

Adam Świątek

Philological School of Higher Education, Wrocław

From a Simple Testing Tool to Advanced Computer-based MLAT Test: A Century of Aptitude Testing

Abstract: The beginning of the aptitude concept, as well as aptitude testing, or, in other words, measuring one's predispositions to foreign languages dates back to the 1920's when the first aptitude tests, like the Iowa Foreign Language Aptitude Examination, or the Luria-Orleans Language Prognosis Test came into existence (Carroll 1962). Since that time, aptitude tests have gone through a multitude of different transformations — from simple testing tools that resembled intelligence tests to an advanced computer-based version of the most influential Modern Language Aptitude Test developed by Carroll and Sapon in 1959 (Dörnyei and Skehan 2003). Apart from them, researchers from different countries have attempted to create their own unique versions of the MLAT, like the Hungarian HUNLAT battery (Safar and Kormos 2008), the Polish version named TUNJO (Rysiewicz 2011), or even the CANAL-F battery based on an artificial language (Kocic 2010). Therefore, the main goal of this paper is to provide a thorough theoretical analysis and review of the available aptitude testing batteries and find the differences and similarities between them. What is more, the paper aims to describe the components of all the possible aptitude tests and discover the potential behind the testing tools that examine one's natural predispositions effectively. Apart from the general knowledge about aptitude testing available anywhere nowadays, it is necessary to understand how the tests work, and what they expect from a participant taking part in such an initiative. As they are often compared with intelligence tests, the purpose of this paper is to show that aptitude tests constitute a different tool, and measure different abilities and skills than a set of intelligence related instruments. To reach this goal, I examine the tools available, describe their properties and potential success rate, analyze their components and compare them with the other batteries.

Keywords: language aptitude, aptitude tests, testing batteries, MLAT, test components, success rate

1. Introduction

Real-life situations and events, when foreign language use is a must for certain people, allow, or even force language users to think about the idea of being predisposed to learning foreign languages more frequently. Failure and the inability

to convey the message in a proper way, which are usually the most probable motor-vehicles of thinking about the idea of potential inborn skills limiting individual effectiveness when conveying a message or fulfilling the established goals, bring a variety of questions that need particular answers. It is a popular trend among language users to ask each other why some face fewer difficulties as far as adopting a foreign language system is concerned. The doubt is related to the fact that for some people it seems to be easier to learn, analyse, understand or join one element in a language to another, whereas for others it might be much more difficult to relate and adopt language elements, such as grammar, vocabulary, or intonation. Therefore, doubts concerning aptitude seem to be justified, which makes it worth considering and analysing their sources.

The current state of research concerning aptitude, or simply predispositions, constitutes a wide range of batteries and an examination of their component parts. Furthermore, the available battery of aptitude testing tools allows researchers to analyse the 'ingredients' of the tests and reconsider the components in a variety of contexts so that innovative aspects can be introduced and new batteries invented. That is why a thorough analysis and discussion concerning aptitude needs to begin with a variety of questions.

Are we predisposed to foreign languages? Are some people more likely to adopt specific linguistic skills faster and more effectively than other language learners and users? Finally, what are the tools available to measure one's abilities and aptitude towards language learning? How are they constructed and how effective do they seem to be?

The above-mentioned research questions are only some of the problems that need answers. The following paper provides an overview of the major available aptitude testing batteries, not only for adults, but also for children who are already able to take part in such tests in order to check their inborn abilities. Therefore, the reporting character of the article aims to gather in one place all the available information concerning the concept of aptitude from a multitude of different sources. The article is a report on the current state of development of the concept of aptitude and contains a componential and qualitative analysis of the concept, including a comparison to intelligence tests, which are frequently contrasted with aptitude batteries. The paper opens with the 1920s, when the first batteries were invented and gave rise to further research, and goes through the 1950s, when the Modern Language Aptitude Test appeared as a revolutionary tool. Furthermore, the paper supplies the reader with an overview of all the aptitude tests that have so far been invented over almost the last 100 years in order to celebrate their upcoming 100th anniversary.

2. The concept of aptitude

The article opens with a variety of definitions concerning the phenomenon of aptitude aimed at familiarising readers with a variety of meanings of that concept, as well as allowing them to understand the information included in later parts of the article properly. The second part aims to provide readers with historical information related to the need for public examination development and the concept of aptitude itself, from early research in the 19th century to the 1950s, when a real breakthrough marked the direction of further research on aptitude testing and batteries.

2.1. Definitions of aptitude

Academic reflection on the concept of aptitude itself, as well as aptitude testing, dates back to the 1920s, when the first aptitude tests were developed (Carroll 1962: 91). Since that time, aptitude tests have been defined in a multitude of different ways, although they all contain a certain degree of similarity, which is why none of the definitions seems to be exceptionally unique. According to Ellis (1994: 36), aptitude is a particular set of inborn abilities to process and adopt language data in a much more efficient way. What is more, he relates aptitude to intelligence, stressing the fact that certain concepts seem to be different, thus making aptitude a distinctive idea. Brown (2000: 391) described the idea as an ability to learn languages more successfully. Furthermore, Biedroń (2011: 472) claims aptitude to be a certain readiness to initialise language absorption, as well as a possible rate of successful completion of the initiative. Another piece of the explanation comes from Snow (cited in Robinson 2007: 256), who claims that aptitude is the specific equipment allowing its owner to deal with particular tasks in a variety of situations. However, one of the most significant descriptions comes from Carroll (1962: 89), who is one of the pioneers of the concept of aptitude, as well as being the co-creator of the Modern Language Aptitude Test. According to Carroll (1962: 89), aptitude is the facilitator of language learning and making use of a particular language in a variety of communicative situations. Aptitude is also a factor which is unrelated to intelligence and, therefore, independent of its characteristics; in other words, intelligence does not influence aptitude and both concepts are independent. Moreover, aptitude is an ability possessed by a limited group of humans. However, there is one more definition of aptitude that comes from Carroll (1962: 89–90), but a little later, and which therefore ought to be mentioned here. He defined it as a predictor of language learning success and progress that a student will soon make and it is independent of the wishes of the language learner (Safar and Kormos 2008: 3). Finally, Dörnyei and Skehan (2003: 590) provide a definition based on another, simplified interpretation of the concept. They claim that aptitude is simply an inborn ability which determines

the degree of success in foreign language learning, as well as differences between students.

Summing up, there is no doubt that all the definitions contain certain common aspects. They define aptitude as an inborn ability and a facilitator of language learning and data processing. Furthermore, aptitude determines an individual rate of success in a variety of communicative situations and readiness to initialise language absorption.

Finally, it should also be mentioned that language aptitude is conceptualised not as a monolithic but as a componential ability and this is the reason why a set of tests are used to measure it. However, if the concept of aptitude has been highly developed, what are the available instruments to measure the rate of probable achievement in foreign language aptitude? What are the components of such tests and how do they work?

The following section of the paper aims to explain the historical background of the development of aptitude testing tools, the reasons for their appearance, and the nature of early aptitude tests developed between the First and Second World War.

2.2. Historical background and early research on aptitude testing

As far as early research on aptitude tests is concerned, it began in the first half of the 20th century when American researchers became interested in and began to focus on the concept of aptitude in greater details and decided to pay more attention to formal testing. However, before that, there were a number of events that led to this decision, or at least set in motion the proper course of action for the development of aptitude testing tools (Spolsky 1994: 26–27).

The initial stage of introducing public examination concentrated on the idea of screening candidates for the army and was entirely devoted to the selection of soldiers who were to serve in the Indian Civil Service, already popular in England in the 19th century when India was under British control. However, to serve in India a soldier had to show a good command of the foreign languages spoken in the colonies. Due to the fact that language training was very expensive, it was necessary to introduce public forms of testing candidates in order to choose the ones who showed language potential and a high rate of success when chosen to be taught the target languages and then serve in the army. Furthermore, public examination was also supported by the French emperor, Napoleon (quoted after Spolsky 1994: 27), who already used public examinations as potential tools for centralised control. Moreover, public examinations were then recognised as rather a political subject, with no specific educational value or purpose (Spolsky 1994: 27).

In the late 1890s, Edgeworth (quoted after Spolsky 1994: 27) claimed that public examinations were not truly reliable and a battery of innovative and fully objective tests was a necessity. As a result, after 1918 a set of objective examination tools were invented and began to dominate in Britain and the USA. There were a

number of breaks caused by numerous conflicts, as well as the First World War (the most serious worldwide conflict up to then). However, the true origin of aptitude testing tools, or, at least tools that attempted to resemble them, was the USA, where the public school system accepted and supported the development of such tools to increase the effectiveness and lower the costs of language training. As a result, American researchers offered an initial set of aptitude tests in the 1920's (Dörnyei 2010: 249).

The most important factor behind the selection of proper candidates in terms of language training was money. The government aimed to be sure that the trainees selected for language programmes were sure to be successful due to the fact that the proper training of a single candidate was very expensive (Dörnyei 2005: 34). That is why it was necessary to select proper groups of candidates who already possessed certain language skills and potential and, therefore, to examine them through the use of aptitude testing batteries to ensure a high rate of efficient and effective language training. The first set of aptitude tests, or prognosis tests, as they were also called at this time appeared between 1925 and 1930 (Spolsky 1994: 29–30).

The first test came in 1925 when Stoddard and Vander Beke (quoted after Carroll 1962: 91) developed the so-called Iowa Foreign Language Aptitude Examination. The tool consisted of six components; three of which dealt with grammatical aspects of the English language, i.e. singular and plural forms, then nominalisation and tenses, whereas the remaining three concentrated on Esperanto and focused on aspects of grammar and vocabulary. However, one of the subcomponents also dealt with the translation of sentences from Esperanto into English (Spolsky 1994: 29–30).

In 1928 another tool was invented. The Luria-Orleans Modern Language Prognosis Test (Carroll 1962: 91) took a different form than the previous test due to its layout. The tool consisted of vocabulary tasks involving word memorisation and searches for cognate word, as well as literal, grammatical translation in two languages, French and English. The test lasted 1 hour and 25 minutes (Spolsky 1994: 30).

Another prominent test at this stage of research on the concept of aptitude was the 1929 George Washington University Language Aptitude Test. The tool was quite distinctive in that it dealt with particular elements of an artificial language. Finally, in 1930 the Symonds Foreign Language Prognosis Test was developed (Spolsky 1994: 30).

Unfortunately, all the above-mentioned batteries were criticised for their similarity to intelligence tests. Kaulfer (quoted after Spolsky 1994: 36) presented a similar opinion in terms of the concept of aptitude and in 1939 published a review of the 1928 and 1930 tests and claimed the instruments to be very similar to intelligence tests. He also stated that the data obtained through the tests was unreliable and lacked sufficient evidence for specific predispositions. Furthermore, he also

drew attention to the fact that aptitude testing tools ought not to measure general linguistic skills, but rather a multitude of particular skills that could be helpful and effective in a variety of contexts and situations. This information might have been the key to further aptitude tests development (Spolsky 1994: 36–37).

The disputes between researchers were interrupted by the Second World War. The war enforced a longer break in research on the concept of aptitude itself, as well as aptitude batteries. However, the discussion was revived after the war and a huge wave of important and justified criticism came no later than in the 1950's from Carroll. As it emerged later, he played the most significant role in the development of aptitude testing (Spolsky 1994: 41–42).

Finally, there was one extra aptitude tool created before Carroll's decisive initiative in that field. It was the Army Language Aptitude Test (ALAT) developed to measure speaking as well as reading aptitude, but only for the purpose of Western Indo-European Languages. The test consisted of 57 questions which were based on artificial syntactic rules that resembled the structures of those languages. Unfortunately, the test was not given too much attention due to the fact that a participant who was ready to adopt a few simple rules or vocabulary items could easily get impressive results (Parry and Child 1988: 34).

The ALAT was a tool that was mainly criticised for its length and the time allowed for it, which was only 27 minutes. Only some of the students participating in it were able to complete the test within time limit. The end of the test contained the most complicated questions, which needed time to be solved, and not many students were able to manage them. However, the tool had one advantage. It was very simple to administer, mainly due to the amount of time necessary to complete the questions (Parry and Child 1988: 35).

Carroll's contribution in the field of the concept of aptitude at the end of the 1950s brought about real changes (Parry and Stansfield 1988: 3).

Early aptitude tests did not exert much influence on the field of aptitude testing. However, they constituted a comprehensive introduction to the further research and development of the concept. They were mainly prognosis tests which focused on general, rather than specific language skills. Moreover, the tools were still very similar to intelligence tests and the components were not fully reliable. Therefore, the evidence was easy to question and raised numerous doubts about the value of the scores obtained by the participants. As far as their strengths are concerned, the tests were easy to administer and provided clear evidence that numerous changes had to be implemented to make them comprehensive aptitude measuring tools.

3. Aptitude batteries

The second section of the article aims to familiarise the reader with all the aptitude batteries that have been invented and implemented in the field of aptitude testing since the end of the 1950s. Furthermore, an overview of the integral parts of particular aptitude batteries, together with their aims and proper fields of application is provided. Therefore, the section opens with the Modern Language Aptitude Test (MLAT) due to its enormous contribution to and influence on this field of linguistics.

3.1. The Modern Language Aptitude Test

Carroll (1962: 89–91) claimed aptitude to be a concept that is totally separate from intelligence and should not be compared to it. He stated that aptitude is a special inborn talent that allows language learners to acquire foreign languages faster and more effectively. What is more, he claimed that aptitude was responsible for the fact that students faced fewer difficulties when solving potential problems and linguistic intricacies. That is why the terms aptitude and intelligence ought not to be used interchangeably. Moreover, he assumed that the concept of aptitude varies in terms of individual effectiveness and it had to be adapted to serve as a reliable basis for constructing tests that would effectively screen candidates for language training courses (the process of developing and administering an aptitude test is both time-consuming and quite costly).

The procedure that led to the creation of the MLAT was quite complex and time-consuming. In the 1950s Carroll and Sapon (quoted after Dörnyei and Skehan 2003: 591) created a variety of tests that were supposed to measure one's natural skills for learning languages faster than other learners. There were more than 40 tasks that were, one by one, solved by students and evaluated by the authors of the research. Then, the two researchers analysed all the tasks the students had solved and searched for specific similarities between the batteries, e.g. a correlating high score performance. As a result, numerous tasks were eliminated and the research provided a multitude of important conclusions (Dörnyei and Skehan 2003: 591–593). According to Dörnyei (2010: 249), the MLAT, as well as its closest relative, the PLAB (to be described later) were developed using the so-called 'trial-and-error' method, which is a suitable way to describe it, taking the above-mentioned information into consideration.

According to Dörnyei (2005: 35) the real breakthrough came in 1959 at Harvard University. After a multitude of trials and extensive research on the concept, under the supervision of the Psychological Corporation, the Modern Language Aptitude Test was published. It turned out to be a revolutionary tool in the field (Biedroń 2011: 472). The development of that test (as well as the PLAB) was often called "the second wave" (Dörnyei 2010: 249) of revival of interest in

apitude testing which, as I mentioned before, occurred after World War II, and its further consequences for the world (Dörnyei 2010: 249).

However, to start with, Carroll (quoted after Dörnyei 2010: 249) developed a 4-dimensional model of the components of aptitude which served as the basis for the establishment of further tests. The model contained the following elements:

- phonetic coding ability
- grammatical sensitivity
- rote learning ability
- inductive language learning ability

(Dörnyei 2010: 249)

The last component of aptitude, i.e. rote learning ability, is also quite frequently given another name — associative memory (Dörnyei and Skehan 2003: 592).

Carroll (1988: 14) claimed that aptitude tests ought to be based on all the skills mentioned above and, thus, tests created on such a basis might be the proper tools to examine one's special abilities in terms of foreign languages. As a result, all the factors were combined together and gave rise to a new and modern tool for measuring foreign language aptitude, i.e. the Modern Language Aptitude Test. The tool consisted of the following components:

- number learning
- phonetic script
- spelling clues
- words in sentences
- paired associates

(Carroll 1988: 14)

At this stage, it is necessary to mention that the second, fourth and fifth of the above-mentioned points represent the components of aptitude mentioned in the first, second, and fourth points of the previous list, whereas number learning and spelling clues resemble the first, third, and fourth components of aptitude (Rysiewicz 2008: 574). Complicated as it seems to be, the model brought about an inevitable change in the way researchers approached the concept itself. Besides, it was the first aptitude test that was considered reliable and worth following, inasmuch as the scores were more helpful than those that resulted from intelligence tests (Kocic 2010: 236).

Since 1959, the Modern Language Aptitude Test has greatly influenced the development of aptitude testing tools and research on it, and has been the dominant model for linguists and students all over the world. Moreover, it can easily be called the driving force behind the concept of aptitude and the tool that gave rise to most, if not all the other tools that appeared later (Kocic 2010: 236).

According to Biedroń (2011: 472), the MLAT became a certain type of paradigm that laid down the right path to follow in the creation of new batteries in the future.

3.2. The Pimsleur Language Aptitude Battery

In 1966, 7 years after the appearance of the revolutionary MLAT, another powerful language aptitude testing tool appeared. The efforts led Pimsleur to the design of the Language Aptitude Battery, which turned out to be the second most influential and important aptitude test after the above mentioned MLAT. It was the first aptitude designed not for adults, but for post-puberty and teenage learners (Robinson 2005: 48). The author of the test claimed that a multitude of high school foreign language learners fail to achieve the target due to problems with auditory skills and, therefore the PLAB was more concentrated on those elements, rather than on memorising vocabulary items. The PLAB's main target was to diagnose those problems and put an immediate emphasis on the problem areas (Dörnyei and Skehan 2003: 594). Together with the MLAT, the PLAB started to be widely used and researched in different environments (Dörnyei 2010: 249).

However, Pimsleur's view on the concept of aptitude differed considerably from the one suggested by Carroll. In his opinion, aptitude consisted of just three elements, as opposed to Carroll's view which contained four. These were:

- verbal intelligence (similar to Carroll's view on aptitude in terms of grammatical sensitivity and inductive language learning ability, especially when it comes to the use of analytical reasoning; however, Pimsleur does not concentrate on memorisation and removes this aspect from his taxonomy, instead targeting inductive skills to a great extent)

- motivation (Pimsleur treats language learning as a process going far beyond the absorption of language data and regards motivation as another prominent factor behind language success)

- auditory ability (resembles Carroll's phonetic coding ability with regard to processing strings of sounds)

(Dörnyei 2005: 39–40)

As far as the components of the test are concerned, Pimsleur Language Aptitude Battery consists of six parts, i.e:

- grade point average
- interest in foreign language learning
- vocabulary
- language analysis
- sound discrimination
- sound symbol association

(Dörnyei 2005: 38–39)

The PLAB test lasts one hour and the amount of time is almost the same as in the case of the MLAT battery which requires between 60–70 minutes to be fully completed (Dörnyei 2005: 38).

As two closely correlated batteries sharing particular elements, the tests were often compared in terms of differences which were more significant at this point and, therefore, made them two different batteries. As a result, a deep analysis in the late 1960s enumerated the considerable number of elements that distinguished both testing tools. First of all, the PLAB concentrates more on auditory skills, whereas memory seems to be a background skill. Then, it has two innovative constituents, like the ‘Grade Point Average’, and the most influential one ‘Interest in foreign language learning’. For the first time, the PLAB highlighted the fact that one’s desire to learn and be successful in language learning, together with one’s foreign language learning background, constitute a component of aptitude and an important message for the researcher (Dörnyei 2005: 36).

Needless to say, both the MLAT and the PLAB are of enormous value and take a special position in the set of available aptitude testing tools (Dörnyei 2005: 35).

3.3. Aptitude testing between 1970–1990

Between 1970 and 1990 the influence of the two significant batteries resulted in the growing interest in and need for aptitude tests and, therefore, different batteries were developed. The first of them was the York Language Aptitude Test (YLAT) invented by Green (quoted after Dörnyei 2005: 41) in 1975; another one was the Defense Language Aptitude Battery (DLAB), developed by Petersen and Al-Haik (cited in Dörnyei 2005: 41) in 1976.

Furthermore, the eighties brought another wave of aptitude tools, and in 1981 the Aptitude Test for Studies in Modern Languages was invented, whereas a year later Miller and Phillips (quoted after Dörnyei 2005: 41) created the German Aptitude Test. Finally, in 1990 the so-called VORD (which is not an acronym, but a word in an artificial language) aptitude tool was developed by Parry and Child (quoted after Dörnyei 2005: 41).

As far as the York Language Aptitude Test is concerned, its main target was to make participants create totally unknown forms from a previously unheard language system. Moreover, the test was mainly focused on inductive skills and, as a result, did not get too much attention. It is often recognised as less comprehensive and reliable than the other much more powerful aptitude testing batteries which are available (Ellis 1994: 495).

The Defense Language Aptitude Battery turned out to be a more influential test than the YLAT and, thus, gained more attention from researchers. The tool was designed for the American military forces, due to the fact that the US Army claimed that the MLAT was not a fully successful idea when considering students

with high level linguistic abilities (Dörnyei and Skehan 2003: 594). Therefore, the DLAB measured only two aspects of language skills:

- visual skills (inducing language notions from pictures)
- auditory skills (recognising foreign language sounds and grammar/sound-symbol association via numerous processes)

(Sternberg and Grigorenko 2002: 138)

However, according to Dörnyei (2005: 41), the DLAB battery was supposed to differentiate effectively between high aptitude students, mainly due to the plateau effect of Carroll's battery. However, the difference was only marginal and there was no remarkable discrepancy or increase in the predictive value of the DLAB. Therefore, the MLAT still maintained its leading position among the available aptitude testing batteries. This evidence was supported by Petersen and Al-Haik's research conducted on more than a thousand participants.

The DLAB and YLAT batteries, together with the aptitude tools that appeared in the 1980s, i.e. the Aptitude Test for Studies in Modern Languages, or the German Aptitude Test did not get much attention. As a result, there is not too much emphasis on them in the literature (Dörnyei 2005: 41).

However, something interesting happened in the case of the so-called VORD battery. The development of the tool started in 1973, whereas between 1987–88 two important American agencies — the Department of Defense together with Central Intelligence Agency — combined their efforts in order to invent a test based on an artificial linguistic system and examine its parameters so that it might be a fully comprehensive and working tool (Parry and Child 1988: 30). The tool was developed by Parry and Child in 1990 and concentrated on grammar. Its main idea was to analyse a grammatical system which was quite similar to Turkish. However, the system was artificial and this is claimed to be the biggest advantage of VORD, due to the fact that a participant is obliged to discover the system for themselves and no support or background knowledge can help. VORD is often said to be a continuation of the earlier developed ALAT (Parry and Child 1988: 35–36). The authors of the tool, i.e. Parry and Child, compared VORD to the available aptitude tests, and stated that the MLAT was still the best and most influential battery, at the same time lowering the value of their own invention (Dörnyei 2005: 41).

However, in 1998 Skehan (quoted after Robinson 2012: 64) conceptualised aptitude in a distinctive way and connected specific components of aptitude (phonemic coding ability, language analysis ability and memory) with stages of information processing and cognitive operations:

Table 1. Skehan's 'Processing Stage' Model of Aptitude

Stage number	SLA processing stage	Aptitude component
1	Noticing	Auditory segmentation Attention management Working memory Phonemic coding
2	Pattern identification	Fast analysis / working memory Grammatical sensitivity
3	Extending	Inductive language learning ability
4	Complexifying	Grammatical sensitivity Inductive language learning ability
5	Integrating	Restructuring capacity
6	Becoming accurate, avoiding error	Automatization Proceduralization
7	Creating a repertoire, achieving salience	Retrieval processes
8	Automatizing rule-based language, achieving fluency	Automatizing Proceduralization
9	Lexicalizing, dual-coding	Memory, chunking, retrieval processes

3.4. A step towards the 21st century: the CANAL-FT battery

In terms of aptitude testing the 21st century began in 2000 with another important tool offered by Grigorenko, Ehrmann and Sternberg. It was the so-called CANAL-FT (Cognitive Ability for Novelty in Acquisition of Language — Foreign Language Test) aptitude test. The name of the test seems to be quite long and, for some linguists, confusing, but its shape and length are justified, as it is an example of a tool that reflects a cognitive theory in terms of foreign language acquisition (Dörnyei 2005: 50).

As a result, theoretically the CANAL-FT was based on a number of processes, which later became the foundation bricks on which the CANAL-FT battery was established:

- knowledge acquisition process
- levels of processing
- modes of input
- memory processes (the theory mainly focused on the concept of working memory, i.e. input manipulation and processing in an active task; then, closely cooperating with the short and long-term storages to store the selected pieces of data)

(Biedroń 2012: 51; Winke 2013: 111)

Apart from those components, the battery was also based on Sternberg's model of intelligence (2005: 191–195) comprising three ingredients, i.e. analytical and

creative, correlated with the practical one. The CANAL-FT was a different and novel tool as it related to more everyday life contexts, rather than only to academically related, formal contexts. As a result, the battery dealt with naturalistic contexts and the set of initial studies gave promising scores and results. Moreover, an aptitude battery ought not to concentrate and rely exclusively on a single general aptitude test. Therefore, it would be easy to state whether the participant possesses a certain degree of natural and practical language ability, or not (Kocic 2010: 239).

The CANAL-FT battery, which is quite distinctive among all the batteries available so far, measures the following skills:

- learning meanings of neologisms from context
- understanding the meaning of passages
- continuous paired-associate learning
- sentential inference
- learning language rules

(Dörnyei 2005: 51)

The test was praised and regarded as unique especially for the fact that the battery focuses on knowledge absorption at the time the test is taken in a more naturalistic context, rather than typical school situations (Dörnyei 2005: 50). Finally, the test was based on an artificial language, like VORD (mentioned earlier in this paper). The language was called Ursulu and, thus, it constitutes another advantage of the early 21st century innovative approach towards the testing of foreign language aptitude (Dörnyei 2005: 52).

The innovative aspect of the CANAL-FT battery is the artificial language, previously unknown to participants. Every section of the battery focuses on Ursulu and students gradually build their consciousness and understanding of the new system. Participants need to comprehend new lexical, morphological or grammatical aspects within a particular period of time, which best shows their skills and potential for the acquisition of a new language system. Therefore, at the end of the test, students possess enough knowledge and awareness of the language to understand a piece of text in the new system. That is why the battery constitutes an innovative approach towards language aptitude conceptualisation (Dörnyei 2005: 52).

From the point of view of the concept of aptitude, the CANAL-FT battery seems to be unique and it has opened another chapter in the field (Rysiewicz 2008: 576–577).

3.5. The Polish contribution to aptitude testing and other modern batteries

In 2008, Jacek Rysiewicz of Adam Mickiewicz University published a Polish Foreign Language Aptitude Test (partly a version of the famous MLAT battery) named Test Uzdolnień do Nauki Języków Obcych (TUNJO), designed for adults above the age of 17. It takes almost an hour to complete the battery, which contains 120 questions and consists of 5 parts (Rysiewicz 2011: 2):

- phonetic alphabet
- hidden words
- words in sentences
- new words
- artificial language

(Rysiewicz 2008: 583)

When compared to the MLAT battery, the last element, i.e. artificial language is not included in Carroll and Sapon's aptitude tool and, thus it constitutes a completely separate element. However, analysing the parts of the MLAT, Rysiewicz decided not to use the number learning component in his test, primarily due to the practical application of the tool. The test was supposed to take no more than a single lesson at school (45 minutes), which meant shortening some of the components (MLAT 2, 3 and 4) and removing others. As a result, the number learning component (MLAT 1) was abandoned. However, the battery still dealt with four components of aptitude. Rysiewicz modified the way inductive reasoning was investigated and introduced the artificial language component based on linguistic data, instead of working with numbers on non-verbal material exclusively. That is why both tests seem to be complementary, but still contain separate elements, therefore making them distinctive batteries (Rysiewicz 2011: 7).

Creating TUNJO required a multitude of trials with high schools students. What is more, loosely adapting the MLAT, Rysiewicz had to modify the existing components and adjust them to Polish conditions and requirements (Rysiewicz 2011: 2).

TUNJO is available for individual teachers, academic institutions and researchers to diagnose and predict language skills. A researcher willing to work with TUNJO needs to contact the owner of the test, get his permission to use the battery in particular contexts and, then, share the results obtained in the study with the author, which constitutes a kind of valuable feedback aimed at improving the battery (Rysiewicz 2011: 16–17).

Among the other well-known batteries, the Hungarian Language Learning Aptitude Test (HUNLAT), designed by Otto (quoted after Hild 2007: 256), constitutes an influential tool in the field of aptitude testing. The original acronym for the test is MENYET and comes from Hungarian. The battery designed in 1996 employs 4 components or, in other words, sub-tests (Hild 2007: 256):

- hidden sounds
- language analysis
- words in sentences
- vocabulary learning

(Hild 2007: 256)

Every sub-test mentioned above corresponds exactly to a particular language aptitude component defined by Carroll and Sapon in the late 1950s. The test lasts

60 minutes and students' answers are recorded on CDs so that reliability is ensured (Hild 2007: 257).

Finally, there are other modern language aptitude tests which constitute particular modifications of the basic MLAT tool. The first one is the so-called Modern Language Aptitude Test-Elementary (MLAT-E) developed for young language learners in grades 3–6, or sometimes 2–7. It was invented as a means of examining children and looking for talented ones, as well as determining whether a particular child is ready to work with the language. Moreover, MLAT-E was mainly identified as a tool providing tutors with a helpful diagnosis and then application of the most appropriate solutions (Suarez-Vilagran 2010: 349).

The same type of test was invented for the needs of Latin American and Spanish students of English in the United States of America, i.e. the Modern Language Aptitude Test-Elementary Spanish (MLAT-ES). The battery was invented by Velasco and reviewed and approved by researchers like Stansfield and Reed, who were the dominating linguists standing behind the creation of the proper measures that might later be adopted for foreign adaptations of the MLAT battery. The test contains exactly the same components as the above-mentioned ones: (Spanish terms are mentioned in brackets) (Suarez-Vilagran 2010: 349):

- hidden words (PALABRAS OCULTAS)
- matching words (PALABRAS QUE SE CORRESPONDEN)
- finding rhymes (PALABRAS QUE RIMAN)
- number learning (APRENDAMOS NUMEROS)

(Suarez-Vilagran 2010: 350–360)

Finally, in 2006, the Computer-Based Modern Language Aptitude Test (CB-MLAT) was introduced by an organisation called the Language Learning and Testing Foundation. The tool is a computer adapted version of the MLAT paper-and-pencil test and includes all the components of the original test. The battery focuses mainly on speaking and listening. Unfortunately, like the traditional MLAT, it is only available to certain agencies cooperating with the government, or psychologists (Computer-Based MLAT).

This paper mainly shows the distance all the aptitude tests have covered since the origins of the concept, as well as the initial aptitude testing tools that seemed to be far away from those available nowadays. However, to be more precise, it is necessary to ask a simple question regarding each battery. Considering them as a whole, are they somehow different from each other, or do they constitute modified versions of the revolutionary MLAT? The following section aims to answer the question and enumerate the differences between particular batteries in the form of my personal analysis. Therefore, for the purpose of the analysis, I decided to adopt the componential approach.

4. Aptitude and intelligence

This section of the paper is devoted to the differences between two prominent concepts that are quite frequently related, that is aptitude and intelligence; however, they generally constitute two separate ideas.

First of all, as already mentioned, numerous researchers such as Carroll (1962), Ellis (1994), Brown (2000), Dörnyei and Skehan (2003), Safar and Kormos (2008) or Biedroń (2011) claim aptitude to be the concept that describes a special inborn ability to adopt languages faster and more effectively, as well as allowing the prediction of one's expected rate of language learning and ultimate success. What is even more important, aptitude is the tricky aspect that allows a person to go ahead, without the participant necessarily being willing to do so; the progress of a particular person may take place subconsciously due to the possession of innate abilities. Therefore, the participant does not need extensive instruction to remember and learn linguistic elements and makes use of them easily later.

As far as intelligence is concerned, Majer (2010: 365) describes it as a general term referring to the ability that covers a multitude of different skills, such as readiness to clarify various problems, handle logic and consider, or judge certain phenomena critically. However, it refers to both linguistic and non-linguistic aspects at the same time. According to Sternberg and Detterman (quoted after Biedroń 2012: 15) intelligence is a combination of specific skills based on three separate phenomena, i.e. the ability to internalise novel concepts and skills on the basis of previously incorporated experience, the ability to accustom to previously unknown conditions and situations, as well as so-called metacognitive ability, which stands for one's consciousness of individual mental processes and capability of controlling them. However, Sternberg (2005: 189) offers another set of interesting definitions for intelligence and claims that the concept relates to the knowledge of how to solve problems and make well-balanced judgments about the sociocultural situation one is in, especially through the use of individual strengths and successful compensation for weaknesses. Finally, the situations require language users to make use of personal analytical, creative and practical skills. That is why, when considering the following definitions, intelligence cannot be related to aptitude and seems to be something different.

There are many varied and balanced definitions of the concepts of intelligence and aptitude that could be mentioned at this stage. However, it is not the aim of this paper to concentrate on intelligence, but rather on aptitude, which is the central topic of the paper. There are researchers who claim that the terms are similar and closely related, as well as researchers who claim the terms are totally separate and independent. In my opinion, the terms refer to separate abilities and it is a good idea to have a look at a set of samples of both types of tests in order to observe some differences and the range of skills that the tests cover. Of course, intelligence tests, like aptitude tools, are not commonly available for the public;

however, with special permission, a researcher may ask the proper institution to provide access to them for scientific purposes and analyse the layout of such tools, therefore observing a number of differences that seem to be evident, even at first glance.

Summing up, there are various similarities and differences between language aptitude and intelligence tests. First of all, aptitude tests are supposed to measure inborn skills, i.e. natural talents. They do not intend to investigate one's interests, although some of the modern batteries offer such a component to some extent. As far as intelligence is concerned, it constitutes an element that can be trained and developed over time due to, for instance, social environment or extensive parental training. Moreover, it is also said that there is a high degree of intelligence inheritability and parents are partly responsible for their children's intelligence quotient. A newly born child develops his or her IQ over time through interaction and proper stimulus, such as extensive education, various games, or parental care. Aptitude is rather a stable factor and no studies have confirmed that it can be an inherited trait. However, from my personal research I have already discovered that aptitude can be subject to stimulation through the use of specific and extensive stimulus the students need to be exposed to, at least in terms of verbal skills; however, the idea still needs further investigation and soon the proper answer may be provided. Finally, intelligence tests measure one's broader range of mental abilities and how they work from, for instance, the logical or perceptual point of view. Aptitude tests, however, aim to predict one's inborn talents which will affect a learner's future performance, strengths and weaknesses.

As far as the similarities are concerned, intelligence and aptitude measure similar types of abilities, i.e. mental. However, this is still in different contexts and from different perspectives, focusing on different types of thinking and reasoning. In my opinion, both concepts ought to be regarded as independent and few similarities can be discovered.

5. Aptitude batteries — how different are they?

To start with, it is necessary to consider the similarities and differences between the leading MLAT battery and its various equivalents. The discussion aims to observe possible changes introduced into the batteries in order to make it clear that every single aptitude testing tool constitutes a separate and unique battery. As a result, a comparison between the MLAT and the other tests, i.e. the HUNLAT, TUNJO, DLAB and YLAT, needs to be conducted. However, I will also compare tests like the PLAB, MLAT-E and MLAT-ES, both separately and together with the above-mentioned batteries, due to their different target audience. Finally, the CANAL-FT battery will be discussed separately, mainly due to its specific construction (it is based on an artificial language system). Altogether, in the initial part

of this section I shall mention three comparisons, in order to finish with a general one that aims to present all the batteries in a form of one set of aptitude testing tools. Unfortunately, not all the tests can be discussed and compared with the other batteries, as little information has been published about them so far.

First of all, let me consider the constituents of the aptitude batteries, starting with the MLAT, which consists of 5 subtests and which, in 1959, set the standard for the other batteries to follow. Therefore, the basic skills that were meant to be measured by aptitude tools consisted of the following tasks: a) number learning, b) phonetic script, c) spelling clues, d) words in sentences, and e) paired associates. What is more, these tasks reflected the basic components of aptitude, i.e. phonemic coding ability, grammatical sensitivity and inductive and rote learning abilities. Are the Hungarian and Polish versions different?

As far as HUNLAT is concerned, the test does not carry 5, but 4 subtests, and differs from the MLAT to a certain extent. First of all, HUNLAT contains one element identical to the MLAT, which is ‘words in sentences’, and presents exactly the same approach towards that component. However, instead of number learning, HUNLAT introduced a similar subtest in a different form, i.e. vocabulary learning; it still corresponds to the aptitude model skills and measures rote learning ability of a particular participant. The other skills are closely correlated; spelling clues (MLAT) and hidden sounds (HUNLAT) cover similar aspects of that aptitude concept. Finally, the MLAT contains one element missing in HUNLAT, which is ‘paired associates’, as the Hungarian authors did not use this subtest in their tool.

Secondly, considering TUNJO together with HUNLAT, and at the same time contrasting them with the MLAT, the Polish tool offers, in my opinion, another approach to aptitude testing; moreover, it provides us with some new subtests as well as some that are already known. First of all, Rysiewicz did not use ‘number learning’ in the primary version of his test, but decided to leave it out and introduce a completely new idea instead. His 5-component tool introduced the so-called ‘artificial language’ component, based on an analysis of innovative content, previously unknown to the user. This element tests inductive learning ability. The other parts of the tool are somehow related to MLAT components in terms of aptitude skills: TUNJO’s phonetic alphabet can be paralleled with MLAT’s phonetic script; TUNJO’s hidden words correspond to MLAT’s spelling clues, while new words (TUNJO) reflect paired associates (MLAT). However, it must be remembered that some of the tasks may require students to deal with slightly different challenges.

The DLAB tool turned out to be a different test in terms of aptitude testing, as it measured one’s potential in only two aspects. It did not correlate too much with the traditional aspects and division into subtests like the MLAT. It focused on the auditory and visual skills of a participant and it might be more accurate to call it a more sensual approach. However, as I have already mentioned, YLAT focused on dealing exclusively with unknown forms and therefore did not attract much

attention. As a result, a comparison of that battery with the above mentioned ones might be a problematic case. The test dealt with inductive skills only and their application can be exclusively limited to a number of specific contexts.

As far as the batteries for a younger audience are concerned, the tasks included in the tools had to be adjusted to the specific age of the test-takers. Both MLAT-E and MLAT-ES contain the same subtests and measure exactly the same parameters, partly sharing the tasks with the original, adult version of the MLAT, i.e. hidden words and number learning. However, the tests also introduce some innovative tasks, and, in my opinion, implement a revived approach towards ways of measuring children and adolescents' potentials, especially through the use of such elements as finding rhymes and matching words. Of course, all the tasks were designed to measure particular aptitude components, remain in agreement with the necessary parameters and reconsider the use of a "proper" aptitude construct.

PLAB, however, comprises 6 subtests altogether, and deals with some commonly known aspects, as well as one revolutionary and previously untested element, that is, interest in foreign languages, which, from the point of view of researchers and teachers, is a priority in order to become successful in language learning. A student who shows no interest in languages might face a multitude of problems when dealing with new things and, thus, achieve a lower level of final success. Nevertheless, this component seems to be really important in terms of one's aptitude and approach towards languages and adopting their individual features, as it measures a person's motivation to succeed in foreign language learning. The other parts concentrate more on verbal and auditory abilities, whereas MLAT-E and MLAT-ES deal more frequently with phonemic coding ability and grammatical sensitivity.

Table 2 outlines the similarities and differences between the batteries described above, including the components of aptitude reflected by the particular parts of the tests.

Finally, it is important to mention the tool working on an artificial system, i.e. the CANAL-FT battery. This tool was based on a different set of assumptions and that is why the tasks it concentrates on are varied as well. Besides, it was the first test that dealt with the idea of comprehending longer pieces of text. However, the subtests also deal with such aspects of aptitude as inductive or rote learning ability, although in a different way than the other tests.

To sum up, all the above-mentioned aptitude tests, when discussed together, will always share some common features, as they deal with a common concept and, thus, need to possess at least some of the features on which the concept of aptitude is based. Every aptitude tool presented in the paper is, in my opinion, unique and constitutes an important chapter in the history of aptitude testing. Each test introduces an innovative approach and contributes to the further development of the concept, therefore, making 'space' for new concepts.

Table 2. Similarities and differences between aptitude batteries ¹

<i>MLAT</i>	<i>MLAT-E MLAT-ES</i>	<i>PLAB</i>	<i>HUNLAT</i>	<i>TUNJO</i>	<i>DLAB</i>	<i>YLAT</i>	<i>Aptitude component</i>
MLAT I Number learning	MLAT-E IV MLAT-ES IV Number learning	X	X	X	X	X	<i>Rote learning ability, inductive reasoning</i>
MLAT II Phonetic script	MLAT-E III MLAT-ES III Finding rhymes ¹	PLAB VI Sound-symbol association	X	TUNJO I Phonetic alphabet	X	X	<i>Phonetic coding</i>
MLAT III Spelling clues	MLAT-E I MLAT-ES I Hidden words	PLAB V Sound discrimination	HUNLAT I Hidden sounds	TUNJO III Hidden words	X	X	<i>Phonetic coding</i>
MLAT IV Words in sentences	MLAT-E II MLAT-ES II Matching words	X	HUNLAT III Words in sentences	TUNJO IV Words in sentences	X	X	<i>Grammatical sensitivity</i>
MLAT V Paired associates	X	PLAB III Vocabulary	HUNLAT IV Vocabulary learning	TUNJO V New words	X	X	<i>Rote learning ability</i>
X	X	PLAB IV Language analysis	HUNLAT II Language analysis	TUNJO II Artificial language	DLAB Artificial language	X	<i>Inductive reasoning</i>
X	X	X	X	X	X	YLAT Unknown language	<i>Inductive reasoning</i>
X	X	PLAB II Interest	X	X	X	X	X
X	X	PLAB I Grade point average	X	X	X	X	X

¹ According to Suarez-Vilagan (2010: 127) it has a relation to phonetic coding ability.

6. Conclusion

This paper constitutes a review and summary of the major aptitude tests that have been invented since the 1920s, together with their components. It shows the distance the tests have covered to adopt their current shape and structure; from simple, even very simplified prognosis tests in the first half of the 20th century, to advanced modern versions relying not only on paper and pencil, but also on computers, which means that the test makers have also accepted technological progress and new ways of managing the tests.

However, it cannot be forgotten that it was not until 1959 that the Modern Language Aptitude Test established the basis, or, in other words, the skeleton for the development of all the latter tests and instruments, as well as proposing the number of skills that ought to be tested through the use of such instruments. Moreover, it still constitutes the most revolutionary tool from the point of view of historical progress; it has also been adopted and modified by a multitude of countries which have developed their own foreign versions of MLAT, with a number of specific modifications in each of them.

The extensive coverage of aptitude testing in scientific literature provides an opportunity to have a look at them from different points of view and individually assess their value and usefulness in a variety of contexts. Measuring the aptitude of our students requires patience and a lot of work from tutors, but it constitutes an effective process for evaluating one's potential in terms of foreign languages, as well as diagnosing the possible problem areas. As a result, it is possible to individualise the teaching and learning process of our students as soon as the aptitude testing process is completed. The batteries currently available, with TUNJO among them, offer innovative elements and updated components, thus making them increasingly reliable tools.

In terms of reliability, the opinions of various researchers as to whether aptitude tests comprise perfect tools and are able to determine one's potential seem to be divided. Thus, some researchers claim them to be still imperfect, whereas another group, or even groups of linguists, describe them as fully comprehensive and worth following. In my opinion, a comprehensive answer will never be given due to the fact that there will always be opponents and supporters of any single concept, and reaching an overall consensus might be a problematic case. Every researcher and person wishing to work or already working with the tests has to answer the question individually and, therefore, decide whether to work with such instruments or not.

As far as the concept of aptitude itself is concerned, there are still many questions that need answers and additional research shedding new light on the concept and opening new areas of aptitude experiments and trials is needed. It would make the concept more accessible for an average person or teacher who does not always need to know how to approach it appropriately. One of the most problematic, up-to-date questions is whether aptitude is a permanent feature, or prone to changes over

time and, therefore, whether it can be somehow modified or stimulated. Another uncertainty focuses on intensive language training and its effect on aptitude. Other questions or problems include opinions about the necessary, thorough revision of the concept due to the fact that it does not concentrate much on such skills as speaking or reading, which deserve more attention and ought to be measured. Finally, it is claimed that the proper batteries of aptitude testing are inaccessible and students have no chance to prepare for them by means of solving tasks similar to the ones in the test. The same situation occurs when it comes to IQ tests. If they were commonly available, one could learn the “correct” answers and then the tests would not test inborn talents but learning efficacy. Moreover, only certain institutions can be granted access to them, having received special permission to use them, and usually they have to share their results with the supervising institution. None of them is officially published on the Internet, or anywhere else. However, in the era of the Internet, a lot of materials that appear on the Network might resemble the structure of the test to a certain extent and describe what to do and how to solve particular tasks. Therefore, how much can we rely on these tests and trust the results? Are new batteries necessary?

So far, aptitude research has shown that numerous people seem to absorb languages faster than others, mainly due to natural talent or so-called inborn skills rather than extensive training. Therefore, it can be said that some language learners are more likely to be predisposed to foreign languages and acquire multiple language skills more effectively and efficiently than others. Numerous tools, such as the MLAT, PLAB, or CANAL-FT (constructed mainly on the model of aptitude suggested by Carroll and Sapon) can be applied to measure and confirm a variety of skills that may not be investigated through the use of any other tools or types of tools, such as intelligence tests. Effective dealing with new things or possessing a ‘sixth-sense’ for sound association or rule inference seem to be rather innate skills, and not learnt ones.

I am sure that within the next few years numerous questions will be raised again. However, current and future research will definitely shed new light on this interesting area of linguistics and maybe new batteries will appear or change the way researchers view this concept. Besides, the idea of aptitude testing is being developed and worked on all the time and, for sure, each question in the paper mentioned above will have been answered when celebrating the 200th anniversary of the concept.

References

- Biedroń, A. 2011. “Personality factors as predictors of foreign language aptitude.” *Studies in Second Language Learning and Teaching* 1.4. 467–489.
- Biedroń, A. 2012. *Cognitive-Affective Profile of Gifted Adult Foreign Language Learners*. Słupsk: Wydawnictwo Naukowe Akademii Pomorskiej w Słupsku.

- Brown, D.H. 2000. *Principles of Language Learning and Teaching*. New York: Addison Wesley Longman.
- Carroll, J.B. 1962. *The Prediction of Success in Intensive Foreign Language Training*. Pittsburgh: University of Pittsburgh Press.
- Carroll, J.B. 1988. "Cognitive Abilities in Foreign Language Aptitude: Then and Now." In: Parry, T.S., Ch.W. Stansfield. *Language Aptitude Reconsidered. Language in Education: Theory and Practice*. Washington, D.C: ERIC Clearinghouse on Language and Linguistics, 74.
- Computer-based MLAT. 2014. <http://lltf.net/aptitude-tests/language-aptitude-tests/computer-based-mlat/>.
- Dörnyei, Z. and P. Skehan. 2003. "Individual Differences in Second Language Learning." In: Doughty, C.J. and M.H. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell, 89–630.
- Dörnyei, Z. 2005. *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. London: Lawrence Erlbaum Associates.
- 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.
- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Hild, G. 2007. "Investigating a Hungarian Language Aptitude Test with Think-Aloud Protocol." In: *UPRT 2007. Empirical Studies in English Applied Linguistics*. Pecs: Lingua Franca Csoport, Hungary, 255–267.
- Kocic, A. 2010. *The Effects of Aptitude on Language Learning*. *Komunikacija i kultura online*: Godina I, broj 1.
- Majer, J. 2010. "Second Language Acquisition and Foreign Language Learning." In: Lewandowska-Tomaszczyk B. (ed.) *New Ways to Language*. Łódź: Łódź University Press, 351–375.
- Parry, T.S. and J.R. Child. 1988. "Preliminary Investigation of the Relationship between VORD, MLAT and Language Proficiency." In: Parry, T.S., Ch.W. Stansfield. *Language Aptitude Reconsidered. Language in Education: Theory and Practice*. Washington, D.C: ERIC Clearinghouse on Language and Linguistics, 74.
- Parry, T.S. and Ch.W. Stansfield. 1988. *Language Aptitude Reconsidered. Language in Education: Theory and Practice*. Washington, D.C: ERIC Clearinghouse on Language and Linguistics, 74.
- Robinson, P. 2005. "Aptitude and Second Language Acquisition." *Annual Review of Applied Linguistics* 25. Cambridge University Press, 46–73.
- Robinson, P. 2007. "Aptitudes, Abilities, Contexts and Practice." In: DeKeyser, R.M. *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology*. Cambridge: Cambridge University Press, 256–286.
- Robinson, P. 2012. "Individual Differences, Aptitude Complexes, SLA Processes and Aptitude Testing Development." In: Pawlak, M. (ed.) *New Perspectives on Individual Differences in Language Learning and Teaching*. Berlin: Springer-Verlag, 57–76.
- Rysiewicz, J. 2008. "Measuring foreign language learning aptitude. Polish adaptation of the modern language aptitude test by Carroll and Sapon." *Poznań Studies in Contemporary Linguistics* 44.4. School of English, Adam Mickiewicz University, Poznań, 569–595.
- Rysiewicz, J. 2011. "Test uzdolnień do nauki języków obcych (TUNJO)." *Opis, analiza, statystyki, procedury*. <https://repozytorium.amu.edu.pl/handle/10593/1398>. (2014).
- Safar, A. and J. Kormos. 2008. "Revisiting problems with foreign language aptitude." *International Review of Applied Linguistics in Language Teaching*, 46.2. 113–136.
- Spolsky, B. 1994. "Prognostication and Language Aptitude Testing, 1925–62." *Language Aptitude Invitational Symposium Program Proceedings*. Arlington, VA, September 25–27.
- Sternberg, R.J. 2005. "The theory of successful intelligence." *Interamerican Journal of Psychology* 39.2. 189–202.

- Sternberg, R.J. and E.L. Grigorenko. 2002. *Dynamic Testing: The Nature and Measurement of Learning Potential*. Cambridge: Cambridge University Press.
- Suarez-Vilagran, M. 2010. *Language Aptitude in Young Learners: The Elementary Modern Language Aptitude Test in Spanish and Catalan*. Unpublished doctoral thesis, Universitat de Barcelona.
- Winke, P. 2013. "An investigation into second language aptitude for advanced Chinese language Learning." *The Modern Language Journal* 97.1. 109–130.