

Chapter 11

Summing up: Individual differences in primary school foreign language learning

Raphael Berthele^a & Isabelle Udry^{a,b}

^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

In Chapter 1, we set the stage for this volume with an outline of the constructs, research methods, and pedagogical relevance of individual differences (IDs) in foreign language teaching and learning. At the close, we wish to revisit the main findings of the chapters of this volume in order to summarize what we have learnt from the investigation and to think about what our insights can mean both for researchers and practitioners.

Our ambition was to provide empirical evidence on learner characteristics that explain (statistically and theoretically) variance in second language skills. The long-term goal of such endeavours is to come to a better understanding of learner variability which in the future could inform pedagogical choices and practices in the foreign language classroom.

In the following, we summarize the main findings drawn from the different chapters and discuss their educational and theoretical implications. The road leading from research as presented in this volume to policy recommendations is long, winding, and often rocky. Language curricula and multilingual education, eminently so in officially multilingual countries, are never entirely “evidence-based”, i.e. they are not simple transpositions of research findings into practice. They are the result of complicated political and other institutional processes, with fluctuating and often inconsistent recourse to scholarly research (see Berthele



2019 for examples). Our goal, however, is to contribute to the growing pool of evidence on the possibilities and the limits of foreign language learning in compulsory primary school curricula.

1 Summary of the findings

In our research, we made an attempt to contribute to three different perspectives on individual differences (IDs) in learning foreign languages: The first is the interest in a better understanding of the variables or constructs that account for the differences in the ability to learn a foreign language: What are the cognitive, affective, and sociological variables that are related to foreign language learning ability, and what is the internal dimensionality of such a broad array of individual difference measures? The second perspective relates to the feasibility, based on the results produced within the first perspective, of drawing on these ID variables in order to prognosticate language development in the foreign language. The third perspective is the interest in change or stability of ID variables over time.

To contribute new evidence to all three perspectives, we examined a range of ID variables deemed to be associated with foreign language learning at primary school beyond the components theorized by Carroll in the 1950s and 60s.

1.1 Dimensions

Of particular interest to the field is the relationship between cognitive ID variables that are more (phonetic coding ability, grammatical sensitivity, inductive ability) or less language related (e.g. fluid intelligence, working memory, field independence). Exploratory and confirmatory factor analyses on two independent samples (LAPS I and II) showed that general and language related cognitive variables load on the same factor. We chose to call this factor *Cognition/Aptitude*. Since the structure of constructs that emerged from the exploratory factor analyses was confirmed in the second, larger sample, we confidently concluded that for the age group of 10- to 12-year-old children examined here, general cognitive abilities and language-oriented abilities represent a single dimension (cf. Chapter 3). In line with other evidence on L2 motivation, we identified a distinction among variables that can broadly be delimited into intrinsic vs. extrinsic facets of motivation. Our analyses yielded two affective factors we named *L2 Academic Emotion* and *Extrinsic* factor. A regression analysis showed that the *L2 Academic Emotion* factor, together with *Cognition/Aptitude*, relates positively

to L2 proficiency. The *Extrinsic* factor, when considered together with the other two factors, was associated negatively with L2 proficiency.

Further explorations of these data are presented in Chapters 5 and 9. In Chapter 9, the analyses compare the linear relationships of individual difference variables with the language of instruction (German) and the foreign language (English), respectively. The analyses yield similar patterns of associations of these individual characteristics with both languages, a result in line with the assumption of an underlying ability to learn and use languages that is not dedicated to a specific language. Moreover, Chapter 9 also explores the contribution of socioeconomic variables and home language use (i.e. individual multilingualism beyond the languages taught at school). These associations are further explored in Chapter 5, by distinguishing between cultural and economic dispositions of the children. The analyses presented in Chapter 5 suggest that such family background characteristics are indeed related to foreign language skills in the target language English. Their contribution, however, is indirect via the two factors that stood out in Chapter 3, namely *Cognition/Aptitude* and *L2 Academic Emotion*. The estimate of the direct path to English as a foreign language is small. The chapter also investigates whether specific features of pupils particularly associated with socio-educational vulnerability, such as being born abroad or not speaking the local language German at home, impact English skills beyond the factorial structure already identified in Chapter 3. These analyses show that the variable cultural and economic dispositions are first and foremost associated with the two factors emerging in Chapter 3 but do not affect L2 English skills beyond these indirect effects.

1.2 Prediction

Identifying predictor variables for L2 proficiency is a way to assess student potential (Chapter 4). Assessing this potential had been the goal of the early language aptitude tests (see Chapter 1 for details). Our study focused on a markedly different educational context than the early and classical aptitude tests did: Two foreign languages are an obligatory part of the curriculum, and the learners are not young adults or adolescents, but children in primary school.

We first trained statistical models with all variables available from the test battery on the LAPS II training set. That is, motivational, general and language related cognitive variables, as well as social background variables were all taken into consideration when training these models. This yielded a *No costs spared model* whose power to prognosticate English skills at T3 was assessed on an unseen test set. Next, we compared this model to simpler, more practical models

with a small set of measures that are relatively easy to take in a classroom. We refer to these models as *Cheap models*. Moreover, since our participants already had English skills, we were also interested in how the English test at T1 fared as a sole predictor for English proficiency at T3.

Overall, our results showed that the English test taken at T1 provides a good forecast of L2 proficiency 1,5 years later at T3. When all variables were considered, i.e. the *No costs spared* model, a linear regression model with 7 predictors fared best: Variables included in this *No costs spared* model are English T1, grade (=year in school), intrinsic motivation, self-concept English, L1 German, MLAT (grammatical sensitivity), and PLAB (inductive ability). The practical, or *Cheap*, models contain small selections of variables that encode either information on learners that is usually readily available in a classroom setting or that is rather easy to collect (e.g. motivational information via questionnaires). The best of these cheap models is one that includes the English test scores at T1 and such easily obtained information. This model performed only slightly worse than the 7 predictor *No costs spared* model at T1. Overall, the comparison of these different predictive models shows that the most informationally rich single measurement for prognostication of skills at T3 is a test of the same skill at T1. A high level of stability within constructs is also what emerges from the analysis in Chapter 10. Here, the stability of two language related measures, a grammatical sensitivity task inspired by an MLAT form and an inductive ability task inspired by a PLAB form are investigated longitudinally: When accounting for measurement error, the scores between T1 and T3 are not perfectly correlated. Thus, the abilities measured are not perfectly stable traits within learners across time. However, the high association ($\rho = 0.74$) suggests a high level of stability of these measurements.

When interpreting the prognostic models in Chapter 4, it is important to keep in mind that variables or dimensions that have not emerged as predictors in these models must not be hastily dismissed as being irrelevant for foreign language learning. Rather, our results suggest that considered together, the variables from the *No costs spared* model give, with some precision, an estimate of how a student's L2 skills are likely to develop between T1 and T3, i.e. between grades 4 and 6. The *English only model* and some of the *Cheap models* fared slightly worse in doing so. As can be seen from the results reported in Chapter 4, the mean error of the models differed within a span of 1.9–2.2 points on a 20-point scale when fitted on the test set. We are not aware of other studies that have attempted to assess in this fashion if and how accurately prognostications about foreign language learning can be made. The main insight emerging from our analyses is

that prognostic testing in primary school foreign language learning is indeed possible, and it can be done with high predictive accuracy.

1.3 Motivation and creativity

Our investigation did not only cover abilities (language related or general cognitive), but also affective, attitudinal and motivational dispositions, as called for by scholars for quite some time (Dörnyei 2010: 267, Parry & Stansfield 1990: 2). Beyond the insights into the different factorial dimensions provided in Chapter 3, two chapters are dedicated to a more fine-grained analysis of different motivational components as well as their development over time (Chapters 7 and 8). The specific setting of our learners, hailing from two different areas of the German-speaking part of officially quadrilingual Switzerland, called for a differentiated analysis of the pupils' motivational stances regarding their own foreign language learning. The comparison of two foreign languages (English and French), both compulsory but introduced in the curriculum in reverse order in the two settings, reveals insights both into the effects of spatial proximity of the target language territory and of foreign language subjects in the curriculum. French is the demographically and culturally dominant language in the region of LAPS I and it is taught as the first L2 in this region's German-speaking areas. The proximity of a French speech community was expected to trigger higher motivation in these pupils, at least with respect to extrinsic or *lingua franca* uses of the language. The results reported in Chapter 7 show that there is no difference between the two areas, that is the motivational dispositions look the same as in the region of LAPS II where French is taught as the second foreign language (L3) and is much further away, both geographically and culturally. The motivation to learn English is higher across all motivational sub-dimensions and in both of the contexts investigated.

Moreover, as in the other developmental analyses, the development across time of the constructs seems rather stable. The most consistent change, as discussed in Chapter 8, seems to be the decline in school-related motivation for both target languages across the 1,5 years investigated. The decline is not dramatic and similar tendencies have been observed in numerous studies on different subjects (see Shan 2020, chapter 2.1 for discussion and references).

Motivation and creativity, more specifically divergent thinking, are investigated in Chapter 6. It is conceivable that task-based learning that requires pupils to generate creative outcomes such as poems or role plays, draws particularly on creative thinking. In such an environment, creative thinking is expected to be associated with the motivation to learn the target language(s). However, as the

analyses reveal, no noteworthy association of the two constructs can be found. Modelling the association of creativity with L2 proficiency in both languages (French and English) in the LAPS I sample reveals a weak but significant association. As discussed in Chapter 6, the causality or contribution of one construct on the other remains a matter of speculation and is difficult to pin down.

2 Discussion of the main findings

2.1 Applying aptitude-related findings

Aptitude tests were initially developed in the US to select able students for language classes (see Chapter 1 for a discussion). In new educational paradigms, the role of predictive testing as a tool of selection seems less fitting. Public education is expected to do its best to even out achievement gaps to which individual differences contribute, and to accommodate different learner needs. At the same time, selection is one of the functions of modern educational systems. Overt or covert selection practices are therefore in place in contemporary educational contexts, also in the Swiss case our study was concerned with. Foreign language learning being part of the compulsory subjects in the Swiss curricula, achievement in these subjects must therefore be considered when educational selection processes are at stake. Achievement measures in foreign language subjects sometimes form a part of high-stakes test results feeding into tracking decisions. In this context, we feel that we can nourish current debates that concern such language-related selection processes. Individual differences are at the centre of attention when discussing whether some students should be exempted from one or both compulsory foreign language subjects if they suffer from learning difficulties (cf. the Linguistic Coding Difference Hypothesis mentioned in Chapter 1). If dispensation from, say, English as a foreign language were to be considered a reasonable practice by teachers and other stakeholders, it would indeed be possible to draw on the results presented here – in the sense that metrics that are associated with L2 skills as discussed in Chapters 3 and 4 could be used to define and assess selection criteria. However, as mentioned in Chapter 1, identifying individual differences in predispositions for learning foreign languages can serve other purposes as well. It can be used, for example, to recognize learners who need more time to acquire similar skills.

In Chapter 1, we critically discussed the aptitude-treatment-interaction (ATI) approach which investigates the mutual influence between language aptitude and teaching methods. ATI is underpinned by the assumption that a) learners do have different aptitude profiles and b) that these profiles are in interaction

with teaching methods, i.e. selecting a method tailored to the aptitude profile will enhance learning. Due to various reasons, ATI has so far failed to deliver enough robust evidence to substantiate these claims. While it is uncontroversial that individual differences in learners exist, and also that learners express different preferences when it comes to choosing ways and methods of learning, the claim that differentiated pedagogical treatments allow learners to systematically draw on strengths and compensate for weaknesses is not based on sufficient robust evidence (see our discussion in Chapter 1). If an ATI approach nevertheless should turn out to be empirically sound, the dimensionality of the aptitude construct from Chapter 3 provides insights on the axes along which learners vary.

If one wishes to assess students' potential L2 development, any of the prognostic models discussed in Chapter 4 are of some avail. Their usability depends on context and purpose. If learners already have basic L2 knowledge, an English test is a suitable and convenient choice for teachers. However, the *Cheap models* are informative if one wishes to gain some insight into affective and cognitive dimensions, an area that is likely to be of interest to practitioners over and above the prediction of L2 skills. Intervention-based research could clarify whether such easy to collect information from the Cheap models could be used to adapt the pedagogical setting in order to attain better learning results longitudinally. If a Cheap model indicates, for instance, that a pupil's L2 motivation is low, appropriate measures could be taken to assist that student.

The extent to which education can influence learner performance at all may not always be as large as educators would like it to be. As shown in Chapter 5, socioeconomic status, which cannot easily be changed by the individual, bears strongly on the two factors that are positively but indirectly associated with L2 proficiency (via the constructs *Cognition/Aptitude* and *L2 Academic Emotion*). If the assumption is that social dispositions contribute causally to one or both of these constructs (and not vice versa), then this points to important hurdles for teachers and schools to change individuals' dispositions with respect to these two important constructs. This result raises concerns about how well an education system whose pledge is equal opportunity can live up to such expectations in real life. In a related vein, if the cognitive and/or linguistic abilities are partially predetermined by genetics, as suggested by Plomin (2019) or Stromswold (2001), this also points to limits of the extent to which individual differences can be pedagogically levelled out, in particular within the restricted possibilities of a dense curriculum in a state school with only limited time at disposal for L2 instruction.

2.2 Theory development

Inquiries into language aptitude usually frame the construct drawing on the Carrollian components, i.e. memory (associative or working memory), phonetic coding ability, grammatical sensitivity, and inductive ability – the latter two being also labelled *language analysis* in Skehan's (1998, see Chapter 2.1) three-component model. Based on the description of learner profiles by Skehan, many researchers distinguish between memory and analysis-oriented adult or adolescent learners.

For primary school children as examined in the LAPS project, a memory vs. language analysis distinction cannot be substantiated: As reported in Chapter 3, all general and language related cognitive variables load on one factor. Here, it is important to emphasize that in a study such as ours that encompasses a multitude of cognitive and affective variables, the measures of the constructs are unavoidably coarse: Digit span-based measures of working memory, for example, allow us to grossly distinguish between categories of students (they reliably identify students with serious working memory problems), but the measure, although widely used, is by no means the state-of-the art test of working memory capacity that one would apply if much more time were at one's disposal (e.g. operation span; cf. Conway et al. 2005 for a comparison and overview). It is important to resist the temptation to essentialize factors emerging from data as representing "scientifically proven reality", as more advanced tests might produce evidence pointing towards a more complex internal dimensionality of our component. Given the temporal and logistic constraints that also govern research as ours, the tests and measures we were able to generate represent what was feasible, not what would be done in an ideal world ruled by researchers. On the other hand, given the converging results of the exploratory and confirmatory factor analyses reported in Chapter 3, we feel confident affirming that for primary school learners aged 10–12, there is solid empirical evidence to postulate that memory and analysis related cognitive variables are highly positively correlated.

As discussed in Chapters 3 and 4, affective variables (particularly those connected to L2 Academic Emotion, i.e. intrinsic motivation, self-concept, and anxiety) make a separate contribution to L2 achievement in addition to cognitive variables. Based on these observations, we argue that future studies, be they within an ATI framework or not, could be enriched by going beyond the cognitive language-related focus of the Carrollian concept to include affective variables along the lines of self-determination theory (Deci & Ryan 2002), as a function of different teaching methods. In this view, assessing the effects of specific

treatment conditions on learner's self-concept, anxiety and pleasure in learning foreign languages seems – unsurprisingly – useful in the search for more efficient foreign language teaching, since our data show that these constructs are linearly associated with skills in the target language.

Focusing on the predictive value of the tests used in the study, aptitude measures for language analysis (grammatical sensitivity and inductive ability) have emerged as being predictive of achievement on several accounts: Both appear in the *No costs spared* model. Moreover, an inductive ability task is part of one of the *Cheap* models (Chapter 4) and a grammatical sensitivity task turns out to be a variable that predicts L2 English and L1 German proficiency as evidenced by the analysis in Chapter 9. On the one hand, these findings emphasize the importance of the factor *Cognition/aptitude* for foreign language learning. On the other, they suggest that within this factor, language-related tests are more strongly associated with L2 outcomes than measures of general intelligence or working memory. This is not surprising, since aptitude tests tap into language-oriented constructs, and language skills were the main outcome variables in our study. Based on our findings from the factor analysis, this indicates that the abilities required to solve tests for language analysis and tests for intelligence and WM overlap, with the language analysis tasks tapping into grammatical sensitivity and induction being more strongly associated with language-related outcomes.

In tackling individual predispositions to language learning comprehensively, we wanted to widen the scope by including potentially relevant, but lesser researched ID variables. Creativity has been hypothesized to play a role in foreign language learning, especially in the task-based approach adopted in Switzerland that relies on the learners' potential to generate ideas that help them solve communicative tasks using their target language skills. In multilingualism research, we find rather strong causal claims based on (weak) positive associations of creativity measures and the individuals' language repertoires (e.g. Fürst & Grin 2018). Our data suggest also that there is a positive linear association of non-verbal creativity and proficiency in foreign languages in the LAPS I test battery (see Chapter 6). The association is weak, though, and the direction of causal links cannot be established based on cross-sectional multivariate analyses. An association of creativity and motivation, however, could not be substantiated.

As we have discussed in Chapter 1, there are several unresolved issues in aptitude research, some of which can be elucidated by our findings. Language aptitude, as conceptualized by Carroll, has been deemed obsolete or at least out-dated by some. For example, in the wake of the *Natural approach*, scholars claimed that language aptitude is irrelevant for learning languages in such communicative, “natural” paradigms, in particular in the case of children (Skehan 2002: 72).

Notwithstanding such claims, the results in the literature reported in Chapter 1 as well as our findings show that “old-fashioned” language related constructs are clearly associated with the outcomes of language learning also in communicative language teaching/learning settings (Chapters 3, 9). The evidence also suggests that such aptitude measures contribute to forecasting learning outcomes in communicative teaching (Chapter 4).

Another matter of debate is the extent to which language aptitude is malleable or a relatively stable trait of the individual. This question is relevant if one considers training aptitude subcomponents, such as metalinguistic skills or language analysis, in order to improve learning outcomes. Testing such an assumption would require experimental settings. Our investigation cannot provide answers beyond the aforementioned relative stability of the MLAT words in sentences and PLAB inductive ability tasks over time (Chapter 10).

2.3 Future perspectives

To date, few studies have dealt with instructed foreign language learning at primary school on a large scale. Throughout this book we have explained our methodological and analytical choices, and we have also directed the readers’ attention to possible improvements to our research plan. We hope these discussions encourage other groups to build upon our work, benefit from our insights, and continue the investigation of individual differences in young language learners. Moreover, we would welcome experimental studies that extend on our results. These could clarify some of the theoretical and educational assumptions currently held about individual differences which are not sufficiently underpinned by empirical evidence. For example, we think that the claims made by ATI and its applicability should be examined further with carefully planned designs along the lines outlined throughout the book. Moreover, we have postulated the ability to solve language analysis tests (as in our grammatical sensitivity and inductive ability tasks) being rather stable within the individual when learning conditions are the same for all participants (i.e. they all followed the same curriculum). This does not license claims about the effects of a specific aptitude or metalinguistic training on proficiency. An experimental design with treatment and control groups would be insightful in this respect.

As discussed in Chapters 3, 7 and 8, motivational dimensions of foreign language learning are undoubtedly an important aspect of individual variability in language learning. Our results corroborate previous findings that children are generally keen to learn foreign languages at school and that their motivational dispositions remain largely stable over the course of two academic school years.

Accordingly, the most notable difference was not identified across measurement points, but between the two target languages English and French. Motivational dispositions were both higher and more stable for English than for French, where they drop slightly over time. Further research, integrating classroom observational and other qualitative data, could investigate whether and which specific pedagogical dispositions could counteract drops in motivation to learn French as a local foreign language vis-à-vis English as a global lingua franca.

Interestingly, the motivational patterns overall are the same, irrespective of the place of residence: The motivation to learn L2 French is not positively affected when children live close to the French speaking community, as in LAPS I (see Chapter 7). This result would be worth following up from a sociolinguistic perspective, to get a better grasp of the mechanisms that shape the children's rapport with the target languages, these languages' status in the community, the extent of language contact, as well as the need to speak them in order to feel integrated.

To conclude, we are cautiously optimistic that the evidence discussed and presented in this book is a valuable contribution towards a better understanding of individual differences in foreign language learning in a state school setting. We are convinced that the quantitative approach chosen by our group is a possible way to identify generalizable associations between variables. In our take on the topic, we use statistical techniques that should help prevent overfitting of models to data. We used factor analyses (first exploratory, then confirmatory), cross-validation techniques, data partitioning in training and test sets to shed light on the internal dimensionality of our constructs and on the prognostic value of the measures that we took. With very few exceptions, our test instruments were not standardized. Therefore, a direct comparison with other findings in other linguistic contexts is difficult. Also, as argued in Chapter 6, it is easier to estimate the association between constructs if the error with which they have been measured is known. In future endeavours such as ours, it would therefore be important to use more standardized material whose reliability is both known and substantial.

3 Individual differences among scholars in individual difference research

This book is the result of a group effort spanning over roughly five years. Many people who did not co-author chapters have contributed to this project (see acknowledgments in Chapter 0). The chapters of this book are the manifest outcome not only of many hours of joint and individual work, but also of a great

many intensive discussions among the different members of our group. Our group, too, is characterized by individual differences – with respect to research interests, epistemological stances, and language teaching, learning and research experiences. These differences in experience and points of view led to rather lively discussions within the group. Some divergences had to be overcome to arrive at a publishable product with which we all can identify:

Many discussions revolved around questions regarding the selection of the optimal statistical model to test a specific hypothesis, and on the number of models and model variants that should be fitted. The most challenging debates, however, were in the implications of the results of our analyses: What can we infer from the list of variables in the *No costs spared* model in Chapter 4? Are the results in the chapters fitting regression models “good enough” or rather disappointing in terms of variance “explained”? When exactly do associations in structural equation models indeed reveal contribution of one factor to the other and not just association without any further causal meaning? On a more global level, formulating and prioritising the objectives of the project was an ongoing source of discussion – was our main goal to assess the feasibility of prognostication or the analysis of dimensions underlying individual differences? And if both were important, to what extent would such multiple goals compromise the efficiency of the research design? How important is it to do research that provides immediate and applicable answers to practical pedagogical issues? Is it better to measure a multitude of constructs and address a multitude of sub-questions in order to do justice to the “complexity” of multilingual language learning or should we concentrate rigorously on one important research question?

Such discussions are both necessary and difficult. To make the interpretation process even more intricate, some of the manuscript’s reviewers pushed for much bolder conclusions and claims to be drawn from our results than those formulated by us.

Our policy was to discuss these problems repeatedly and extensively. The ultimate decision on the matters pertaining to modelling strategies and to the limits of interpretation, however, were left to the first authors of the chapters respectively. Overall, this approach led to increasingly cautious interpretations of our findings, which may make the read of the book appear overly modest to some. Although the empirical effort might at times seem disproportionate when put in relation to our conclusions, we prefer it that way.

The field of Second Language Acquisition and Multilingualism studies has produced a great wealth of mostly small-scale studies that are often associated with relatively bold and general claims when it comes to the interpretation of the

results. Some such studies from our discipline are (selectively, occasionally) referred to by policy makers, with respect to various aspects of pedagogical policy decisions (curricula, definition of learning outcomes, teaching materials, etc). Research that informs policy must worry about the generalizability of its findings and interpretations. As in other disciplines (cf. Ritchie 2020 for examples mostly from science and psychology), the interpretation of “significant” (or non-significant) results in small-scale studies presents particular and often unrecognized risks for generalizability.

Our group tried to address these problems responsibly by acquiring as much data as was feasible, by collecting two independent data sets from two different contexts and testing patterns emerging from the first with the second, and by carefully cross-validating and testing the models according to the best of our statistical knowledge. To make the results verifiable and future research possible, we publish our data and scripts on the *osf.io* platform, for fellow researchers to download, use and most importantly improve them. We look forward to more discussion with a wider scholarly audience.

References

- Berthele, Raphael. 2019. Policy recommendations for language learning: Linguists’ contributions between scholarly debates and pseudoscience. *Journal of the European Second Language Association* 3(1). 1–11. DOI: [10.22599/jesla.50](https://doi.org/10.22599/jesla.50).
- Conway, Andrew R. A., Michael J. Kane, Michael F. Bunting, D. Zach Hambrick, Oliver Wilhelm & Randall W. Engle. 2005. Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review* 12(5). 769–786. DOI: [10.3758/BF03196772](https://doi.org/10.3758/BF03196772).
- Deci, Edward L. & Richard M. Ryan (eds.). 2002. *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Dörnyei, Zoltán. 2010. The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In Ernesto Macaro (ed.), *The Bloomsbury companion to second language acquisition*, 247–267. London: Bloomsbury Academic.
- Fürst, Guillaume & François Grin. 2018. Multilingualism and creativity: A multivariate approach. *Journal of Multilingual and Multicultural Development* 39(4). 341–355. DOI: [10.1080/01434632.2017.1389948](https://doi.org/10.1080/01434632.2017.1389948).
- Parry, Thomas S. & Charles W. Stansfield. 1990. *Language aptitude reconsidered*. Englewood, NJ: Prentice Hall Regents.
- Plomin, Robert. 2019. *Blueprint, with a new afterword: How DNA makes us who we are*. London: Penguin Books.

- Ritchie, Stuart. 2020. *Science fictions: How fraud, bias, negligence, and hype undermine the search for truth*. New York: Metropolitan Books.
- Shan, Yiming. 2020. Whether successful language learners require intrinsic motivation. *Open Journal of Modern Linguistics* 10. 549–559. DOI: 10.4236/ojml.2020.105031.
- Skehan, Peter. 2002. Theorising and updating aptitude. In Peter Robinson (ed.), *Individual differences and instructed language learning*, 69–93. Amsterdam/Philadelphia: John Benjamins.
- Stromswold, Karin. 2001. The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language* 77(4). 647–723. DOI: 10.1353/lan.2001.0247.