Jacek Rysiewicz
Instytut Filologii Angielskiej
Uniwersytet im. A. Mickiewicza
Poznań

FOREIGN LANGUAGE APTITUDE TEST -POLISH (FLAT-PL)

Test Uzdolnień do Nauki Języków Obcych

(TUNJO)

General characteristics, description, analysis, statistics and test administration procedures

Charakterystyka ogólna, opis, analiza, statystyki i procedury przeprowadzania testu

Poznań, 2011

DESCRIPTION OF THE TEST

Name:	*Test Uzdolnień do Nauki Języków Obcych* (Foreign Language Aptitude Test - Polish)
Test's uses:	- diagnosis of a learners' aptitude profile - prognosis of FL learning success - research
Test type:	test of cognitive dimension predominantly power test (time limits set for tasks) 'paper and pencil' test
Target population:	young adults, adults (from 17 onwards)
Testing time:	on average 60 min.
Test administration:	group and/or individual
Test's availability:	available after contacting the author
Number of parts/items:	5 parts / 120 items
Test materials:	test booklet, answer sheet, audio material to Part I (*and VI*)

TUNJO (PLAT) and MLAT compared

TUNJO is a battery of five tests measuring foreign language learning ability of (young) adult native speakers of Polish aged 17 upwards. The battery is, to a large extent, an adaptation of the Modern Language Aptitude Test (MLAT) by Carroll and Sapon (2002, [1959]). As described in more detail elsewhere (Rysiewicz 2008) the Polish adaptation of MLAT involved three major decisions. The first was to paraphrase the construct of the MLAT 3 (Spelling Clues) and MLAT 4 (Words in Sentences), which meant recreating it in Polish. The second decision was to take across the board without major changes MLAT 2 (Phonetic Script) and MLAT 5 (Paired Associates) and to include them, with minor changes to the stimulus material and mode of presentation, into TUNJO. The last decision involved creating a new task to fully represent the inductive language learning ability - a component of FL aptitude which was only implicitly present in MLAT (in MLAT 1 Number Learning). This also meant that MLAT 1 (Number Learning) had to be dropped from the Polish adaptation. This is tabulated in Table 1 below.

Table 1. Comparison of MLAT and TUNJO

MLAT	**aptitude component/s**	**TUNJO**
MLAT I „Number Learning"	phonological memory inductive learning phonemic coding	*TUNJO VI "Uczenie się Liczb" - to be included in a new version; piloting to finish in January 2012*
MLAT II „Phonetic Script"	phonemic coding	TUNJO I Phonetic Alphabet „*Alfabet Fonetyczny*"
MLAT III „Spelling Clues"	phonemic coding native vocabulary	TUNJO III Hidden Words „*Ukryte Słowa*"
MLAT IV „Words in Sentences"	grammatical sensitivity	TUNJO IV Words in Sentences „*Słowa w Zdaniach*"
MLAT V „Paired Associates"	memory	TUNJO V New Words[1] „*Nowe Słowa*"
xxxxxxxxxxxxxxxxx	inductive learning	TUNJO II Artificial Language „*Sztuczny Język*"

[1]Work is in progress aimed at including in the battery a test of working memory to substitute or supplement the existing Part V. Anticipated end of piloting – May 2012.

In the tables below the following abbreviations to denote the parts of TUNJO are used:

PhA - Phonetic Alphabet *(Alfabet Fonetyczny)*
AL - Artificial Language *(Sztuczny Język)*
HW - Hidden Words *(Ukryte Słowa)*
WinS - Words in Sentences *(Słowa w Zdaniach)*
PA - Paired Associates *(Nowe Słowa)*

Psychometric characteristics of TUNJO

The norming stage of what at the time seemed to be the final version of TUNJO began in March 2007. However, the analysis of the test's results obtained from 535 subjects (1st year students of English, 12th graders in a secondary vocational school, army soldiers and cadets of a firefighting school) revealed that the test needed some more refinement (the report on this preliminary validation study is to be found in Rysiewicz 2008: http://versita.metapress.com/content/l760666h47rh8414/?p=111911cbfdb84d7ba732a183dd6008ec&pi=6). The final version of the test which resulted from this stage was then administered to a less heterogeneous group of 650 subjects of 1st year students (English, linguistics and economics) and secondary school pupils. Although the age span of the subjects in this stage was quite big and ranged from 16:10 to 24:05 the sample studied was not representative of the whole target population of adults/young adults and as such hardly any norming of the test was possible. Table 2 below summarizes this stage of work on the test. The statistics discussed further are based on this sample only.

Table 2. Groups of subjects tested with TUNJO in 2008-2010

	Year 1 students of			Secondary School		Total
	English	linguis-tics	econo-mics	11th grade	12th grade	
2008/2009	171	89	52	34		346
2009/2010	128					128
2010/2011	148				28	176
Total	447	89	52	34	28	**650**

Descriptive statistics

Tables 3 and 4 below contains basic descriptive statistics for 650 subjects tested in the years 2008 - 2010. The largest group were 1 year students of English (69%), then students of linguistics (13%), students of economics (8%), 11th graders (5,4%) and 12 graders (4,6%). All the groups of subjects were always tested between October and December i.e., in the first three months of an academic/school year.

Table 3. Descriptive statistics for the final version of TUNJO for the groups tested (in italics - average percentage score)

Tests	**Mean**						**Standard Deviation**					
	1eng	1econ	1ling	3sec	2sec	total	1eng	1econ	1ling	2sec	3sec	total
PhA	25,7 *88,6*	25,5 *87,9*	24,2 *83,4*	24,9 *85,9*	23,9 *82,4*	25,4 *87,6*	2,60	2,40	2,96	2,77	3,38	2,74
AL	10,1 *59,4*	9,5 *55,9*	10,1 *59,4*	8,8 *51,8*	11,3 *66,5*	10,1 *59,4*	3,38	3,27	3,42	2,91	3,41	3,37
HW	22,1 *73,7*	20,7 *69*	19,9 *66,3*	22,1 *73,7*	22,3 *74,3*	21,7 *72,3*	3,70	4,06	4,52	3,60	3,53	3,91
WinS	14,5	12,3	14,7	12,8	13,9	14,2	3,25	3,32	3,43	4,09	3,33	3,39

	63	*53,5*	*63,9*	*55,7*	*60,4*	*61,7*						
PA	17,0 *70,8*	15,9 *66,3*	17,0 *70,8*	16,3 *67,9*	17,8 *74,2*	16,9 *70,4*	5,09	4,86	4,41	4,15	4,31	4,91
TOTAL	89,4 *72,7*	83,8 *68,1*	86,0 *69,9*	84,9 *69*	89,1 *72,4*	88,2 *71,7*	11,60	9,60	12,64	12,39	13,04	11,83

1eng - 1st year of English N=447; average age 19:08
1econ - 1st year of economics N=52; average age 19:10
1ling - 1st year of linguistics N=89; average age 19:08
2sec - 11th grade of secondary school N=34; average age 17:04
3sec - 12th grade of secondary school N=28; average age 18:04
total N=650

Table 4. Descriptive statistics for the groups tested - women and men.

Tests	**Mean**										**Standard Deviation**									
	1eng		1econ		1ling		sec		total		1eng		1econ		1ling		sec		total	
	W	M	W	M	W	M	W	M	W	M	W	M	W	M	W	M	W	M	W	M
N	332	115	20	32	73	16	39	23	464	186	332	115	20	32	73	16	39	23	464	186
PhA	25,6	25,9	24,4	26,2	24,3	23,8	24,7	24,1	25,3	25,5	2,65	2,48	2,62	2,00	2,66	4,16	2,89	3,40	2,72	2,81
AL	10,2	9,9	10,1	9,2	10,0	10,6	9,6	10,5	10,1	9,9	3,36	3,45	3,15	3,35	3,27	4,16	3,23	3,55	3,32	3,51
HW	22,0	22,4	19,9	21,2	20,1	19,1	22,3	21,9	21,6	21,9	3,62	3,88	3,69	4,25	4,61	4,11	3,50	3,67	3,86	4,03
WinS	14,8	13,7	13,7	11,4	14,8	13,8	14,1	12,0	14,7	13,1	3,06	3,64	3,06	3,20	3,33	3,84	3,48	3,98	3,14	3,72
PA	17,4	15,6	16,1	15,7	17,4	15,3	17,0	16,9	17,3	15,7	4,89	5,41	6,07	4,03	4,36	4,33	4,45	4,00	4,82	4,94
TUNJO	90,1	87,4	84,1	83,7	86,7	82,4	87,7	85,4	89,1	86,1	11,3	12,4	10,1	9,5	12,1	14,7	12,2	13,9	11,5	12,4

1eng - 1st year of English
1econ - 1st year of economics
1ling - 1st year of linguistics
sec - 11th and 12th grade of secondary school

The inspection of the means for the groups in Tables 3 and 4 shows that the differences between them are not big, and indeed, the analysis showed that those differences did not achieve statistical significance for any of the groups and/or sexes included in the sample. This was to be expected as the sort of 'convenience-sampling' that the present researcher was limited to did not cater for much variability across the groups studied with reference to social, demographic and educational factors. Consequently, because all the groups tested with TUNJO performed pretty much similarly to each other with respect to the dimension studied their unsuitability for the norming purposes was apparent. More heterogeneous samples have to be included in the norming stage to be done some time in the future so that the test's target population is adequately represented.

Reliability

The split-half method was used to estimate TUNJO's reliability. From among many reliability estimators the split-half seems to be an appropriate choice for the type of test represented by TUNJO: predominantly power test of cognitive dimensions with the difficulty of items progressively increasing through the task. The halving of the test was done using the 'odd-even' method and the correlations between the two halves were calculated and the result was corrected for length using the Spearman-Brown formula. Table 5 below shows the coefficients for the groups tested. All the values are within the high/very high range of the reliability estimate.

Table 5. Split-half reliability for groups tested with TUNJO

	1eng08	1eng09	1eng10	1econ	1ling	2sec08	3sec10
N	163	136	148	52	89	34	28
Age	19:08	19:09	19:08	19:10	19:08	17:4	18:4
TUNJO	**0,86**	**0,90**	**0,88**	**0,82**	**0,90**	**0,86**	**0,91**

1eng08/09/10 - 1st year students of English tested in respective years
2sec08 - 11th grade of secondary school tested in 2008
3lo10 - 12th grade of secondary school tested in 2010

Intercorrelations of parts

As a battery of tests measuring foreign language learning potential TUNJO is a multidimentional instrument whereby each of its five parts is an operationalization of a slightly different aspect of FL aptitude, thus tapping onto a different, but related, language learning ability. The analysis of the intercorrelations among the parts of the battery and the total score seem to corroborate the relative independence of the abilities captured in each test of the battery while, at the same time, contributing strongly to the more generalized, higher-order construct of aptitude. This can be seen in Tables 6, 7 and 8 where the values of the correlation coefficients are always lower among the parts of the battery and higher for the parts versus total.

Table 6. Intercorrelations of TUNJO parts for 1st year students of:

English

N=447	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,32				
HW	0,25	0,26			
WinS	0,21	0,33	0,24		
PA	0,19	0,31	0,24	0,23	
TUNJO	0,54	0,67	0,63	0,61	0,72

economics

N=52	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	**0,21***				
HW	**0,14**	**-0,12**			
WinS	**-0,07**	**0,03**	**0,07**		
PA	**0,13**	**0,21**	**0,13**	**0,18**	
TUNJO	0,42	0,46	0,51	0,46	0,73

linguistics

N=89	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,40				
HW	0,30	**0,15**			
WinS	0,28	0,51	0,26		
PA	0,28	0,34	0,33	0,34	
TUNJO	0,62	0,68	0,60	0,69	0,72

Table 7. Intercorrelations of TUNJO parts for secondary school pupils:

11th & 12th grade together

N=62	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,42				
HW	0,42	**0,24**			
WinS	0,41	0,44	0,47		
PA	0,45	0,39	0,28	0,29	
TUNJO	0,74	0,69	0,67	0,73	0,71

11th grade (age 17:04)

N=34	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,42				
HW	0,50	**0,17**			
WinS	0,53	0,34	0,46		
PA	0,41	**0,28**	**0,28**	0,35	
TUNJO	0,78	0,59	0,69	0,78	0,69

12 grade (age: 18:05)

N=28	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,61				
HW	**0,37**	**0,32**			
WinS	**0,34**	0,55	0,49		
PA	0,57	0,44	**0,28**	**0,16**	
TUNJO	0,80	0,79	0,67	0,67	0,71

Table 8. Intercorrelations of TUNJO parts for all subjects

N=650	**PhA**	**AL**	**HW**	**WinS**	**PA**
AL	0,33				
HW	**0,29**	**0,21**			
WinS	**0,22**	0,34	0,25		
PA	0,21	0,31	**0,25**	**0,25**	
TUNJO	0,57	0,66	0,63	0,62	0,71

* insignificant correlations in bold

Validity

In order to determine whether TUNJO is a valid test of the postulated ability criterion measures data in the form of scores from a set of language tests were obtained for 239 subjects out of 346 tested in 2008/2009 (see Table 2). Table 9 below shows the values of correlation coefficients for aptitude - proficiency variables for the four groups tested.

Table 9. Predictive validity data for the groups tested in 2008/2009

	Group	**N**	**Criterion**	**r**
1.	1st year students of English	106	test of grammar and vocabulary	0,33
			essay	0,31
			oral exam	0,20

			final-total	0,35
2.	1st year students of linguistics	48	test of grammar and vocabulary	0,63
3.	1st year students of economics	52	test of grammar and vocabulary skills test (reading/listening) **final-total**	0,11* 0,28 0,18*
4.	11th grade secondary school pupils	33	test of grammar and vocabulary skills test (reading/listening) **final-total**	0,01* 0,30* 0,11*

* statistically insignificant correlations

The scores for the criterion measures were obtained by the present researcher from the groups' respective language teachers/examiners who administered them to their students/pupils as part of a regular EFL course assessment. In the case of students of English and students of linguistics the tests were a part of a year final assessment, whereas for the remaining two groups they formed a part of continuous assessment. Because it was not possible to administer a uniform set of language tests to all groups participating in the validation study the criterion measures varied from group to group not only in terms of language level, content and form but also with respect to language areas tested. As regards the last point the only common language area covered by the tests in all groups was grammar and lexis. Reading and listening comprehension test scores were obtained only for the economics students group and secondary pupils, while essay and oral exam were administered only to the English group. Where more than one measure was used a value for the total test/exam score is also reported. Before the validity coefficients were calculated the scores from all the criterion measures had been standardised. Unfortunately, it was not possible to obtain the criterion tests' item-level data for reliability estimates.

In connection with the data in Table 9 it is to be observed that TUNJO's validity (low to medium) has only been demonstrated for the first two groups whereas in the case of the two remaining groups the validity coefficients did not, with the exception of the skills test for 1st year students of economics, reach statistical significance. And although the results of the same test for the group of 11th graders were close to the critical values of r for 32 df, the logic of statistical inference compels us to attribute this relationship to chance rather than to a real association between the variables studied. The pattern of the correlations in Table 9 suggests another observation, namely that there are in fact two larger groups of subjects, one being a 'language' group (groups 1 and 2) and the other a 'non-language' group (group 3 and 4). When the correlations for those two, collapsed groups were re-calculated the difference between them became even more pronounced. Table 10 below shows the result for the 'collapsed groups'.

Table 10. Validity for 'language' versus 'non-language' groups

	Group	**N**	**Criterion**	**r**
1+2	students of English and linguistics	154	test of grammar and vocabulary	0,44
3+4	students of economics + 11th grade secondary school pupils	85	test of grammar and vocabulary skills test (reading/listening) **final-total**	0,01* 0,22 0,12*

* statistically insignificant correlations

The explanation of the disparate behaviour of the two groups on the test of grammar and lexis seems to be related to the type of the language measure used (achievement versus proficiency) rather than to the nature of the sample (homogenous versus heterogeneous). As is happens,

the language tests for the 'language' group were mainly of the proficiency type, whereas for the 'non-language' group they were typical achievement tests. In the case of the first type of tests the postulated relationship between the aptitude measure and the language measure had a greater chance to be revealed as the tested domain depends more on the aptitude factor (among others) than on other, aptitude-external factors influencing success in language testing. In the case of language achievement tests the relationship sought didn't necessarily have to reveal itself, as the domain tested in the case of 'non-language' groups was the portion of a certain material covered earlier. A high score on an achievement measure is usually related to factors other than language ability, such as diligence, regularity of studying, perseverance and/or cramming before the exam. In other words, it may have happened that a person who scored low on TUNJO test but had prepared himself well to an achievement test passed it with flying colours. As a consequence of this correlational analysis will not reveal a meaningful relation between the two variables.

The tables 11 and 12 below the correlation coefficients for: TUNJO (and its parts) and a criterion measure. Additionally, intercorrelations among the parts of the TUNJO test are reported for the 'language' as well as the 'non-language' group.

Table 11. Intercorrelations for TUNJO and criterion for 'language' students (N=154)

	PhA	**AL**	**HW**	**WinS**	**PA**	**PhA**
AL	0,32					
HW	0,39	0,27				
WinS	0,11*	0,36	0,12*			
PA	0,12*	0,34	0,21	0,35		
TUNJO	0,54	0,71	0,64	0,59	0,69	
test gr/voc	0,26	0,34	0,34	0,23	0,22	0,44

Table 12. Intercorrelations for TUNJO and criterion for 'non-language' students (N=85)

	PhA	**AL**	**HW**	**WinS**	**PA**	**PhA**	**gr/voc test**	**skills**
AL	0,29							
HW	0,25	-0,04*						
WinS	0,22	0,16*	0,23					
PA	0,23	0,23	0,18*	0,24				
TUNJO	0,59	0,53	0,55	0,62	0,69			
gr/voc test	-0,04*	0,11*	-0,08*	0,02*	-0,01*	0,00*		
skills	0,04*	0,02*	0,22	0,22	0,14*	0,22	-0,25	
L2 total	-0,01*	0,12*	0,04*	0,14*	0,06*	0,12*	0,86	0,29

* statistically insignificant correlations

Bibliography

Alderson, J. C., Clapham, C., Wall, D. 1995. *Language test construction and evaluation*. Cambridge: CUP

Anastasi, A., Urbina, S. 1999. *Testy psychologiczne.* Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.

Brzeziński, J. 2002. *Metodologia badań psychologicznych*. Warszawa: PWN.

Brzeziński, J. (red.) 2004. *Metodologia badań psychologicznych*. *Wybór tekstów.* Warszawa: PWN.

Carroll, J. B., Sapon, S. M. 2002, [1959]. *Modern Language Aptitude Test*. Bethesda, MD: Second Language Testing.
Dornyei, Z. 2001. *Teaching and researching motivation.* Harlow: Longman.
Dornyei, Z. 2003. *Questionnaires in second language research: Construction, administration, and processing.* Mahwah: Lawrence Erlbaum Associates.
Dornyei, Z. 2005. *The psychology of the language learner*. Mahwah: Lawrence Erlbaum Associates.
Góralski, A. 1974. *Metody opisu i wnioskowania statystycznego w psychologii*. Warszawa: PWN, p. 34.
Grigorenko, E. L., Sternberg, R. J., Ehrman, M. E. 2000. "A theory-based approach to the measurement of foreign language learning ability: The CANAL-F theory and test". *Modern Language Journal 84/33*. p. 390–405.
Guilford, J. G. 1964. *Podstawowe metody statystyczne w psychologii i pedagogice*. Warszawa: PWN, s. 157.
Hornowska, E. 2003. *Testy psychologiczne. Teoria i praktyka*. Warszawa: Wydawnictwo Naukowe Scholar.
Hornowska, E., Paluchowski, W. J. 2004. *Kulturowa adaptacja testów psychologicznych.* W: Brzeziński, J. (ed.). *Metodologia badań psychologicznych*. *Wybór tekstów.* Warszawa: PWN. p. 151-191.
Pimsleur, P. 1966. [2003]. *The Pimsleur Language Aptitude Battery*. N. Bethesda, Maryland: Second Language Testing, Inc.
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*. p. 569-595.
Rysiewicz, J. 2009. "Measuring aptitude - Polish adaptation of Modern Language Aptitude Test by Carroll and Sapon". W: Edwardes, M. (ed.). Materiały pokonferencyjne BAAL Annual Conference 2008. London: Scitsiugnil Press. p. 93.
Rysiewicz, J. 2010. „Krytyczna analiza dwóch narzędzi do pomiaru uzdolnień do nauki języków obcych". *Studia Psychologiczne 48/3*. p. 49-60.
Skehan, P. 1989. *Individual differences in language learning*. London: Arnold.
Stansfield, C.W., D.J. Reed. 2003. "Adaptation of the Modern Language Aptitude Test and The Pimsleur Language Aptitude Battery for examinees whose first language is not English". (paper read at an annual conference of East Coast Organization of Language Testers, 20–21 March 2003, Washington, D.C.)
Strelau, J. (ed.) 2004. *Psychologia: podręcznik akademicki. Tom 2*. Gdańsk: GWP.