution might be equally necessary since the two separate components of working memory may require different intervention methods). Tsai, Au & Jaeggi (2016) review a series of studies which have reported working memory training in relation to L2 learning, and they do indeed focus on success in improving central executive components of working memory. They also report that it is important to consider how close the processing operations are in the trained and target use of working memory to facilitate transfer of the effect (cf. Klingberg 2010). (In passing it might be noted that research on the mutability of working memory has not been matched by work on the malleability of other aptitude components. This remains a challenge for the future.)

Gathercole's research group based in the UK has done an impressive amount of work in assessing, identifying, treating and training working memory capacity among young children whose general academic performance is severely affected by their poor working memory (Gathercole & Alloway 2008). From an educational/pedagogic perspective, working memory can be trained and improved to help students' general learning (Dehn 2008, 2015; Gathercole & Alloway 2008; Klingberg 2010). That is to say, future research can aim to evaluate whether these results can be generalized to L2 learning and to what extent they can inform L2 instruction (Doughty et al. 2010).

Therefore, when viewed within this comprehensive standard of constructing an aptitude theory (Snow 1992), the working memory perspective on FL aptitude not only complies with the criteria for 'the testing phase' (thus achieving 'predictive validity') but is most compatible

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

with the latest efforts to redefine the construct (i.e. with a major goal to 'explain' the process), as well as possessing great potential for informing educational and pedagogical practice, as exemplified in the recent Hi-LAB research (Doughty, Jackson & Hughes 2015; Jackson & Doughty 2015). The incorporation of working memory into FL aptitude therefore may push the research field into new and entirely desirable directions (i.e. from 'predicting' and 'explaining' to 'pedagogical execution'). It is fair to argue then that the concept of FL aptitude has developed from being seen as a stable and unitary fixed trait (Carroll's time) to being considered as a more dynamic and multiple set of abilities which interact with other internal or external factors (Robinson 2007; Aguado 2012). For these reasons, we argue that the 'working memory-based perspective on FL aptitude', to use the words from Snow's (1992) title, represents an aptitude theory 'for tomorrow'.

4.2 Micro approaches to language aptitude: Towards systematicity

Another major change that has taken place with FL aptitude research in the last 20 years or so is a shift from 'macro' to 'micro' research designs (Skehan 2016). Many studies have been published (see Li 2015 for a review) which are (a) of relatively or even very short duration, (b) focused on a particular focal grammar point (or two), (c) experimental or quasi-experimental in nature, generally manipulating certain variables relating to type of instruction or type of feedback, and (d) incorporate aptitude measures to explore speed or success in learning in relation to these variables. In addition, a contrast which underlies virtually all the studies reviewed in Li (2015) is between explicit and implicit learning. The former involved instructional treatments with explicit presentation of focused materials and/or feedback, together with some degree of awareness. The latter were based on materials which contained implicit structure but without attention being drawn to this structure in presentation or feedback. (The same distinction will recur later in relation to the focus of aptitude sub-tests themselves).

The studies in Li's (2015) review have broadly indicated that aptitude has an effect even in such short-duration, focused studies, and that aptitude tends to generate larger effect sizes with explicit rather than implicit treatments (see Li 2015 for a meta-analytical review). In a narrative review of such studies, Skehan (2015a) argues that one can reflect upon the range of focal grammar points and offer tentative generalizations. With explicit instruction or feedback, aptitude is more likely to enter into stronger relationships with learning when focal grammar points are redundant for communication. In contrast, with implicit feedback, aptitude links with complexity of L1-L2 relationship. In addition, he notes that a relationship with aptitude may have a different time course with explicit and implicit conditions, growing in impact with implicit conditions after showing little relationship initially, whereas effects with explicit manipulations emerge sooner, and do not grow (cf. Serafini & Sanz 2015).

A first point to make about this research is that aptitude has shown effects over quite short time intervals, and in a way more clearly related to acquisitional processes. Taken as a whole, these results are remarkable in that they show that aptitude does not exert its effects only over the quite long time intervals which are typical of more conventional macro research. They do suggest that aptitude can be an important explanatory variable at

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

a process level. But an important additional point to make is that, given that this research is quite recent, the reviews now available (Li 2015; Skehan 2015a) indicate that different researchers, unsurprisingly, have chosen to relate aptitude to different focal grammar points, as well as different interpretations of feedback and instructional treatment. The suggestions in Skehan (2015a) are post-hoc interpretations on the basis of the findings which have been reported.

In light of these recent developments, we believe what is needed now is a greater degree of systematic coverage of potentially important variables, such as salience and redundancy in this research paradigm, to give clearer insights as to whether aptitude relates to some focal grammar points better than others. This might give us insight into exactly how aptitude makes the contributions that it does. For example, we need to tease out how different lengths of treatment time have different relationships with different components of aptitude, particularly as this relates to the implicit-explicit contrast, with the possibility that some aptitude components influence learning relatively quickly while others require longer periods (cf. specifically the study by De Graaff (1997), which suggested aptitudinal factors take more time to manifest their influence when instruction is implicit). This line of research is exciting as it suggests that a new paradigm for aptitude research has been unlocked. That it has delivered such interesting results already is encouraging and very suggestive that more systematic research of this sort will be very revealing as to the more precise contribution that aptitude can make.

4.3 Knowledge vs control, and implications for aptitude tests

With the development of CANAL-F and Hi-LAB, as well as the availability of LLAMA (Meara 2005) as an MLAT substitute, we now have a range of aptitude tests to supplement the MLAT. Earlier, though, we saw Skehan's (2002) acquisitional sequence as a framework for explicating aptitude sub-tests. Skehan (2016) argues that despite these relatively new tests, there is something of an imbalance between aptitude subtests, such as grammatical sensitivity, which emphasize knowledge acquisition (which link with processes such as noticing, pattern identification, restructuring, processing of feedback and the like) – tests which exist in reasonable numbers – and aptitude sub-tests, for example, of implicit learning, which focus on control over the system which has been developed. These are not so plentiful. One can wonder, therefore, what such tests would be like. First of all, this would mean creating tests which are associated with being quicker to achieve error-free performance, even if this were slow and effortful. This would represent relatively limited control, in other words, with accuracy somewhat at the expense of fluency and speed (Skehan 2015b). Beyond that, control is likely to involve speed of automatizing the knowledge which has been partially achieved so that not only accuracy but also fluency and speed of delivery are involved. Then, even beyond such a level of control, there are issues of developing aptitude tests which might predict native-likeness in terms of the lexical choices made. Finally, there is the capacity to use lexicalized language to ease processing pressures even further. These are difficult challenges, and it is difficult to see what such tests would look like. But new tests are needed if we are to be able to account for, and predict, advanced level language achievement (see also Doughty et al.

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

2010 for a similar argument). Currently it seems there only exist some Hi-LAB sub-tests, and possibly LLAMA D.

One might examine existing tests to see what they have to offer. CANAL-F, to some degree, but Hi-LAB, in particular, has made important progress in developing some relevant aptitude tests. For example, Hi-LAB (Doughty et al. 2010; Linck et al. 2013) contains sub-tests on long-term memory retrieval (through priming), implicit learning and processing speed, which are relevant. Many of the specific CANAL-F sub-tests have an implicit component and this may apply especially strongly to the delayed retrieval variant of several of the sub-tests. Finally, LLAMA D (Meara 2005) concerns the ability to recognize sounds in a spoken language. Granena (2013, 2015) argues that LLAMA D is a measure of implicit learning since it contains no explicit presentation and was also distinct from other (more explicit) aptitude measures in a factor analytic study while nonetheless loading similarly to other implicit measures. But there is still a lot to do if FL aptitude research is to go beyond how language is acquired, and to measure skills that are relevant to how it is used (Williams 2015). So, in coming years, it would be good to see more tests developed to capture these later processes of L2 development. They would be relevant to all levels of L2 proficiency, since beginner and also intermediate learners have to go beyond slow and effortful performance, even with a limited language repertoire. They too have to make accessible (Skehan 1995) and usable what they have acquired. But such tests would certainly be essential in predicting high levels of accomplishment in an L2, where the nature of improvement is less the learning of new things but rather the capacity to sustain real-time processing, and to use language more idiomatically.

4.4 Domain-specificity/generality and explicit vs implicit processes

At a deeper level, two contrasts have importance and their resolution would have implications not solely for aptitude testing itself but for SLA and applied linguistics more generally. The first concerns whether there is something special about language in language aptitude testing, on the one hand, or whether, in post-critical period learning, we are dealing with a general cognitive activity; in which case, while language may be a face-validity-oriented veneer in aptitude tests, the real processes (of pattern recognition, or automatization and so on) are general cognitive processes, and language has no special role at all. In other words, L2 language learning might be no different, essentially, from learning in other more general domains. The second fundamental issue is whether such learning is primarily explicit or implicit. In this case, the contrast is between learning, which was originally the focus of attention, and possibly rule-based, compared with learning below the level of consciousness, where judgements about correctness could be made, but without the capacity to explicate any rules that might underlie performance.

Regarding the domain-specificity argument, the practical issue concerns whether language aptitude tests should be based on language material, or whether non-linguistic material would be equally appropriate. There are two broad justifications for the former approach. First, it could be that there is the residue of a critical period, or even that a critical period does not come to an end completely: individual differences in the extent to which this occurs

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

was a speculation of Carroll (1973). As a result, the processing of language material is accomplished at least partly through specific language-equipped capacities. Second, but not necessarily completely distinct, a hierarchical theory of intelligence – sub-abilities such as verbal, mathematical, musical, mechanical were another area to which Carroll (1993) made enormous contributions in psychology – might suggest that there are specialized abilities in the verbal domain which contribute to language aptitude. Whatever the nature and status of any critical period, if these accounts are correct, then FL aptitude tests should be based on verbal material (as is the case with most existing aptitude tests).

In contrast, one might consider that a critical period has completely ended, with no residue in terms of the parameters set. Alternatively, but with similar effect, one might consider a critical period never existed, given the view that learning is unified (MacWhinney 2005) and that all learning (L1 or L2, and learning in non-language domains) follows similar processes, and is frequency or usage based (Ellis 2012b). Aptitude tests, according to this view, would not inherently require language material – general learning tasks could be used to detect abilities to learn new material and automatize it. Hi-LAB (Doughty et al. 2010; Linck et al. 2013) is closest to this view. Some sub-tests do contain language material but this is generally incidental, and the qualities of the language material are not central to the way the sub-test is designed. Other sub-tests within Hi-LAB do not contain language material at all, consistent with this general cognition approach. In other words, choices about aptitude test content have connections with major issues regarding SLA and the nature of language itself. It follows that the nature of the aptitude tests which work best will have implications for how we think about language – they will be a reflection of how learning takes place in the real world (cf. Williams 2015). Certainly it is the case that aptitude tests which are language-based thereby would become more difficult to use in different L1 contexts since they would need to be translated and re-validated.

Added insight on this analysis comes from recent critical period work and the connection between L1 and L2 learning (Skehan 1986a, 1986b). Of course, proposals that there is a critical period for language learning are not without controversy. It is claimed that the evidence in support of a critical period is questionable (Bongaerts 2005; Birdsong 2007; Moyer 2007) and that alternative theoretical accounts are possible. If we do assume, though, for this discussion, that a critical period may exist, several studies have suggested (DeKeyser 2000; Abrahamsson & Hyltenstam 2008; Biedroń & Pawlak 2016) that high aptitude has a compensating function and is associated with the capacity to reach close to native-like L2 levels after a critical period has closed (also see Doughty et al. 2010). On this basis, DeKeyser (2000) proposes that such compensation derives from a high level of talent for explicit language learning. There is even an aptitude connection with resistance to L1 attrition in post-critical period speakers where exposure to the L1 ends or is minimal (Bylund, Abrahamsson & Hyltenstam 2010). However, there are studies that are not so supportive of this position. Granena & Long (2013) report a less strong compensatory role for aptitude. One factor in accounting for this discrepancy is that the methods used to establish near-native-likeness may be an issue, since grammaticality judgement tests are widely used for this, and it is possible that these vary in the extent to which they are based on explicit or implicit processes, with this contrast accounting for the discrepancy in the results. The competing claims will not be discussed here (but see Skehan 2015a for a discussion). The point is to bring out the centrality

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

of aptitude research for a major issue in linguistics: whether aptitude can compensate for the end of a critical period, and whether we are looking at explicit and/or implicit learning in each case.

We turn next to the explicit-implicit contrast, principally in L2 instructed settings. There are, broadly, three possible positions (Ellis 2015): (a) the 'Strong Interface' position is that explicit learning is important and gradually becomes proceduralized so that it is produced automatically (although this is simply explicit knowledge produced very fast); (b) the 'Weak Interface' position is that explicit learning facilitates implicit learning, for example, through noticing, so that it can be used autonomously and below the level of consciousness from an implicit knowledge-base; and (c) the 'No Interface' position is that direct, albeit slow, implicit learning is the most important type of learning.[1] These are theoretical positions but they have implications for aptitude tests. The first two positions are probably the most consistent with existing aptitude test batteries, since batteries like the MLAT do present material and allow explicit processes to operate upon this material. The third position is not so well represented at present. Granena (2013, 2015) has provided evidence to suggest that one sub-test from LLAMA (Meara 2005), the sound recognition sub-test, is based on implicit learning. Hi-LAB (Doughty et al. 2010; Linck et al. 2013) also contains sub-tests of implicit learning (as well as others of a more explicit nature). If a goal for aptitude tests is to achieve a balance of coverage of implicit and explicit areas, there is certainly scope to develop more implicitly oriented aptitude tests (Williams 2015).

But the analysis demonstrates that FL aptitude does not exist in a vacuum. The nature of aptitude sub-tests, and more especially the nature of the sub-tests that work, provides a commentary on fundamental issues in language learning. This is part of the move, already mentioned, of how aptitude research can stimulate experimental and hypothesis-testing research studies, rather than simply pre-post correlational research designs. In this case, the comparison is between the performance of aptitude sub-tests based on explicit learning and implicit learning respectively. The tests that work better could then provide a commentary on the various interface positions, suggesting whether, for example, explicitly oriented tests outperform (or vice versa) the implicitly oriented tests. This would not simply be a practical method for educational prediction and, even, intervention; it help clarify what language learning actually is.

Two further points are worth discussing. First, language learning contexts vary. It is vital therefore that the range of aptitude sub-tests that have been developed are used in a variety of contexts in order to explore possibilities of ATIs, as proposed by Robinson (2007). This general strategy may be particularly revealing regarding the comparative performance of explicit and implicit approaches to aptitude, since each may have areas of greater relevance, e.g. explicit for instructional contexts, implicit for more naturalistic domains. Second, there is scope here to link aptitude work to the sort of neurolinguistic research covered earlier, since there have been proposals that suggest different brain connectivity networks for implicit and explicit learning (Yang et al. 2015). Aptitude could be linked to this research (Schumann 2004; Carpenter 2008; Xiang et al. 2012). Such research may provide clues as to possible brain

[1] One variant of the 'No Interface' position is that there is, separately, scope for explicit learning to occur, but that this learning is less important and never becomes implicit.

https://doi.org/10.1017/S0261444816000276 Published online by Cambridge University Press

localization and neural substrates linked to performance on explicit or implicitly oriented aptitude tests (Li & Grant 2015).

The impact of this analysis is that the type of questions and research designs that aptitude measures over the coming years could be very different to those of the past. Aptitude could become an important variable within more complex research designs which probe the fundamental issues of domain-specificity and explicit-implicit tensions. Such research designs will change how FL aptitude is perceived more widely within applied linguistics. For this reason, we argue that language aptitude research directed towards these two contrasts may bring back the concept of FL aptitude into the mainstream of SLA and applied linguistics.

5. Conclusion

To conclude, through this historical and integrative review of FL aptitude theory and testing research in the past, present and future, it has become evident that the concept of language aptitude is still a relevant, vigorous and viable construct for SLA and applied linguistics. As a result of these renewed research efforts, recent years have witnessed significant progress in both language aptitude test construction (e.g. from MLAT to CANAL-F, LLAMA and Hi-LAB) and theoretical conceptualization as well as research paradigms (from MLAT-based 'predictive' correlation studies to current interactional studies that aim to both predict and 'explain' SLA).

Despite this encouraging progress, due to the marginalized interest in the concept of FL aptitude, advances in both theoretical conception and methodological innovation have been limited (particularly dwarfed when compared to the enormous developments with its individual difference counterpart of L2 motivation). So it is also argued that the proposal of incorporating working memory as a central component of FL aptitude, and the theoretical linkages between aptitude components and learning or acquisition processes, can greatly compensate for some limitations in previous interpretations of the construct. That being so, it is our final hope that the proposals outlined in this paper will serve as a springboard for reopening the research agenda and reorienting future research on FL aptitude so that concerted efforts from multiple disciplines can be directed towards its greater explanatory power in terms of SLA and more innovative research methodology that ultimately can predict, explain and inform L2 learning, training and pedagogy. This seems to be a major line of development with current Hi-LAB research where aptitude information is used to diagnose learner strengths and weaknesses (Jackson & Doughty 2015).

Acknowledgements

The authors wish to thank the editor, Graeme Porte, and the anonymous reviewers for their detailed reading and constructive suggestions made to earlier drafts. In particular, we thank Catherine Doughty and Jing Yang for providing us with additional materials that have helped clarify and improve the paper. All remaining limitations and errors are our own.

References

Abrahamsson, N. & K. Hyltenstam (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30.4, 481–509.

Aguado, K. (2012). Language learning aptitude and foreign language learning. *Individual characteristics in foreign language education: Perspectives from research and practice. Proceedings of the CLaSIC Conference*. Berlin etc.: De Gruyter, 51–69.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders* 36, 189–208.

Baddeley, A. D. (2012). Working memory: Theories, models and controversies. *Annual Review of Psychology* 63, 1–30.

Baddeley, A. D. (2015). Working memory in second language learning. In Z. Wen, M. B. Mota & A. McNeill (eds.), *Working memory in second language acquisition and processing*. Bristol: Multilingual Matters, 17–28.

Baddeley, A. D., S. Gathercole & C. Papagno (1998). The phonological loop as a language acquisition device. *Psychological Review* 105, 158–173.

Biedroń, A. (2012). Memory in gifted foreign language learners. In M. Pawlak (ed.), *New perspectives on individual differences in language learning and teaching*. Heidelberg-Berlin: Springer, 77–95.

Biedroń, A. (2015). Neurology of foreign language aptitude. *Studies in Second Language Learning and Teaching* 5.1, 13–40.

Biedroń, A. & A. Szczepaniak (2012). Working-memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal* 96, 290–306.

Biedroń, A. & M. Pawlak (2016). New conceptualizations of linguistic giftedness. *Language Teaching* 49.2, 151–185.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O. S. Bohn & M. J. Munro (eds.), *Language experience in second language speech learning*. Amsterdam: John Benjamins, 99–116.

Bongaerts, T. (2005). Introduction: Ultimate attainment and the critical period hypothesis for second language acquisition. *IRAL* 43, 259–267.

Bongaerts, T., S. Mennen & F. van der Silk (2000). Authenticity of pronunciation in naturalistic second language acquisition. The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54, 298–308.

Bunting, M. F., N. Cowan & J. S. Saults (2006). How does running memory span work? *Quarterly Journal of Experimental Psychology* 59, 1691–1700.

Bunting, M. F. & R. Engle (2015). Foreword. In Z. Wen, M. B. Mota & A. McNeill (eds.), xvii–xxiv.

Bylund, E., N. Abrahamsson & K. Hyltenstam (2010). The role of language aptitude in first language attrition: The case of prepubescent attriters. *Applied Linguistics* 31, 443–464.

Carpenter, H. (2008). A behavioral and electrophysiological investigation of different aptitudes for L2 grammar in learners equated for proficiency level. Ph.D. dissertation, Georgetown University.

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

Carroll, J. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of Psycholinguistics* 2, 5–14.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House, 83–118.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. Parry & C. W. Stansfield (eds.), *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice Hall, 11–29.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Carroll, J. B. & S. M. Sapon (1959/2002). *Modern Language Aptitude Test (MLAT)*. New York, NY: The Psychological Corporation. (Reprinted in 2002.)

Chan, E., P. Skehan & G. Gong (2011). Working memory, phonemic coding ability and foreign language aptitude: Potential for construction of specific language aptitude tests – the case of Cantonese. *Ilha Do Desterro: A Journal of English Language, Literatures and Cultural Studies* 60, 45–73.

Conway, A., M. Kane, M. Bunting, D. Hambrick, O. Wilhelm & R. Engle (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12, 769–786.

Cowan, N. (2015). Second language use, theories of working memory and the Vennian mind. In Z. Wen, M. B. Mota & A. McNeill (eds.), 29–40.
De Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition* 19.2, 249–276.
Dehn, M. J. (2008). *Working memory and academic learning: Assessment and intervention*. Hoboken, NJ: John Wiley & Sons, Inc.
Dehn, M. J. (2015). *Essentials of working memory assessment and intervention*. Hoboken, NJ: John Wiley & Sons, Inc.
DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22, 499–533.
DeKeyser, R. M. (ed.) (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.
DeKeyser, R. M., I. Alfi-Shabtay & D. Ravid (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31.3, 413–438.
DeKeyser, R. M. & J. Koeth (2011). Cognitive aptitudes for second language learning. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning*. New York, NY: Routledge, 395–407.
Dien, J., A. Weinberg, S. Blok, P. O'Rourke, K. Kayton & N. Hamedani (2011). *Cognitive neuroscience of second language acquisition: The Department of Defense research landscape*. Center for Advanced Study of Language, University of Maryland. Washington, DC: Georgetown University.
Dörnyei, Z. (2001). *Teaching and researching motivation*. Harlow: Pearson Education.
Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (ed.), *Continuum companion to second language acquisition*. London: Continuum, 247–267.
Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching* 47.1, 80–91.
Dörnyei, Z., A. Henry & C. Muir (2016). *Motivational currents in language learning: Frameworks for focused interventions*. New York, NY: Routledge.
Dörnyei, Z., P. MacIntyre & A. Henry (2015). *Motivational dynamics in language learning*. Bristol: Multilingual Matters.
Dörnyei, Z. & P. Skehan (2003). Individual differences in second language learning. In C. Doughty & M. H. Long (eds.), *The handbook of second language acquisition*. Malden, MA: Blackwell Publishing, 589–630.
Doughty, C. J. (2013). Optimizing post-critical-period language learning. In G. Granena & M. H. Long (eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins, 153–175.
Doughty, C. J. (2014). Assessing aptitude. In A. Kunnan (ed.), *The companion to language assessment*. Oxford, UK: Wiley-Blackwell, 25–46.
Doughty, C. J., S. G. Campbell, M. A. Mislevy, M. F. Bunting, A. R. Bowles & J. T. Koeth (2010). Predicting near-native ability: The factor structure and reliability of Hi-LAB. In M. T. Prior, Y. Watanabe & S-K. Lee (eds.), *Selected proceedings of the 2008 Second Language Research Forum*. Somerville, MA: Cascadilla Proceedings Project, 10–31.
Doughty, C. J., S, Jackson & M. Hughes (2015). *Implementation of the high-level language aptitude battery at the Foreign Service Institute and the National Security Agency*. Paper presented at the East Coast Organization of Language Testers Conference. 9 October 2015.
Ehrman, M. E. (1998). The Modern Language Aptitude Test for predicting learning success and advising students. *Applied Language Learning* 9.1–2, 31–70.
Ehrman, M. E. & R. L. Oxford (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal* 79.1, 67–89.
Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition* 18, 91–126.
Ellis, N. C. (2012a). Formulaic language and second language acquisition: Zipf and the phrasal Teddy Bear. *Annual Review of Applied Linguistics* 32, 17–44.
Ellis, N. C. (2012b). Frequency-based accounts of SLA. In S. Gass & A. Mackey (eds.), *The Routledge handbook of second language acquisition*. London: Taylor &Francis/Routledge, 193–210.

Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics*. Oxford: Blackwell, 525–552.
Ellis, R. (2012). *Language teaching research and language pedagogy*. Chichester, West Sussex; Malden, MA: Wiley-Blackwell.
Ellis, R. (2015). *Understanding second language acquisition* (2nd edn.). Oxford: Oxford University Press.
Ellis, R. & N. Shintani (2013) *Exploring language pedagogy through second language acquisition research*. London: Routledge.
Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research* 9.2, 147–171.
French, L. (2006). *Phonological working memory and second language acquisition: A developmental study of francophone children learning English in Quebec*. Lewiston, NY: Edwin Mellen Press.
French, L. M & I. O'Brien (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics* 29, 463–487.
Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics* 27, 513–543.
Gathercole, S. E. & T. P. Alloway (2008). Working memory and classroom learning. In K. Thurman & K. Fiorello (eds.), *Cognitive development in K-3 classroom learning: Research applications*. Mahwah, NJ: Lawrence Erlbaum, 15–38.
Golestani, N. (2014). Brain structural correlates of individual differences at low to high levels of the language processing hierarchy: A review of new approaches to imaging research. *International Journal of Bilingualism* 18.1, 6–34.
Golestani, N. & R. J. Zatorre (2009). Individual differences in the acquisition of second language phonology. *Brain and Language* 109, 55–67.
Golestani, N., C. J. Price & S. K. Scott (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience* 31.11, 4213–4220.
Granena, G. (2013). Reexamining the robustness of aptitude in second language acquisition. In G. Granena & M. H. Long (eds.), 179–204.
Granena, G. (2015). Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied Psycholinguistics*. Online First View. doi:10.1017/S0142716415000120.
Granena, G. & M. Long (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29, 311–343.
Green, P. S. (1975). *The language laboratory in school: The York study*. Edinburgh: Olivier and Boyd.
Grigorenko, E. L., R. J. Sternberg & M. E. Ehrman (2000). A theory based approach to the measurement of foreign language learning ability: The CANAL-F theory and test. *The Modern Language Journal* 84, 390–405.
Gullberg, M. & P. Indefrey (eds.) (2006). *The cognitive neuroscience of second language acquisition*. Oxford: Blackwell Publishing.
Harley, B. & D. Hart (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19.3, 379–400.
Harley, B. & D. Hart (2002). Age, aptitude, and second language learning on a bilingual exchange. In P. Robinson (ed.), *Individual differences and instructed language learning*. Amsterdam: John Benjamins, 302–330.
Hu, X., H. Ackermann, J. A. Martin, M. Erb, S. Winkler & S. M. Reiterer (2013). Language aptitude for pronunciation in advanced second language learners: Behavioural predictors and neural substrates. *Brain and Language* 127.3, 366–376.
Hummel, K. M. & L. French (2010). Phonological memory and implications for the second language classroom. *Canadian Modern Language Review* 66.3, 371–391.
Hwu, F. & S. Sun (2012). The aptitude-treatment interaction effects on the learning of grammar rules. *System* 40.4, 505–521.
Jackson, S. R. & C. J. Doughty (2015). *Predicting language-learning success at FSI: Validity of MLAT and Hi-LAB*. Technical report: Center for Advanced Study of Language. University of Maryland, College Park.

Jackson, S. R., S. Swee, A. Bloomfield & C. J. Doughty (2015). *A simulation model of aptitude-by-treatment interaction in language training: Quantifying the impact of tailored language training*. Technical Report: Center for Advanced Study of Language. University of Maryland, College Park.
Juffs, A. & M. Harrington (2011). Aspects of working memory in L2 learning. *Language Teaching* 44, 137–166.
Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences* 14, 317–324.
Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena & M. H. Long (eds.), 131–152.
Levelt, W. J. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (eds.), *Neurocognition of language*. Oxford, England: Oxford University Press, 83–122.
Li, P. & A. Grant (2015). Second language learning success revealed by brain networks. *Bilingualism: Language and cognition*. Online First View. doi:10.1017/S1366728915000280.
Li, P., J. Legault & K. A. Litcofsky (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex* 58, 301–324.
Li, S. (2011). Interactions between feedback type, proficiency, linguistic target, and individual differences in language analytic ability and working memory. Ph.D. dissertation, Michigan State University.
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.
Linck, J. A., M. M. Hughes, S. G. Campbell, N. H. Silbert, M. Tare, S. R. Jackson & C. J. Doughty (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning* 63.3, 530–566.
Linck, J. A., P. Osthus, J. T. Koeth & M. F. Bunting (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review* 21.4, 861–883.
Long, M. & P. Robinson (1998). Focus-on-form: Theory, research and practice. In C. Doughty & J. Williams (eds.), *Focus-on-form in classroom second language acquisition*. Cambridge: Cambridge University Press, 15–41.
MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press, 49–67.
Martin, K. I. & N. C. Ellis (2012). The roles of phonological STM and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition* 34.3, 379–413.
Masgoret, A-M. & R. C. Gardner (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning* 53, 123–163.
Meara, P. (2005). *LLAMA language aptitude tests*. Swansea, UK: Lognostics.
Miller, G. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Science* 7.3, 141–144.
Miyake, A. & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (eds.), *Foreign language learning*. Mahwah, NJ: Lawrence Erlbaum, 339–364.
Moyer, A. (2007). Empirical considerations on the age factor in L2 phonology. *Issues in Applied Linguistics* 15, 109–129.
Parry, T. S. & J. R. Child (1990). Preliminary investigation of the relationship between VORD, MLAT, and language proficiency. In T. S. Parry & C. W. Stansfield (eds.), *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice-Hall Regents, 30–66.
Peterson, C. R. & A. R. Al-Haik (1976). The development of the defense language aptitude battery (DLAB). *Educational and Psychological Measurement* 36, 369–380.
Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery*. New York, NY: Harcourt Brace Jovanovich.
Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (ed.), 159–180.
Reves, T. (1982). What makes a good language learner? Ph.D. dissertation, Hebrew University of Jerusalem, Israel.
Robinson, P. (1995). Aptitude, awareness and the fundamental similarity of implicit and explicit foreign language learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawai'i Press, 303–357.
Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search and instructed condition. *Studies in Second Language Acquisition* 18, 27–67.

Robinson, P. (2002). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (ed.), 113–133.
Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics* 25, 46–73.
Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. M. DeKeyser (ed.), *Practice in second language*. Cambridge: Cambridge University Press, 256–286.
Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak (ed.), *New perspectives on individual differences in language learning and teaching*. Berlin Heidelberg: Springer-Verlag, 57–75.
Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York, NY: Peter Lang.
Sawyer, M. & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 319–354.
Schumann, J. H. (2004). The neurobiology of aptitude. In J. H. Schumann (ed.), *The neurobiology of learning: Perspectives from second language acquisition*. Mahwah, NJ: Lawrence Erlbaum, 7–23.
Serafini, E. J. & C. Sanz (2015). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, 1–40. Online First View. doi:10.1017/S0272263115000327.
Sheen, Y. (2007a). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly* 41.2, 255–283.
Sheen, Y. (2007b). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (ed.), *Conversational interaction in second language acquisition*. New York: Oxford University Press, 301–322.
Singleton, D. (2014). Apt to change: The problematic of language awareness and language aptitude in age-related research. *Studies in Second Language Learning and Teaching* 4.3, 557–571.
Skehan, P. (1982). Memory and motivation in language aptitude testing. Ph.D. dissertation, University of London, UK.
Skehan, P. (1986a). The role of foreign language aptitude in a model of school learning. *Language Testing* 3, 188–221.
Skehan, P. (1986b). Where does language aptitude come from? In P. Meara (ed.), *Spoken language*. London: Center for Information on Language Teaching, 95–113.
Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.
Skehan, P. (1995). Analysability, accessibility, and ability for use. In G. Cook & B. Seidlehofer (eds.), *Principle and practice in applied linguistics*. Oxford: Oxford University Press, 91–106.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (ed.), 69–95.
Skehan, P. (2012). Language aptitude. In S. Gass & A. Mackey (eds.), 381–395.
Skehan, P. (2015a). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics* 36.3, 367–384.
Skehan, P. (2015b). Working memory and second language performance. In Wen, Z., M. Mota & A. McNeill (eds.), 189–201.
Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson & Y. Yilmaz (eds.), *Cognitive individual differences in L2 processing and acquisition*. Amsterdam: John Benjamins, 15–38.
Snow, R. E. (1987). Aptitude complexes. In R. Snow & M. Farr (eds.), *Aptitude, learning and instruction, cognitive process analyses of aptitude* (2 vols.). Hillsdale, NJ: Erlbaum Associates, 11–34.
Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist* 27, 5–32.
Snow, R. E. (1994). Abilities in academic tasks. In R. J. Sternberg & R. K. Wagner (eds.), *Mind in context: Interactionist perspectives on human intelligence*. New York, NY: Cambridge University Press, 3–37.
Sparks, R. L. & L. Ganschow (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *The Modern Language Journal* 7, 3–16.
Sparks, R. & L. Ganschow (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics* 21, 90–111.
Sparks, R. L., L. Ganschow & J. Patton (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology* 87, 638–655.

Sparks, R., L. Ganschow & J. Patton (2008). L1 and L2 literacy, aptitude, and affective variables as discriminators among high- and low-achieving L2 learners. In J. Kormos & E. Kontra (eds.), *Language learners with special needs: An international perspective*. London: Multilingual Matters, 11–35.
Sparks, R., J. Patton, L. Ganschow & N. Humbach (2011). Subcomponents of L2 aptitude and L2 proficiency. *The Modern Language Journal* 95, 253–273.
Sternberg, R. J. (1997). *Successful intelligence*. New York, NY: Plume.
Sternberg, R. J. (2002). The theory of Successful Intelligence and its implications for language aptitude testing. In P. Robinson (ed.), 13–43.
Sternberg, R. J. & E. L. Grigorenko (2002). The theory of successful intelligence as a basis for gifted education. *Gifted Child Quarterly* 46, 265–277.
Tan, L., L. Chen, V. Yip, A. H. Chan, J. Yang, J. H. Gao & W. T. Siok (2011). *Activity levels in the left hemisphere caudate-fusiform circuit predict how well a second language will be learned*. Proceedings of the National Academy of Sciences, USA 108, 2540–2544.
Tare, M., C. Bonilla, K. Vatz, M. Clark, J. Linck & C. J. Doughty (2013). *Building a language learner's profile: Characteristics which inform training and pedagogy*. CASL technical report. College Park, MD: Center for Advanced Study of Language.
Trofimovich, P., A. Ammar & E. Gatbonton (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (ed.), *Conversational interaction in second language acquisition*. New York: Oxford University Press, 171–195.
Tsai, N., J. Au & S. M. Jaeggi (2016). Malleability of working memory and implications for second language acquisition. In Granena et al. (eds.), 67–87.
VanPatten, B. & S. Borst (2012a). The roles of explicit information and grammatical sensitivity in processing instruction: Nominative-accusative case marking and word order in German L2. *Foreign Language Annals* 45.1, 92–109.
VanPatten, B. & S. Borst (2012b). The roles of explicit information and grammatical sensitivity in the processing of clitic direct object pronouns and word order in Spanish L2. *Hispania* 95.2, 270–284.
Vatz, K., M. Tare, S. R. Jackson & C. Doughty (2013). Aptitude-treatment interaction in second language acquisition: Findings and methodology. In G. Granena & M. H. Long (eds.), 273–292.
Wen, Z. (2012a). Foreign language aptitude. *ELT Journal* 66.2, 233–235.
Wen, Z. (2012b). Working memory and second language learning. *International Journal of Applied Linguistics* 22, 1–22.
Wen, Z. (2014). Theorizing and measuring working memory in first and second language research. *Language Teaching* 47.2, 173–190.
Wen, Z. (2015). Working memory in second language acquisition and processing: The Phonological/Executive model. In Wen, Z., M. B. Mota & A. McNeill (eds.), 41–62.
Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Bristol: Multilingual Matters.
Wen, Z. & P. Skehan (2011). A new perspective on foreign language aptitude: Building and supporting a case for 'working memory as language aptitude'. *Ilha Do Desterro: A Journal of English Language,Literatures and Cultural Studies* 60, 15–44.
Wen, Z., M. B. Mota & A. McNeill (2013). Working memory and SLA: Innovations in theory and research. *Asian Journal of English Language Teaching* 23, 1–102.
Wen, Z., M. B. Mota & A. McNeill (eds.) (2015). *Working memory in second language acquisition and processing*. Bristol, UK: Multilingual Matters.
Wesche, M. B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (ed.), 119–154.
Williams, J. N. (2012). Working memory and SLA. In S. Gass & A. Mackey (eds.), 427–441.
Williams, J. N. (2015). Working memory in SLA research: Challenges and prospects. In Z. Wen, M. B. Mota & A. McNeill (eds.), 301–307.
Winke, P. (2013). An investigation into L2 aptitude for advanced Chinese language learning. *The Modern Language Journal* 97.1, 109–130.
Xiang, H., D. Dediu, L. Roberts, E. V. Oort, D. G. Norris & P. Hagoort (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Language Learning* 62, 110–130.
Yang, J. & P. Li (2012). Brain networks of explicit and implicit learning. *PLoS One* 7.8, e42993.

Yang, J., K. M. Gates, P. Molenaar & P. Li (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics* 33. 29–49.
Yilmaz, Y. (2013). The role of working memory capacity and language analytic ability in the effectiveness of explicit correction and recasts. *Applied Linguistics* 34.3, 344–368.

ZHISHENG (EDWARD) WEN is Associate Professor at the School of Languages and Translation, Macao Polytechnic Institute, Macau SAR, China. Dr Wen has lectured, researched and published extensively in applied linguistics and psycholinguistics and his current research foci are theoretical and methodological issues surrounding language aptitude and working memory in SLA. His recent books include a research monograph, *Working memory and second language learning*, 2016 and an edited volume, *Working memory in second language acquisition and processing*, 2015, with Mailce Mota and Arthur McNeill, both published by Multilingual Matters.

ADRIANA BIEDROŃ is Professor in English Philology Department at the Pomeranian Academy in Słupsk. Professor Biedroń's research interests include applied psycholinguistics and second language acquisition theory. Her research focuses on individual differences in SLA, in particular, foreign language aptitude and cognitive and personality factors in gifted L2 learners. Her recent publications include a research monograph, *Cognitive-affective profile of gifted adult foreign language learners* 2012, Słupsk) and a state-of-the-art article 'New conceptualizations of linguistic giftedness' (with M. Pawlak) in *Language Teaching* (2016).

PETER SKEHAN is a Professorial Research Fellow at St Mary's University College, Twickenham, London having previously worked at Auckland University, The Chinese University of Hong Kong, King's College London and Thames Valley University. Professor Skehan's research interests and extensive publications are in major areas of applied linguistics such as SLA, L2 task-based language teaching & learning as well as individual differences in L2 learning (especially language aptitude). His recent book was an edited volume, *Investigating a processing perspective on task performance*, published by John Benjamins (2014).