

Eötvös Loránd Tudományegyetem

**Towards the Validation of Translation as an Intermediate Language
Proficiency Exam task**

A fordítás mint középfokú nyelvvizsgafeladat validási lehetőségei

Summary of the Dissertation
A doktori (PhD) értekezés tézisei

Fekete Hajnal

Témavezető: Dr. Heltai Pál, CSc. Habil., egyetemi docens

Neveléstudományi Doktori Iskola
Nyelvpedagógiai Doktori Program
Budapest, 2007

Contents

Summary in English

1. The need for the present research, the theme and the aims of the of the dissertation
2. Thesis overview
3. Summary of the theoretical background
 - 3.1. Summary of the findings on test validation
 - 3.2. Summary of the findings on the construct of translation
 - 3.3. Summary of the findings on the assessment of translation
4. A tentative model of the construct of pedagogic translation
5. Summary of quantitative research
 - 5.1. Summary of findings from statistical analysis
 - 5.2. Summary of process-based research
 - 5.3. Summary of product-based research
6. Final summary and implications

Tartalomjegyzék

Magyar nyelvű összefoglaló

1. A kutatás szükségessége, a disszertáció témája és célkitűzései
2. A disszertáció felépítése
3. Elméleti háttér – összefoglaló
 - 3.1. A tesztvalidálással kapcsolatos megállapítások
 - 3.2. A fordítás konstrumával kapcsolatos megállapítások
 - 3.3. A fordítás értékelésével kapcsolatos megállapítások
4. A pedagógiai fordítás konstruktumának ideiglenes modellje
5. A kvantitatív kutatás eredményei – összefoglaló
 - 5.1. A statisztikai elemzések megállapításai
 - 5.2. A folyamat alapú kutatás megállapításai
 - 5.3. A produktum alapú kutatás megállapításai
6. Összefoglalás és implikációk

- Pym, A. (2002a). *Redefining translation competence in an electronic age. In defence of a minimalist approach*. Pre-print version 2.1.[On-line]. Retrieved October 14, 2002, from: Anthony Pym's homepage, <http://www.fut.es/~apym/on-line/training.html>
- Pym, A. (2002b). Localization and the Training of Linguistic Mediators for the Third Millenium. Paper presented to the conference "The Challenges of Translation & Interpretation In the Third Millenium", Zouk Mosbeh, Lebanon, May 17, 2002. [On-line]. Retrieved October 12, 2002, from Anthony Pym's homepage, <http://www.fut.es/~apym/on-line/training.html>
- Pym, A. (2003) *A Theory of Cross-Cultural Communication*. Pre-print version 3.3. [On.line]. Retrieved March 11, 2003, from Anthony Pym's homepage, <http://www.fut.es/~apym/on-line/cross-cultural.pdf>
- Risku, H. (2002). Situatedness in Translation Studies. *Cognitive Systems Research*, 3, 523-533.
- Toury, G. (1986). Monitoring Discourse Transfer: A Test-case for a Developmental Model of Translation. In House, J. and Blum-Kulka, S. (Eds.), *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies* (pp. 79-94). Tübingen: Gunter Narr.
- Toury, G. (1995). *Descriptive Translation Studies and Beyond*. Amsterdam and Philadelphia: John Benjamins.
- Weir, C. J. & Shaw, S. (2005). Establishing the Validity of Cambridge ESOL Writing Tests: towards the implementation of a socio-cognitive model for test validation. *Research Notes*, 21, August 2005, 10-14. Retrieved October 12, 2006, from http://www.cambridgeesol.org/rs_hsts22pdf.

Gero & McNeill, (1998). The Process-oriented Coding Scheme. Retrieved January 24, 2006, from http://www.arch.usyd.edu.au/~john/DESC9099_Tutorial_1.htm

Gile, D. (1994). The process-oriented approach in translation training. In C. Dollerup .& A. Lindengaard (Eds.), *Teaching translation and interpreting 3* (pp. 107-112). Amsterdam: John Benjamin Publishing Company.

Hatim (2001). *Teaching and Researching Translation*. Harlow, UK: Pearson Education and Longman.

Kaiser-Cooke, M. (2002). How They Do It. The Nature of Translation Expertise. In *Across Languages and Culture, 3(1)*, 59-70.

Kenny, D. (2001). Equivalence. In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 77-80). London and New York: Routledge.

Kiraly, D. (2005). Situating Praxis in Translator Education. In Károly, K. & Fóris, Á. (Eds.), *New Trends in Translation Studies* (pp. 117-138). Budapest: Akadémiai Kiadó.

Klaudy, K. (2004/2006). *Bevezetés a fordítás elméletébe*. [Introduction into the Theory of Translation]. Budapest: Scholastica.

Lörsher, W. (1991). *Translation Performance, Translation Process and Translation Strategies: A Psycholinguistic Investigation*. Tübingen: Gunther Narr.

Lörsher, W. (1993). Translation Process Analysis. In Gambier & Tömmola (Eds.), *Translation and Knowledge: Proceedings of the 1992 Scandinavian Symposium on Translation Theory*. Turku: Centre for Translation and Interpreting.

Malkjaer, K. (2001). Units of translation. In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies* (286-288). London and New York: Routledge.

Manual (see Figueras et al., 2003)

Melis, N. M. & Albir, A. H. (2001). Assessment in Translation Studies: Research Needs. *Meta, XLVI, 2*, 272-287.

Mislevy, R. J., Steinberg, L. S., and Almond, R. G. (2002). Design and analysis in task-based language assessment. *Language Testing, 19(4)*, 477-496.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement. Third edition*. (pp. 13-103). New York: MacMillan.

Newman, P. E. (1994). Translation Equivalence: Nature. In Ascher, R. E. and Simpson, J. M. Y. (Eds.), *The Encyclopedia of Language and Linguistics*. Oxford and New York: Pergamon Press.

PACTE 2003. Building a Translation Competence Model. In Alves, F. (Ed.), *Triangulating Translation* (pp. 43-46). Amsterdam: John Benjamins.

RETS (see Baker, eds., 2001)

Pym, A. (1992). Translation error analysis and the interface with language teaching. In Dollerup and Loddegaard (Eds.), *Teaching Translation and Interpreting I: Training, Talent and Experience* (pp. 279-288). Amsterdam and Philadelphia: Benjamins.

1. The need for the present research, the theme and the aims of the dissertation

The question of the relevance of testing mediation, including translation as an exam task, became a controversial issue in Hungary in the 80s and 90s, with competing approaches aiming at restructuring the scene of testing foreign language proficiency in Hungary. Criticism of mediation, and translation (from L2 to L1) as an exam task, mainly concentrated on validity and reliability issues, most of them concerning theoretical and methodological aspects of construct validity. The main concerns about the pedagogical use of translation for language testing centred on doubts a) if translation can fit the context of communicative language teaching, b) if it measures something else than foreign language proficiency, c) if assessment of translation performance can be done on a valid and reliable basis, d) if performance in the translation exam tasks needs strategies the teaching of which do not necessarily enhance foreign language proficiency in general. Thus the question that concerned testers and teachers alike was if the use of the mother tongue in testing foreign language competence, within that L2 to L1 translation was desirable and justifiable in an exam construct, especially at intermediate level.

The need for validation studies into the construct of pedagogic translation is further emphasised by the fact that accredited foreign language exams in Hungary have started the process of linking their exam levels to Common European Framework levels, the process of which needs internal and external validation. Thus the challenge for exam boards using translation as an exam task in Hungary is to find the theoretical foundation and methodological answers to the validation of their translation exam tasks.

The primary aim of the present research was to contribute to the construct validation of translation through a) exploring key aspects of the construct of translation in translation research literature, b) giving an overview of key concepts in the assessment of translation performance that concern construct validity, c) exploring exam data for relevant theoretical and methodological issues of construct validity, d) probing into new methodologies (think-aloud method and corpus-based research methods) that can contribute to the theoretical and methodological issues of construct validation.

The general aim of the present research is to sensitise teachers and test developers to key theoretical issues behind pedagogical translation and to enable them to engage in meaningful and theoretically better founded discussions about the role of translation and mediation in language testing.

The expected theory-based outcome of the present research was a) a list of the key aspects of translation from translation research literature that can be identified for the construct validity of pedagogic translation, b) a preliminary model of the construct of pedagogic translation, and c) key concepts explored in the assessment of translation performance that can contribute to construct validity. The expected outcome from empirical research in the present dissertation was methods explored for a) analysing exam data for construct validity and reliability, b) probing into the potential of think-aloud protocols in addressing response validity (a part of construct validity) and c) probing into the potential of corpus-based research for contributing to scoring validity (a part of construct validity).

2. Thesis overview

Table 1: Research design - an overview of methods used

Method	Type of method	Type of data	Type of analysis	Outcome
Literature review	Qualitative	Studies in validation	Critical reading	Validity types identified for the present research
Literature review	Qualitative	Studies in translation literature	Critical reading Exploratory	Aspects of the construct of translation, a preliminary model
Literature review	Qualitative	Studies in translation literature	Critical reading Exploratory	Key concepts in assessment of translation performance
Statistical analysis	Quantitative empirical	Exam data (annual and specific)	SPSS analysis	The methodology of statistical analysis of exam data for construct validity
Introspective (Think-aloud protocols)	Qualitative empirical	Test takers' think-aloud protocols	Protocol analysis experimental	An analytic framework for process-based analysis of translation strategies, a preliminary list of successful and unsuccessful translation strategies
Corpus-based research	Qualitative (with quantitative aspects) empirical	Test takers' translation scripts	Concordance programme analysis experimental	Potential methods recommended for scoring validation based on product-based analysis of translation performance

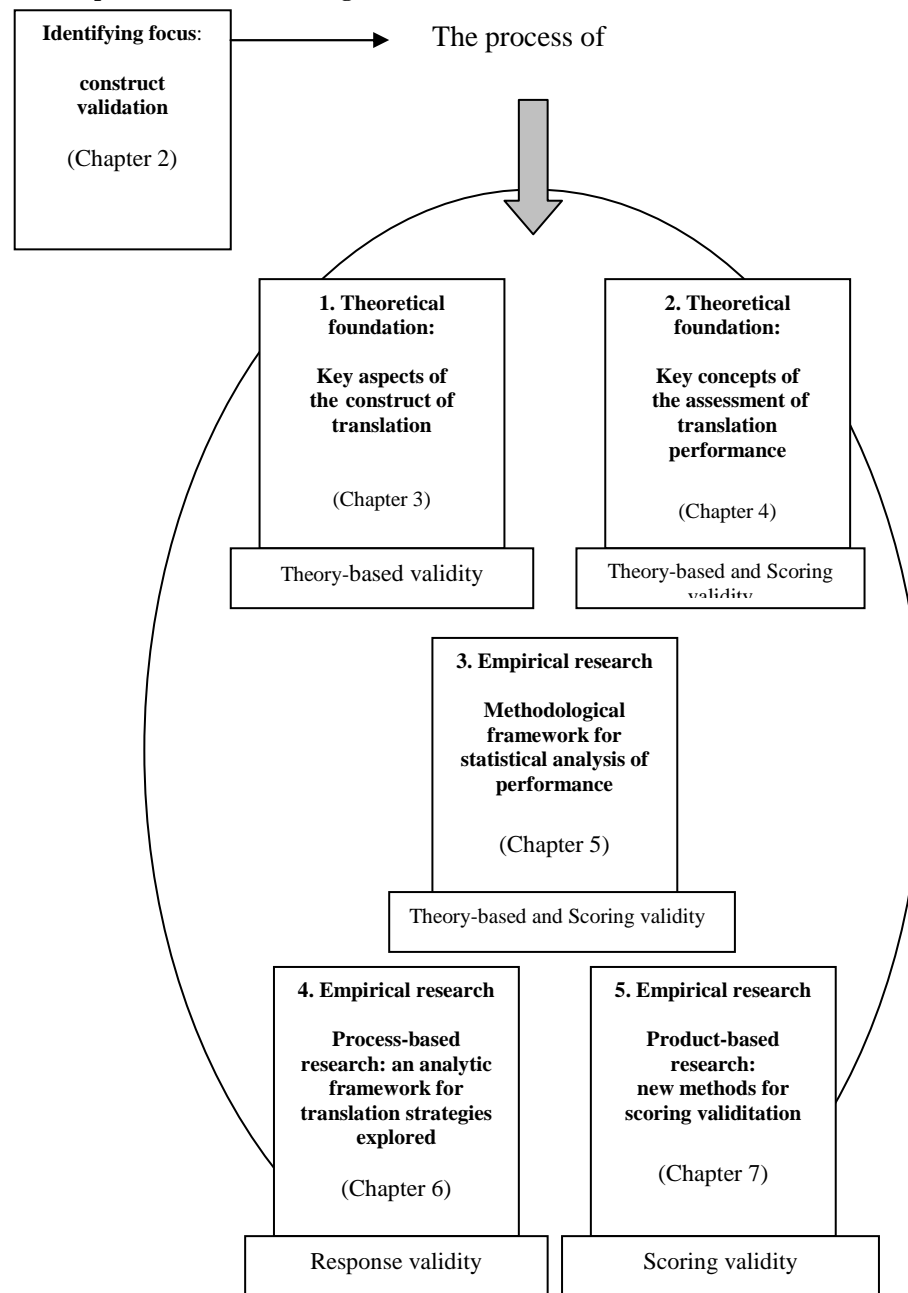
References Hivatkozások

- Alderson, J. C. & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79-113.
- Bachman, L. F. (1991). *Fundamental Considerations in Language Testing* (2nd ed.). Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baker, M. (1992). *In Other Words: A Course-book on Translation*. London and New York: The Routledge.
- Baker, M. (2001). Norms. In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 163-165). London and New York: Routledge.
- Baker, M. (Ed.). (2001). *Routledge Encyclopedia of Translation Studies (RETS)*. (Paperback edition). London and New York: Routledge.
- Bell, R. T. (2001). Psycholinguistic/cognitive approaches. In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 185-190). London and New York: Routledge.
- Bernardini, S. (2001). Think-aloud protocols in translation research: Achievements, limits, future prospects. *Target*, 13(2), 241-263.
- Brown, J. D. (2000). What is construct validity? *JALT Testing and Evaluation SIG Newsletter*, 4(2), 7-10.
- Campbell, S. (1998). *Translation into the Second Language*. London: Longman.
- Catford, J. C. (1965). *A Linguistic Theory of Translation: An Essay in Applied Linguistics*. London: Oxford University Press.
- Catford, J. C. (1994). Translation: Overview. In Asher, R. E. & Simpson, J. M. Y. (Eds.), *The Encyclopedia of Language and Linguistics*. Oxford and New York: Pergamon Press.
- Chesterman, A. (1993). From 'Is' to 'Ought': Laws, Norms and Strategies in Translation Studies'. *Target* 5(1), 1-21.
- Chesterman, A. (2005). Problems with strategies. In Károlyi, K. & Fóris, Á. (Eds.), *New Trends in Translation Studies. In Honour of Kinga Klaudy* (pp. 17-28). Budapest: Akadémiai Kiadó.
- Common European Framework of References: Learning, Teaching and Assessment* (2001). Cambridge: Cambridge University Press.
- Dróth, J. (2001). A fordítások nyelvi megvalósításának értékelése a fordításoktatás folyamatában [Assessing translation in translator training]. *Nyelvi MÉRCE*, 1(1-2), 30-36.
- Figueras, N., North, B., Takala, S., Avermaet, P. V. & Verhelst, N. (2003, September). *Manual. Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Preliminary Pilot Version*. Strasbourg: Language Policy Division.

d) a fordítási vizsgafeladat megoldásához nem szükséges olyan stratégia, amely ne lenne elhelyezhető az idegennyelv tanításának általános kontextusában.

A jelen kutatás elméleti és módszertani eredményei hozzájárulhatnak ahhoz a konstruktum validálási folyamathoz, amelyet a vizsgaközpontok jelenleg végeznek vizsgarendszerük belső validálásával kapcsolatosan a Közös Európai Keretrendszer szintjeihez illesztés során. Felhasználhatóak továbbá fordítási teljesítményt leíró értékelési skálák fejlesztési folyamatában, amely munka jelenleg a fordítás nyelvi szintjeinek KER szintekhez történő illesztése kapcsán időszerű. E két terület segítségével a közvetítőkézség validálásában alkalmazva a nyelvtudásmérésnek egy olyan területéhez járulhat hozzá, amely eddig erősen elhanyagoltnak bizonyult a nyelvvizsgáztatási kutatásban.

Figure 1: The structure of the present research
Chapter 1: Theoretical background to test validation



3. Summary of the theoretical background

3.1. Summary of the findings on test validation

Validity, in short, is systematic gathering of empirical and non-empirical evidence that, in a justifiable way, support the claims testers make in connection with the construction, administration, evaluation and use of their tests.

Messick's (1989) unified notion of construct validity has been emphasised in validation literature, which notion has led to the acceptance that there is no *one* best way to validate inferences to be made from test scores for a particular purpose (Alderson and Banerjee's, 2002). Instead there are a *variety* of different perspectives from which evidence for validity can be accumulated. These types of validity today are seen as subsumed into construct validity (Brown, 2000). The concept of construct validity is explored in depth in Bachman's (1991) seminal work on language testing.

Evidence gathered in construct validation can be *quantitative* and *qualitative* (Bachman, 1991). The quantitative types of evidence are: *correlational evidence* (correlation, factor analysis, multitrait-multimethod matrix) and experimental evidence. The qualitative types of evidence include *analysis of the process underlying test performance* (e.g. protocol analyses).

Approaches to organising validity procedures into conceptual frameworks vary (Bachman and Palmer, 1996; Mitlevy et al., 2002; Manual, 2004; Weir and Shaw, 2005). The closest to reflecting the actual process of the test design and evaluation cycle is Weir and Shaw's (2005) *theoretical socio-cognitive framework for an evidence-based validity approach*.

3.2. Summary of the findings on the construct of translation

Translation research literature has proliferated in the past few decades (as evidenced by St. Jerome Publishing, listing 27 categories in which the bibliography of translation studies itself is divided), thus there is ample literature on translation to turn to.

In the history of translation research different periods can be distinguished (Klaudy, 2004). All these periods have formulated different answers to the basic questions in translation research: a) what is actually translated?, b) for whom?, and c) in what context?

These historical periods have brought along their specific focuses (overviewed by Hatim, 2001), specific terminologies (discussed by Chesterman, 2005), distinct research methods (studies in RETS, Baker 2001) and different models (overviewed by Hatim, 2001;

a vizsgafejlesztés folyamatába rendkívül hasznos információkkal csatolhat vissza, az ilyen elemzés, amennyiben a nyilvánosság számára is hozzáférhető, segítheti a felkészítő tanárokat és a vizsgázókat is annak megértésben, hogy valójában mi okoz nehézséget a fordítási feladatokban, és így hozzájárulhatna a fordítási feladat demisztifikálásához.

Az előfordulások gyakoriságának számszerűsítésével az is vizsgálható lenne, hogy az értékelési rendszerben javasolt esetleges konkrét változtatások hogyan érintenék a vizsgázói teljesítményeket (**4. kutatási kérdés**), ez akár a pretesztelési szakaszban, adott értékelési rendszer tervezett változtatása esetén, akár vizsgafejlesztési szakaszban, az értékelési rendszer véglegesítésével kapcsolatban is vizsgálható lenne. Az előfordulások gyakoriságának vizsgálatával nemcsak az állapítható meg előre, hogy a tervezett módosítások hogyan befolyásolják a feladat nehézségét és a megfelelési arányt, hanem az is, hogy a gyengébb vagy a jobb teljesítményű vizsgázókkal szemben diszkriminálnak majd erőteljesebben.

A korpuszkutatási módszer felhasználásának legizgalmasabb területe a fordítási stratégiák empirikus kutatásának lehetősége (**5. kutatási kérdés**). Az előfordulások gyakoriságának számszerűsítésével próbaelemzés alapján megállapítható, hogy korpuszkutatás segítségével azonosíthatóak a középfokú dolgozatokban ilyen stratégiák. Megfelelő annotálással produktum-alapon nagy számú vizsgadolgozaton lehetne kutatni a sikeres és nem sikeres fordítási stratégiákat, feltárva azokat az automatizált stratégiákat, amelyek a folyamat-alapú kutatás hatókörén kívül esnek.

6. Összefoglalás és implikációk

A jelen kutatás elméleti alapú, valamint empirikus kutatásra építő megállapításai alapján a következőekben összegezhetőek a pedagógiai fordítás nyelvvizsgáztatásban való felhasználására vonatkozó válaszok:

- a) a pedagógiai fordítás beilleszthető a kommunikatív nyelvtanítás kontextusába, erre elméleti alapú és empirikus bizonyítékok is szolgálnak,
- b) a fordítás középfokon nem mér alapvetően mást, mint nyelvtudást (korreláció, MTMM mátrix, faktorelemzés), bár a szinthez képest nehéznek bizonyuló fordítási feladat esetén egy gyenge, második faktorként megjelenik,
- c) a konstruktum kutatásban felhasználható újabb kutatási módszerek (hangos gondolkodás és korpuszkutatás) hozzájárulhatnak ahhoz, hogy növekedjen a fordítási teljesítmény értékelésének validitása és megbízhatósága,

felderítésére - időhiány és terjedelmi korlátok miatt nem lehetett az összegyűjtött adatokban megjelenő minden lehetséges stratégiát beazonosítani és felsorolni. Negyedszer, a két elemzett fordítási feladatban nem jelenhet meg az összes fordítási típus probléma, így a két konkrét szöveg esetében a szövegtípus és a szöveg jellemzői eleve meghatározhatták az alkalmazott stratégiák körét.

5.3. A produktum-alapú kutatás megállapításai

A kutatási kérdések fókuszja megírt vizsgafordítások konkordancia program segítségével végzett empirikus kutatásában rejlő lehetőségek feltárása volt. A korábbi fejezetekben elemzett és a hangos gondolkodás során használt három fordítási szöveg éles vizsgadolgozatai szerepeltek a kutatásban.

Az **1.kutatási kérdésben** szereplő hipotézist, miszerint erős összefüggés mutatható ki a korpusz általánosan leírható statisztikai jellemzői másrészt a fordítási feladat nehézsége között, el kellett utasítani, minthogy sem a szó/mondat arány, sem a típus/token arány nem bizonyult a szöveg tényleges nehézségét jelző indikátornak, ha a szöveg nehézségére vonatkozó leíró statisztikai adatokkal vetjük össze azokat.

A **2. kutatási kérdés** fókuszja az volt, hogy az előfordulási gyakoriság számszerűsítése milyen potenciális felhasználási lehetőséget rejt egyrészt egyes fordítási egységek esetében a jó és hibás megoldások variációinak vizsgálatában, másrészt a vizsgázók által adott tipikus fordítás rekonstruálásában. Az elemzett adatok azt bizonyítják, hogy mind az elfogadható, mind a nem elfogadható megoldások hihetetlen változatosságot mutatnak, empirikusan is bizonyítva a fordítás idioszinkretikus, egyénekre egyedien jellemző természetét. A középfokú fordítási feladatok vizsgázók által adott megoldásait vizsgálva igazolhatónak tűnik Pym (2002) definíciója: hogy a fordítás potenciálisan azt jelenti, hogy végtelen számú célnyelvi megoldást generálunk, majd ezek közül kiválasztjuk a legmegfelelőbbet.

A **3. kutatási kérdés** a „fordítási item” potenciálisan lehetséges itemelemzését vizsgálta. A konkordancia program egy irányított legyűjtési funkcióját használva (sort by headword) valamint az előfordulások gyakoriságát vizsgálva elemezni lehet az egyes „fordítási itemeket” jósági mutató és diszkrimináció szempontjából. Az ilyen típusú, korpusz kutatás alapján igazolható eredmény, amennyiben ilyen elemzések rendszeresen végezhetőek lennének, véget vehetne az értékelők végtelenbe nyúló vitáinak arról, hogy az adott fordítási feladatban valójában mi és mennyire volt igazán nehéz a vizsgázók számára. Azon túl, hogy

discussed by Kiraly, 2005). What can be recognised and is emphasised by several scholars (Pym, 2002a, 2002b, 2003; Kiraly, 2005) is the changing social need for translation that also calls for the need to redefine basic concepts in translation research, bringing along new research methods to explore the new concepts. These historic periods have resulted in an abundance of concepts, terms, definitions and models that have all enriched our understanding of the complexity of the translation process (as reflected on in Hatim, 2001; and Kiraly, 2005) but have also brought about terminological confusion in the area of definitions of terms and concepts (reflected on by Chesterman, 2005; Kiraly, 2005).

Among recent theoretical approaches three major directions could be distinguished: *competence based approaches* (e.g. Campbell, 1998; the PACTE model, 2003), *minimalist approaches* (e.g. Pym), and *socio-cognitive approaches* (CEFR, 2001; Kaiser-Cooke, 2002; and Risku, 2002), with socio-cognitive approaches taking ground

In translation research today the interdisciplinarity nature of translation research has been widely accepted and emphasised (Baker, 2001; Klaudy, 2004) and, together with the socio-cognitive view of translation, it can be expected to come even more to the foreground, researching cognitive processes and the working of the translator's mind (Kaiser-Cooke, 2002; Risku, 2002). Such research can satisfy the need for data-based definitions of translation as opposed to speculative models, a need voiced by Campbell (1998).

3.3. Summary of the findings on the assessment of translation

The central criteria for quality assessment in translation are “equivalence” and “translation norms”, both controversial and complex issues.

When the quality of translation is addressed, linguistic and source text-based approaches seem to concentrate on some kind of definition of *equivalence* to establish relationship between ST and TT (Kenny, 2001: a *rank*-based [word, sentence or text level] vs. *meaning-based* equivalence, Baker, 1992: *textual equivalence*, Newman, 1994: *functional equivalence*). The *criteria* for defining the nature of equivalence can be *extralinguistic* (Catford, 1965, 1994) and *linguistic* (Pym, 1992). On the other hand, however, target-text and process-based approaches concentrate on the translator as observing accepted *norms* in using translation strategies and decision making (Chesterman, 1993 and Baker, 2001, referring to Toury, 1995). As in a complex and comprehensive framework to the definition of translation one can see translation as both a process and the product of such a process. In establishing the quality of any translation one can only accept

the simultaneous use of both conceptual frameworks for quality assessment (equivalence as a text-based category and norms as a process-based category).

Typical contexts for translation assessment in Hungary seem to be: a) the language teaching scene, b) the LSP translator courses scene in higher education, and c) the professional translator courses scene. These different contexts call for different considerations in assessment, different constructs underlying the notion of what is to be measured, and therefore different sets of assessment criteria for assessment (the concept of construct validity addressed) (Dróth, 2001).

The concept of *error* is more focused on in applied linguistics (foreign language acquisition) than in translation research literature. Pym's (1992) categorisation of binary and non-binary errors is an exception. On the language teaching scene Melis and Albir (2001) suggest that error should be approached from a functionalist perspective, and call for *partial assessment* (not taking into account all the factors involved in a translation) in the teaching context.

The notion of *unit of translation* is defined in the process-based approach as a relative term (Lörscher, 1993). Observations from both process (Lörscher, 1993) and product-based (Toury, 1986) empirical studies suggest that there are pedagogic implications of what is identified as a translation unit (single word lexical units vs. phrases, clauses or sentences), and also how long these translation units are. Malmkjaer (2001) suggests that the clause seems to be the generally accepted unit for translation.

Finally studies on translation as a testing device were overviewed. The studies published in English or Hungarian mainly focus on specific tasks or testing problems within testing translation or mediation. No comprehensive validation studies were found.

4. A tentative model of the construct of pedagogic translation

Based on some basic considerations from the literature review to test validation (Chapter 1) and translation research literature (Chapter 2 and Chapter 3), a *tentative model* could be suggested for the *construct of pedagogic translation* in language proficiency exams. The model assumes but does not express explicitly the findings, specified below, from the exploratory reading of translation research literature.

The model is an adaptation of Bachman's (1991) model, in which the facets of language proficiency testing are accounted for. As far as the facets of the definition of translation, the following emerged from the literature review: 1) translation has been

Az 1-6 pont megállapításai mindkét fordítási feladatra érvényesnek bizonyultak, nem csak egy konkrét szövegre.

2. kutatási kérdés: *a hangos gondolkodás révén hogyan azonosíthatók be azok a fordítási stratégiák, amelyek középfokon megfelelő fordítást eredményeznek.* Első lépésként a leírt jegyzőkönyvekben a stratégiák kerültek beazonosításra, majd két alapvető csoportra: sikeres és nem sikeres stratégiákra osztottam őket - ezek teljes listáját a függelék tartalmazta. A stratégiák listáját egy olyan tentatív, empirikus megalapozottságú listának lehet tekinteni, amely a sikeres középfokú fordítás szempontjából potenciálisan sikeresnek és sikertelennek tekinthető stratégiákat tartalmaz. Az adatok további elemzése és további kutatás szükséges ahhoz, hogy a lista véglegesíthető legyen. Egy ilyen további kutatás újabb szövegtípusokra és témákra is ki kell, hogy terjedjen.

3. kutatási kérdés: *A nyelvvizsgáztatás elméleti kutatása hogyan hasznosíthatja a fordítás, mint vizsgafeladat kutatásában hangos gondolkodás segítségével kapott eredményeket?*

Az összegyűjtött adatok segítségével az egyes vizsgázókra jellemző gondolati sémák leírhatóak. Ehhez be kell azonosítani és számszerűsíteni kell előbb a lépéseket, majd leírni a lépésekből felépülő gondolati sémákat, végül pedig leírni azt, hogy az egyes vizsgázókat ezek milyen kombinációja jellemzi - ez további kutatási lehetőségeket is rejti.

A jegyzőkönyvekben beazonosítható a főbb szakaszok jellemző száma és hossza, valamint az, hogy vizsgázónként ezek hány lépésből állnak. A fordítási szövegek nehézségére is következtetni lehet abból, hogy a vizsgázók hány lépésben tudnak megoldani egy fordítást, másrészt az egyes vizsgázókra jellemző egyéni profil is megrajzolható.

A hangos gondolkodás korlátai a középfokú fordítással kapcsolatban végzett kutatásban: először is, a vizsgálatban szereplő nők és férfiak általában eltérő módon verbalizáltak. A nők esetében sokkal több adat keletkezett, és több stratégia volt beazonosítható, így a kutatás megállapításai inkább az általuk használt stratégiákra vonatkoznak. Viszont ugyanezen fordítások produktum alapú értékelésében a férfiak voltak sikeresebbek (amit tendenciaként a 4. fejezet statisztikai elemzései is megerősítenek). Másodszor, a sikeres stratégiákat tartalmazó listát csak mint lehetséges javaslatot vagy opciókat tartalmazó ajánlást lehet elfogadni, minthogy a jegyzőkönyvek elemzése azt bizonyítja, hogy a probléma megoldásban használt gondolati sémák kombinációja egyéni jellegzetességeket mutat, így az „uniformizált” megoldások erőltetése valószínűleg kerülendő. Harmadszor, a jelen kutatás csak első lépés a módszerben rejlő lehetőségek

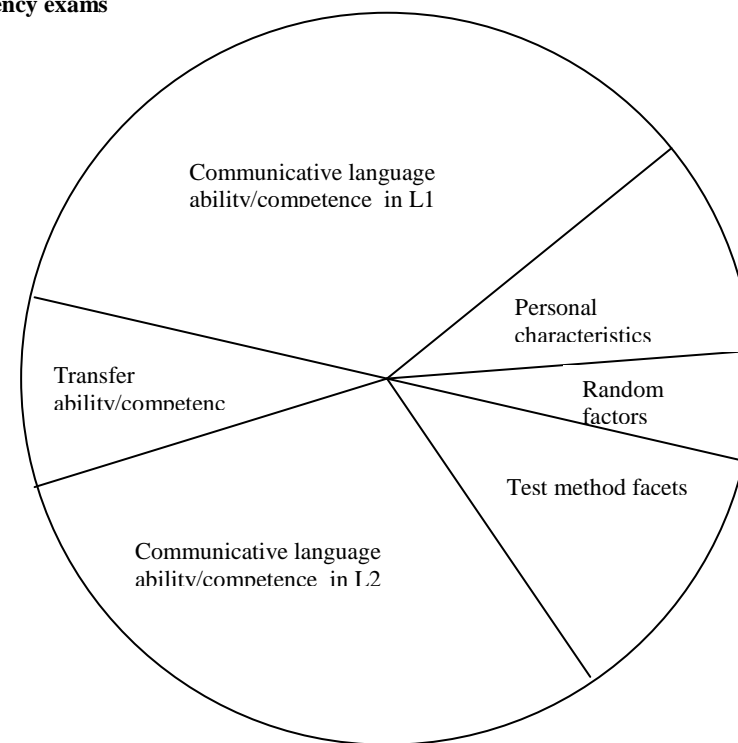
Az egyes konkrét kutatási kérdésekben az alábbi eredmények születtek:

1. kutatási kérdés: *Mi történik valójában, amikor középfokú vizsgázók fordítási vizsgafeladatot fordítanak?* Megállapítható, hogy a fordítás során nem minden kerül verbalizálásra a hangos gondolkodás révén, azok az automatikus folyamatok, amelyek a középfokú vizsgázóknál is működnek, nem kutathatók ilyen módon. Más módszerek szükségesek az automatikus folyamatok eredményének feltárásához (pl. produktum-alapú kutatás). Ami verbalizálásra kerül, alapvetően problémamegoldó jellegű. Ami valójában történik a középfokú fordítást végző vizsgázók fordítási folyamatának leírása közben, az *alapvetően nem más*, mint amit a fordításkutatási irodalom leír (Lörscher, 1991; Gile, 1994; Bell, 2001): a fordító mikro-stratégiákat (a mellékmondatok vagy az alatti szinten), és makro-stratégiákat (vagy explicit stratégiákat, szövegszinten) használ, háttértudására támaszkodik (a szöveg által nem megadott külső információra: magáról a nyelvről, a témáról, a forrásnyelvi vagy célnyelvi kultúráról, a világról, stb.), és időnként magára a fordítási folyamatra is reflektál.

A hangos gondolkodás jegyzőkönyvekben szereplő valamennyi lépése leírható a fordítási folyamat azon közös modelljében, amelyet Gile (1994) folyamatot ábrázoló modellje és Gero és McNeill (1998) probléma megoldásra alkalmazható modellje egységesítéséből létrehozható, és azonosítható a következő szakaszok valamelyikeként: megértési szakasz (a probléma azonosítása, a probléma elemzése) vagy megfogalmazási szakasz (megoldási javaslat, a megoldás elemzése). A középfokú vizsgázóknál megfigyelt fordítási folyamat azonban nem lineáris, és nem képesek hosszabb fordítási egységek automatikus folyamatok segítségével történő fordítására. A folyamat nagyon is dinamikus: gyakori a váltás a forrásnyelvi szöveg és a célnyelvi szöveg között, rövid fordítási egységeket alkalmazva, ciklikusan haladva az elemzés, szintézis és revízió folyamatában, és közben visszatérve a meg nem oldott problémákhoz. Elsősorban lexikai egységeket (szó és kifejezés) azonosítanak, mint problémás fordítási egységet (az ismeretlen szavakra és a szótározásra koncentrálnak). Kevesebb figyelmet szentelnek a mondatok szegmentálásának, nyelvtani elemzésének és egyéb makro-stratégiák használatának. Néhány vizsgázó (tanult stratégiákra támaszkodva) azonban használ ilyen explicit stratégiákat, és következetesen tudja használni őket. Az összegyűjtött adatok további ilyen jellegű elemzésében számszerűsítené is lehet majd ezek gyakoriságát. A kódolt jegyzőkönyvek segítségével a vizsgázóra jellemző gondolati sémák használata is beazonosítható, és az egyes vizsgázók probléma- megoldására jellemző képletek is felírhatóak.

increasingly defined as a *form of communication* between cultures through the language, 2) the *componentiality* of translation competence has to be addressed, 3) the *directionality* of translation (from L2 into L1 or vice versa) may have an effect on the definition of the competence and on performance, 4) there is no consensus on what the primary aim of pedagogic translation is (language teaching vs. translator training), 5) the *distance between language pairs* is an issue, 6) a *hypothesised language proficiency threshold* may exist for teaching and testing translation competence, 7) *research methods* in translation studies vary (process-, product-, and pedagogic-based approaches), 8) *new research methods* are expected to result in data-based definitions of translation, 9) an *increase in real life needs for teaching intercultural communication* at earlier levels than professional translator training is expressed in a more mobile European context, 10) a clear-cut distinction between 'school translation' and 'real translation' might be replaced by the idea of a *continuum of translation competence* from lower proficiency levels to professional translators.

Figure 2: A tentative model for the construct of pedagogic translation in language proficiency exams



5. Summary of quantitative research

5.1. Summary of findings from statistical analysis

In this chapter the main research question from the point of view of construct validation was: *To what extent does the translation task measure the same or different foreign language competence as other task types in the intermediate exam? Is there anything else measured in the translation task than language proficiency?* To be able to develop a methodological research framework for the construct validation of empirical data, the exam data from ELTE ITK intermediate exams of English were used in the chapters that followed.

Data from real exams (ITK ORIGÓ, intermediate written exams in English) were used for analyses, from two sources: annual data from academic years between 2000 – 2006 and specific exam data from three separate exam dates. The types of data used were scores in the five tasks, the total score, test takers' sexes and their age. In the analyses, the effect of six variables were examined: task difficulty, the effect of the sexes of test takers (m/f), the effect of age, task types, and underlying traits. The methods used were: Descriptive statistics Compare means, T-Test, Correlations (Pearson), Reliability analysis (Cronbach alpha), a modified Multitrait-Multimethod matrix, Multiple regression, Principal component analysis and Principal factor analysis. 6 specific research questions were formulated.

Research question 1 aimed to find out how the difficulty of translation as a task type compared to the difficulty of the other task types in the intermediate ORIGÓ exam of English. *Translation* together with *Reading Comprehension* have been found to be consistently the least two difficult task types in the six academic years, with the tendency for *Translation* to become increasingly easier.

Research question 2 aimed to examine whether the sexes of test takers (male/female) affected translation performance. Both the annual data and the specific exam data showed constant differences in the Means, both in the annual and the specific exam data. T-tests for the three *Translation* texts in the specific exams showed that a significant difference was found with the three texts (*Leonardo*, *Mayor urged* and *Arctic Meltdown*). The implication is that women may need more training in exam preparatory classes.

Research question 3 aimed to explore if age had an effect on translation performance. Practically hardly any correlation was found between age and performance in the three translation tasks, or sometimes it was even minus (worse performance as age increases). The effect of age was not found automatic or linear, but parabolic. When test

5.2. A folyamat-alapú kutatás megállapításai

A fő kutatási kérdés: *Milyen potenciális felhasználhatósága van a hangos gondolkodás (TAP) módszerének a fordítás mint középfokú nyelvvizsgafeladat válaszadási validitásának vizsgálatában. Az alap kutatási kérdés további részkérdésekre bontható: 1. Mi történik valójában, amikor középfokú vizsgázók fordítási vizsgafeladatot fordítanak? 2. A hangos gondolkodás, mint módszer használata mennyiben segíthet annak feltérképezésében, hogy milyen fordítási stratégiák használatára van szükség a fordítás feladatban való megfeleléshez? 3. A nyelvvizsgáztatás elméleti kutatása hogyan hasznosíthatja a fordítás, mint vizsgafeladat kutatásában a hangos gondolkodás segítségével kapott eredményeket?*

A kérdések megválaszolásához középfokú próbavizsgázók empirikus adataira volt szükség, hangfelvételen rögzítve a hangos gondolkodás folyamatát, amelyben a középfokú vizsgafeladat fordítása közben folyamatosan hangosan megfogalmazták azt, amit éppen csináltak (a fő kutatásban összesen 8 próbavizsgázó vett részt az első, és 7 próbavizsgázó a második fordítási szöveg esetében).

Az 1. kutatási kérdés megválaszolásához a felvett hanganyagot át kellett írni szöveggé (Bernardini (2001) átírási kódolási jelelt adaptáltam). Gero és McNeill (1998) általános probléma megoldásra kifejlesztett kódolási rendszerére alapozva egy árnyaltabb kódolási rendszert kellett kifejleszteni, a kifejlesztett rendszert összhangba hozva Gile (1994) fordításoktatásban használható fordítási modelljével. Végül az átírt jegyzőkönyveket táblázatos formába rendezve a mikro-stratégiák (lépések) beazonosítása történt, egyes lépések mellett kutatói megjegyzések megtételével. Majd a mikro-stratégiákat nagyobb egységekké vagy makro-stratégiákká lehetett szervezni. A felhasznált két feladat a 4. fejezetben statisztikailag elemzett feladatok közül került ki (*Arctic Meltdown* és *Mayor*). Az elemzések eredményeképpen egy tentatív, empirikus alapon meghatározott listáját lehetett összeállítani azon potenciálisan sikeres fordítási stratégiáknak, amelyek alkalmazása a középfokú nyelvvizsga feladatban sikeres fordítást eredményezhetnek.

A hosszas kutatás eredményeképpen alapvetően három fő tanulság fogalmazható meg határozottan a hangos gondolkodásnak a vizsgaközpontok gyakorlatában történő potenciális felhasználhatóságára vonatkozóan: a) a hangos gondolkodás nagyon hasznos és ígéretes adatokkal tud szolgálni a pedagógiai fordításkutatásban, b) a módszer használata új dimenziókat nyit a vizsgázók fordítási stratégiái és az általuk használt gondolati sémák kutatásában, c) a fordítás validálás kutatásához új és ígéretes eredményekkel járulhat hozzá.

4. kutatási kérdés: mennyire megbízhatóan méri az ORIGÓ középfokú nyelvvizsga fordítási feladata a nyelvtudást. Az elemzés alapján a feladatsor összemegbízhatósága nem növekedne, ha fordítást, mint vizsgafeladatot törölnénk a feladatsorból – ez igazolta a vizsgafejlesztők azon állítását, hogy a fordítás, mint feladattípus nem csökkenti az ORIGÓ vizsgák megbízhatóságát. A további megbízhatósági elemzések azt mutatták, hogy a fordítás mint feladat megbízhatósága a feladattípuson kívüli, egyéb tényezőktől függ.

5. kutatási kérdés: milyen kapcsolat van a fordításban és az egyéb feladattípusokban nyújtott teljesítmény között, valamint mennyire határozza meg a fordításban elért eredmény az összeredményt. A Pearson korrelációs elemzés azt mutatta, hogy a fordítási feladatok a nyelvtudás más aspektusait mérik, mint a többi feladattípus (mérsékelt korreláció).

Az MTMM mátrix segítségével végzett elemzésben, ahol a konstruktum validitásra vonatkozó állítások a konvergencia és a diszkriminációs validitások összevetésén keresztül vizsgálhatóak, csak egy gyenge korrelációs összefüggés volt tapasztalható arra vonatkozóan, hogy a fordítás és a szövegértés egyrészt, másrészt a teszt és a magyar nyelvről fordítás (500n), és végül a magyar nyelvről fordítás és az íráskészség feladattípusok a nyelvtudásnak olyan aspektusait mérik-e, amelyek erősebb összefüggést mutatnak, mint az öt feladattípus egyéb kombinációi.

A többtényezős regressziós elemzés azt mutatta, hogy mindhárom egyedi vizsgaalkalomban mind az öt feladat (jósló változók) szignifikáns volt, egyiket sem zárta ki az elemzés azon teljesítményt elől jelző változók közül, amelyek szignifikánsan növelik a modell prediktív erejét. A kapott Standardizált Beta együtthatók az egyes feladatok (jósló változók) modellre gyakorolt relatív hatását mutatták – a fordítás az első két legerősebb indikátornak bizonyult a vizsgán nyújtott teljesítmény megjósolhatóságára vonatkozóan.

6. kutatási kérdés: hány fő komponens vagy faktor található az írásbeli vizsgastruktúra egészében. A faktorelemzés fő komponens elemzés típusa (Stepwise módszer használva) azt mutatta, hogy csak egyetlen fő komponens található, amely a vizsgaeredményeket magyarázza, tehát a vizsga egy fő tényezőt (a nyelvtudást) méri. Egy további faktorelemzés (Principal Axis Factoring – Varimax rotáció) futtatásával hét változó egyidejű vizsgálata volt lehetséges: az öt feladaté, valamint az életkor és a nemek. Az elemzés célja itt is az volt, hogy megállapítsa: hány faktor/rejtett tényező mutatható ki a vizsga szerkezetében. Az elemzés eredménye azt mutatta, hogy egyes fordítási feladatok (nem a feladattípus maga), hatással lehetnek a vizsga konstruktmára, és gyenge második faktorként működhetnek.

takers were grouped further, then weighted Means showed the oldest age group (26-31) to perform constantly the best. The differences of Means between test takers aged (14-19) and those aged (19 and above), were found to be significant with two texts (*Leonardo and Arctic Meltdown*), and not significant with the *Mayor* text, a different finding from the effect of the sexes (male/female) on performance.

Research question 4 addressed the question how reliable the measures of translation performance are in ORIGÓ intermediate exams. Overall reliability of the exam would not increase if translation were deleted from the exam, thus test developers' claim that *Translation* does not decrease overall reliability in ORIGÓ exams was confirmed. Further measures of reliability showed that reliability of *Translation* seems to depend on other factors than the task type.

Research question 5 examined what kind of relationship exists between performance in the translation exam task and performance in other tasks (types), and also, to what extent translation determines the overall performance in the exam. Based on *Correlation* (Pearson) analyses, one can conclude that the *Translation* tasks do not measure the same aspects of language proficiency as the other tasks (types) do.

From the *modified MTMM matrices*, in which the concept of construct validity was explored with the help of convergent and discriminant validities, only a very weak pattern has emerged that can confirm assumptions in the construct that *Translation* and *Reading comprehension* as well as *M/C Test* and *Inverse translation* and finally *Inverse Translation* and *Writing* measure aspects of language proficiency that are more related than other combinations of these five task types.

From *Multiple Regression* analysis the conclusion is that all the five tasks (predictor variables) in all the three exams are significant, none of them were removed from among the predictor variables that significantly increase the predictive power of the model. The Standardised Beta coefficients showed the relative contribution of the tasks (predictor variables) to the model (predicted variable), and *Translation* was found among the first two best predictors of performance in the exam.

Research question 6 sought to find out how many principal factors or components can be identified in the overall exam structure. In *Principal Component Analysis* (Stepwise method) of the exam structure for the five task (types) it was found that in all the three exams only one principal component was extracted. Principal Axis Factoring (Varimax rotation) was run with 7 variables entered: the 5 tasks, plus *Age* and the *Sexes*, as two

additional variables. The purpose of the analysis was to find out how many factors/underlying traits can be detected in the exam structure. The analysis from Principal Axis Factoring suggests that certain translation tasks (not the task type) have an impact on the construct of the exam, and may function as a second factor.

5.2. Summary of process-based research

The overall research question is: *What is the potential of the use of think aloud (TAP) research as a method in establishing response validity for translation as an intermediate exam task?* This basic research question was broken down into more manageable units: 1. *What does actually happen when intermediate students translate exam tasks?* 2. *How can the use of think-aloud research help to understand what translation strategies are needed to produce acceptable translations at intermediate level?* 3. *How can language testing research profit from the use of think-aloud research of translation exam tasks?*

To be able to answer the above questions, empirical data was collected from intermediate mock exam students in the form of recorded think-aloud processes, in which they were verbalising what they were doing while translating an intermediate exam task (8 students with Text1 and 7 students with Text2 in the main research).

To answer the *Research question 1*, the recorded think-aloud processes were transcribed (transcription codes adopted from Bernardini, 2001), a coding system was developed based on a problem-solving coding scheme (Gero and McNeill, 1998), the coding scheme was designed in a way to fit Gile's (1994) sequential model of translation. Gile's model was adopted as both its relative simplicity and the categories used were expected to make it possible to fit students' terminology into this model. Finally the transcribed protocols were presented in a structured way in the form of tables, micro-strategies (moves) identified, and comments added. Later, these micro-strategies were joined into larger sequences of moves, making up either larger units (episodes) or macro-strategies. The instruments were 2 of the translation tasks statistically analysed in Chapter 4. Based on the analysis, a tentative, empirically based list of potentially successful strategies was compiled that can produce acceptable translations in intermediate translation tasks.

What emerged from my prolonged engagement in TAP research in connection with its potential use for exam centres is three considerations: a) TAP research method can yield very rich and promising data, b) this method seems to offer new dimensions for exploring

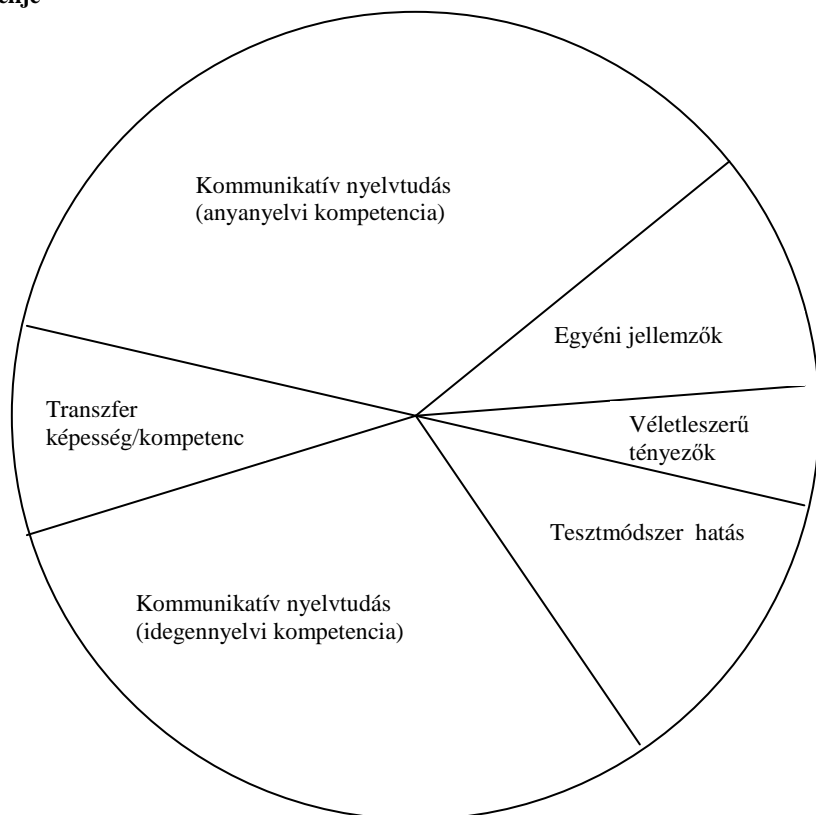
feladattípusban elért pontszámok, a összeredmény pontszáma, a vizsgázók neme és életkora. Az elemzésekben öt változó hatásának vizsgálata szerepelt: a feladat nehézsége, a vizsgázók nemének hatása az eredményre, az életkor hatása, a feladattípus hatása és a nyelvtudáson kívüli esetleges egyéb tényező. Az elemzések típusai: leíró statisztikai elemzés, átlagpontszámok összehasonlítása nemek és életkor alapján, t-próba, Pearson korreláció, megbízhatósági elemzés (Cronbach alpha), módosított MTMM mátrix elemzés, többtényezős regresszió, és a faktor elemzés két típusa (Principal component és Principal factor elemzés). Hat konkrét kutatási kérdés volt a vizsgálat tárgya.

1. kutatási kérdés: az *ORIGÓ* angol nyelvi középfokú fordítás feladat nehézsége hogyan viszonyul a többi feladattípus nehézségéhez. Az elemzésben a fordítás a szövegértéssel együtt következetesen a két legkönnyebb feladattípusnak bizonyult. A hat éves adatai alapján, megfigyelhető volt az a tendencia, hogy a fordítás az évek során egyre könnyebbé vált.

2. kutatási kérdés: a vizsgázók neme hatással van-e a fordításban nyújtott teljesítményre. Mind az éves, mind az egyedi vizsgaalkalmak adatai azt bizonyították, hogy az átlagpontok közötti különbség konstansnak bizonyult. A három konkrét fordítási szövegen elért eredményeken elvégzett t-próba azt bizonyította, hogy szignifikáns különbség volt tapasztalható mindhárom szöveg esetén (*Leonardo*, *Mayor urged* és *Arctic Meltdown*). Következtetésként levonható, hogy a nők úgy tűnik több felkészítést igényelnek a vizsgára készülésben.

3. kutatási kérdés: hatással van-e az életkor a fordítás eredményére. Az elemzés során korcsoporti bontás nélkül gyakorlatilag semmilyen korrelációs összefüggés nem volt az életkor és a teljesítmény között a három fordítási feladatban, sőt esetenként mínusz korreláció is előfordult (az életkor emelkedésével romlott a teljesítmény). Az életkor hatása nem mutatott lineáris vagy automatikus összefüggést, inkább parabolikus jellegű volt. Ha viszont a vizsgázókat életkor szerint három csoportba bontva vizsgáltuk, akkor a legidősebb csoport (26-31 év között) súlyozott átlaga alapján ők teljesítettek következetesen a legjobban. A populációt két csoportba bontva a 14-19 év közötti korcsoport átlagpontszáma a 19 – és feletti csoporthoz viszonyítva két fordítási szöveg esetén (*Leonardo* és *Arctic Meltdown*) szignifikáns különbséget mutatott, viszont nem volt szignifikáns a *Mayor* szöveg esetén – így az eredmény valamelyest különbözik a nemek hatásával kapcsolatban tapasztalhatótól.

2. ábra: A nyelvvizsgáztatásban használt pedagógiai fordítás konstruktumának tentatív modellje



5. A kvantitatív kutatás eredményei – összefoglaló

5.1. A statisztikai elemzések megállapításai

Ebben a fejezetben a konstruktum validitás szempontjából fő kutatási kérdés: *Középfokon a fordítás mint nyelvvizsgafeladat mennyiben mér más vagy ugyanolyan idegennyelv tudást, mint a többi feladattípus? Mér-e a fordítás feladat mást, mint nyelvtudást?* A további fejezetekben a módszertani keret kialakításához szükséges empirikus adatokat az ELTE ITK angol nyelvi középfokú vizsgájának vizsgálata adta.

Az elemzésekben az ITK ORIGÓ nyelvvizsgaközpont valós középfokú írásbeli nyelvvizsga adatai szerepelnek, kétféle forrásból: a 2000-2006 közötti éves adatok, valamint három különböző, önálló vizsgaidőpont adatai. Az elemzett adatok típusai: az öt

students' translation strategies and thought patterns, offering enormous research potential, and c) it can contribute with new and promising findings to translation validation research.

In the present research the following findings emerged in the three specific research questions:

Research question 1 sought to find out *what actually happens when intermediate students translate exam tasks*.

It was found that not everything gets verbalised in TAP, automatic processes with intermediate students cannot be researched in this way. Other research methods are needed to reveal automatised processes (e.g. product-based research). What gets verbalised is inherently problem solving. What happens when intermediate students translate exam tasks is *inherently not different from* how the process of *translation* is described in translation research literature (Lörscher, 1991; Gile, 1994; Bell, 2001): the translator using micro-strategies (at the level of clause, or under) and macro-strategies (or explicit strategies, level of text), relying on background knowledge (external information not given in the text: information about the language, the topic, the ST or TT culture, or the world, etc.), and occasionally reflecting on the translation process, itself.

All the moves in the TAPS could be categorised in a combined model of the translation process (Gile, 1994) and problem solving in general (Gero and McNeill's model, 1998), belonging either to the comprehension phase (identifying problem, analysing problem) or the reformulation phase (proposing solution, analysing solution). The translation process with intermediate students is not linear (longer translation units transferred in automatic processes), but very dynamic, with frequent and constant switches between ST and TT, in short translation units, proceeding in cycles of analysis, synthesis and revision, and returning to problems unsolved. Mainly lexical items (words, noun phrases) are identified as problematic translation units (with a special focus on unfamiliar words and dictionary work). Less emphasis is given to segmentation of sentences, grammatical analysis, and other macro-strategies. Some students (relying on previous training) use such explicit strategies and can use them consistently. Further analysis of the TAP data collected could quantify the above observations. Thought patterns can be identified in the coded protocols, and thought pattern formulas created that can characterise problem solving strategies and particular test takers.

Findings 1-6 apply to TAPs of both translation tasks, and thus were not found as text specific.

Research question 2 addressed the issue of translation strategies: *how the use of think-aloud research can help to understand what translation strategies are needed to produce acceptable translations at intermediate level.*

First strategies were identified in the scripted TAPs, then grouped into two basic categories: successful and unsuccessful strategies, finally a method was suggested for further research. The list of strategies suggested should be considered as a *tentative, empirically based list* of potentially successful and unsuccessful strategies that can produce acceptable translations in intermediate translation tasks. Further exploration of the data and further research is needed, to make the list more definitive. Such further research should involve further text types and translation topics.

Research question 3 aimed to explore *how language testing research can profit from the use of think-aloud research of translation exam tasks.*

Thought patterns characteristic of particular students can be established by exploring all the TAP data collected. To establish what is typical of a student, one should be able to count and describe all the patterns of moves performed by the student and establish typical patterns afterwards, which is a potential direction for further research of the data collected.

The number and length of major phases in which the translation happens when students cope with translation exam tasks can be identified in TAPs. By comparing the number of phases and the number of moves these phases include across tasks and across students, both the difficulty of translation texts can be compared - if the same students translate different texts-, and individual students can also be profiled for the major structure and macro-strategies they use in their translation process.

Limitations, however, to TAP findings at intermediate level in general, and to the list of successful and unsuccessful strategies in particular, should be acknowledged. First of all, male and female participants tend to verbalise differently, female participants tend to provide a lot more data and more strategies, thus findings will tend to be based on female participants' contribution, whereas in the final analysis of outcome male participants tend to be more successful in translation tasks (Statistical analysis in Chapter 4 also confirmed that). Secondly, any such list of successful and unsuccessful strategies should be considered as a collection of options and suggestions only, as TAP research indicates that thought patterns in problem solving tend to be characteristic of individuals, thus no 'uniform' solutions could be recommended. Thirdly, the present research is only a first step to explore the potential of the method, and an exhaustive listing of all possible strategies used by students in the amount of

Végül a fordítást nyelvvizsgafeladatként vizsgáló tanulmányok áttekintése során megállapítható volt, hogy a teszteléssel és fordítással foglalkozó angol nyelven megjelenő vagy hazai cikkek általában konkrét fordítási vagy közvetítési feladatokkal vagy tesztelési problémákkal foglalkoztak. Pedagógiai fordításra vonatkozó átfogó validálási tanulmányt nem sikerült találni.

4. A pedagógiai fordítás konstruktumának tentatív modellje

A teszt validálással foglalkozó irodalom (1. fejezet) valamint a fordításkutatással foglalkozó irodalom (2. és 3. fejezet) áttekintése során néhány alapvető megállapítás azonosítható, amelyet a szakirodalom elfogad, és amelyekre alapozva a nyelvvizsgáztatásban használt pedagógiai fordítás konstruktumának tentatív modellje felállítható.

A modell Bachman (1991) modelljének adaptációja, amelyben szerepelnek a nyelvvizsgáztatást általában meghatározó fontosabb tényezők. Az adaptált modell alapfeltételezésként bár nem explicit módon, de olyan megállapításokra épít, amelyek a pedagógiai fordítás definíciójának potenciális elemeiként összegezhetők az irodalmi áttekintés alapján: 1) a fordításra egyre inkább úgy tekint a szakirodalom, mint a a nyelven keresztül megvalósuló, *kultúrák közötti kommunikáció egy formájára*, 2) a fordítási kompetencia *összetett* fogalom, amely további részekből áll, 3) a *fordítás iránya* (idegen nyelvről anyanyelvre vagy ellenkező irányba) hatással lehet a fordítási kompetencia meghatározására és a teljesítményre, 4) nincs egyértelmű konszenzus arra vonatkozóan, hogy mi a pedagógiai fordítás *alapvető célja* (nyelvtanítás vagy fordítás tanítás), 5) a *nyelvi párok* (nyelvi távolság) kérdése árnyalhatja a definíciót, 6) létezhet egy feltételezett *nyelvi küszöb* a fordítás tanításában és a fordítási kompetencia tesztelésében, 7) különböző *kutatási módszerek* léteznek a fordítás kutatásában (folyamat-, produktum-, és pedagógiai-alapú megközelítések), 8) az *új kutatási módszerek* várhatóan a fordítás empirikus alapokra épülő definícióját adhatják majd, 9) az egyre mobilisabbá váló európai kontextusban *erősödik a valós nyelvhasználati igény* arra, hogy az interkulturális kommunikáció formáit már a hivatásos fordító szint alatti szinten is tanítsák, 10) az „iskolai fordítás” és a „valódi fordítás” éles szembeállítás helyett alkalmazhatóbbnak bizonyulhat a *fordítási kompetencia mint kontinuum skála* fogalma, ahol az alacsonyabb kompetencia szinttől a hivatásos fordítói szintig különböző szintek különíthetők el.

amellyel definiálni lehet a forrásnyelvi szöveg és a célnyelvi szöveg viszonyát (Kenny, 2001: *a szöveg szintjeire* [szó, mondat, szövegszint] alapozó ekvivalencia vs. *jelentés-alapú* ekvivalencia, Baker, 1992: *szövegszintű* ekvivalencia, Newman, 1994: *funkcionális* ekvivalencia). Az ekvivalencia jellegének meghatározására használt kritérium lehet *nyelvészetten kívüli* (Catford, 1965, 1994) és *nyelvészeti* (Pym, 1992).

A másik oldalon a célnyelvi szövegből kiindulva a folyamat-alapú szemléletmód a fordító személyére koncentrál, mint aki elfogadott *normákat* követ a megfelelő fordítási stratégiák kiválasztásában és a döntéshozatalban (Chesterman, 1993 és Toury, 1995: idézi Bakers, 2001).

A fordítás komplex és átfogó definíciójában elkerülhetetlen, hogy a fordítást folyamatként és az adott folyamat termékeként egyszerre definiáljuk. Bármely fordítás minőségének meghatározásakor mindkét megközelítést csak együttesen alkalmazhatjuk (az ekvivalenciát mint szöveg-alapú kategóriát és a fordítási normákat, mint folyamat-alapú kategóriát).

A fordításértékelés tipikusan elkülönülő szinterei Magyarországon: a) a nyelvtanítás általában, b) a felsőoktatás szakfordítói képzése, és c) a hivatásos fordítói képzések. Ezek az eltérő kontextusok eltérő szempontokból értékelik a fordítást, eltérő fogalmat alkotnak arról, hogy mit mérnek, és ennek megfelelően eltérő értékelési kritériumokat fejlesztenek ki (a konstruktum validitás kérdése a teljesítmény mérésben) (Dróth, 2001).

A *hiba* fogalmának vizsgálata hangsúlyosabban kerül előtérbe az alkalmazott nyelvészeti kutatásokban (idegennyelv elsajátítás), mint a fordításkutatási szakirodalomban. Kivételt képez Pym (1992) felfogása, aki bináris és nem-bináris fordítási hibát különböztet meg. A nyelvtanításban Melis és Albir (2001) azt javasolják, hogy a hibát funkcionista szempontból kellene megközelíteni, és a tanításban *részleges értékelés* módszerét alkalmazni (azaz nem értékelni valamennyi lehetséges tényezőt a fordítás értékelésében).

Egy további lényeges fogalom a *fordítási egység*, ennek meghatározása a folyamat-alapú megközelítésben mint relatív fogalom történik (Lörscher, 1993). Mind a folyamat-alapú kutatás (Lörscher, 1993) mind a produktum-alapú kutatás (Toury, 1986) megfigyelései azt sugallják, hogy pedagógiai szempontból komoly jelentősége van annak, hogy mit tekintünk fordítási egységnek (egy szóból álló lexikai egység vs. kifejezések, mellékmondat vagy mondat), és milyen hosszúak ezek a fordítási egységek. Malmkjaer (2001) megállapítása szerint a mellékmondat tűnik a legáltalánosabban elfogadott fordítási egységnek.

data collected was not possible, because of lack of time. Fourthly, in two translation tasks not all possible translation problems appear, thus the type and nature of the two texts may have determined the scope of strategies explored.

5.3. Summary of product-based research

The research questions focused on the potential of exploiting a concordance programme for empirically supported analysis of translation scripts. The exam tasks researched here were real translation scripts of the three translation tasks already analysed in the statistical and think-aloud chapters.

The hypothesis in *Research question 1*, as to any high correlation exists between the overall statistical properties of the corpus and the overall difficulty of the translation task, had to be rejected as neither the word/sentence ratio nor the type/token ratio seems to be an indicator of overall task difficulty, when compared to data from descriptive statistics.

In *Research question 2*, the potential of frequency counts relating to the variety and types of learner errors and good translations, as well as the possibility of creating a typical learner translation of a sentence, or even of the whole text, were examined. Data analysed proves that an incredible variety exists of both acceptable and unacceptable translations of translation units, providing empirical evidence for the idiosyncratic nature of translation. Pym's (2002) definition of translation seems to be substantiated in the context of intermediate test takers' translation of translation exam tasks: translation potentially means generating an infinite number of target text versions and deciding on the most appropriate one.

Research question 3 focused on the potential of item analysis for translation items. It was found that, based on the *sort by headword* function of a concordance programme and also frequency counts, a particular translation item can be analysed for item difficulty and discrimination. That kind of corpus-based evidence, if routinely provided, could end endless debates between markers as to what really constituted difficulty for the candidates in the given translation task, and to what extent. In addition to its extreme use for test developers and item writers, such analysis, if made public, could also help teachers and candidates to better understand what constitutes real difficulty in translation tasks, and could contribute to demystifying the translation exam task.

Frequency counts can also be the basis for trying out alternative marking schemes (*Research question 4*), either in the pre-testing period or before finalising the marking

scheme, as well as in the development of new tests or in reform periods. By the help of frequency counts it is possible to foresee to what extent certain modifications would change the overall task difficulty or the pass rate, as well as whether they would discriminate against the weaker or the better candidates.

The most exciting area is doing large scale empirically supported research on translation strategies (**Research question 5**). Minimal analysis is already possible based on frequency counts and insight from the researcher. If, however, proper annotation for strategies was added, successful and unsuccessful translation strategies could be explored from a product-based approach, revealing automatized strategies process-based research cannot tap.

6. Final summary and implications

Based on the theory-based and empirical research in the present dissertation, the final answers about the pedagogical use of translation for language proficiency testing are:

a) pedagogic translation can fit the context of communicative language teaching, as both theory-based and empirical research provided evidence for that,

b) translation at intermediate level does not measure any other factor than foreign language proficiency (correlation, multitrait-multimethod matrix, principal component analysis), although with difficult translation tasks it emerges as a second weak factor (principal axis factoring),

c) new research methods (think-aloud and corpus research) in construct validation can contribute to increasing the validity and reliability of the assessment of translation performance,

d) performance in the translation exam tasks does not need strategies the teaching of which could not be placed within the context of teaching foreign language proficiency in general.

Both the theoretical and the methodological outcome of the present research could contribute to the construct validation part of internal validation that language proficiency exam centres are going through at the moment in Hungary, when working on establishing validity for their exams in linking their levels to CEFR levels. They could also contribute to developing translation performance scales in Hungary for the purpose of linking translation to CEFR levels again, thus contributing to an area in language testing (the validation of mediation) that has been under researched in mainstream language testing literature so far.

A fordításkutatásban különböző történelmi szakaszokat különböztethetünk meg (Klaudy, 2004). Mindegyik korszak a maga különböző válaszait adta meg a fordításkutatás alapkérdéseire: a) mit fordítunk valójában?, b) kinek?, és c) milyen kontextusban?

Mindegyik történelmi szakasz a maga sajátos fókuszát helyezte előtérbe (áttekinti Hatim, 2001), a rá jellemző terminológiát használta (elemzi Chesterman, 2005), eltérő kutatási módszereket alkalmazott (RETS, 2001 tanulmányai) és különböző modelleket állított fel (áttekinti Hatim, 2001; elemzi Király, 2005). Amit számos elemző (Pym, 2002a, 2002b, 2003; Király, 2005) felismer és hangsúlyoz az az, hogy folyamatosan változik a fordítással szemben támasztott társadalmi igény, ez a fordításkutatás alapfogalmainak folyamatos újra definiálásához vezet - ez pedig új kutatási módszereket hoz magával, amellyel az új fogalmak mélységeiben feltárhatóak. E történelmi szakaszok a fogalmak, terminusok, definíciók és modellek bőségét eredményezték, amelyek mind gazdagították a fordítás folyamatának komplexitásáról való tudásunkat (ahogy erre Hatim, 2001 és Király, 2005 reflektál), ugyanakkor ez a bőség a fogalmak és terminusok definíciójában terminológiai zűrzavart is eredményezett (lásd Chesterman, 2005; Király, 2005), és nagyban hozzájárult a fordítás definiálhatósága körül kialakult elbizonytalanodáshoz.

Az újabb elméleti megközelítések között három fő irányvonal körvonalazódik: *kompetencia alapú megközelítések* (pl. Campbell, 1998; a PACTE modell, 2003), *minimalista* meghatározások (pl. Pym) és *szocio-kognitív* megközelítések (KER, 2001; Kaiser-Cooke, 2002; és Risku, 2002) - érezhető a szocio-kognitív megközelítés térnyerése.

Ma a fordításkutatásban általánosan elfogadott és hangsúlyos az *interdiszciplináris* megközelítés létjogosultsága (Baker, 2001; Klaudy, 2004), és a fordítás szocio-kognitív felfogásával együtt várhatóan mindjobban előtérbe kerül majd, a fordítás kognitív folyamatait és a fordító gondolkodásában lezajló folyamatokat kutatva (Kaiser-Cooke, 2002; Risku, 2002). Interdiszciplináris kutatással lehetővé válik majd a korábbi spekulatív alapon létrejövő definíciókkal szemben az empirikus adatokra épülő fordítás meghatározás, ahogy azt annak idején Campbell (1998) is sürgette.

3.3. A fordítás értékelésével kapcsolatos megállapítások

A fordítás értékelés központi kategóriái az „ekvivalencia” és a „fordítási normák” - mindkettő sokat vitatott és komplex fogalom.

A nyelvészeti és a forrásnyelvi szöveg alapú megközelítések a fordítás minőségének vizsgálatában úgy tűnik, valamilyen ekvivalencia típusú meghatározásra törekcsenek,

3. Elméleti háttér – összefoglaló

3.1. A tesztvalidálással kapcsolatos megállapítások

A validitás röviden meghatározva olyan empirikus vagy nem-empirikus adatok gyűjtését jelenti, amelyek igazolható módon alátámasztják a vizsgafejlesztők és vizsgáztatók állításait, amelyeket a vizsgák előállítással, lebonyolítással, az eredmények értékelésével és a vizsgaeredmények felhasználhatóságával kapcsolatban tesznek.

A validálással foglalkozó szakirodalomban alapvető hangsúlyt kap Messick (1989) konstruktum validitás meghatározása, aki egy alapvető, mindent átfogó érvényességeként definiálja a fogalmat. Ennek széleskörű elfogadását követte később az a gondolat, hogy nincs csupán *egyetlen* üdvözítő módja annak, hogy megadott céllal a vizsgaeredményekből levont következtetéseket hogyan validáljuk (Alderson és Banerjee, 2002). Ehelyett inkább számos olyan szempont létezik, ami alapján bizonyítékokat gyűjthetünk a validitás igazolására. A fogalom fejlődésének eredményeként ma már a konstruktum validitásra úgy tekintünk, mint egyéb validitás típusokat magába foglaló, átfogó fogalomra (Brown, 2000). A konstruktum validitás fogalmát a nyelvvizsgáztatásban korábban részletesen és mélységeiben Bachman (1991) vizsgálta a nyelvvizsgáztatásról szóló meghatározó művében.

A konstruktum validálás során gyűjtött bizonyítékok lehetnek kvalitatív és kvantitatív jellegűek (Bachman, 1991). Kvantitatív típusba tartoznak: korrelációs jellegű elemzések (korreláció, faktor analízis, MTMM mátrix) és kísérleti eredmények. Kvalitatív típusúak közé tartozik a vizsgateljesítményt eredményező folyamat elemzése (pl. hangos gondolkodás segítségével).

Változatos megközelítéseket találunk arra vonatkozóan, hogy a különböző validálási eljárásokat hogyan lehet egy elméleti keretrendszerbe rendezni (Bachman és Palmer, 1996; Mislevy et al., 2002; Kézikönyv [Manual], 2003; Weir és Shaw, 2005). Weir és Shaw (2005) evidencia-alapú elméleti szocio-kognitív modellje áll legközelebb ahhoz, hogy a vizsgakészítés és értékelés tényleges folyamatát időrendiségében is tükrözze.

3.2. A fordítás konstruktumával kapcsolatos megállapítások

A fordításkutatással foglalkozó irodalom exponenciálisan nőtt az elmúlt néhány évtizedben (mint ahogy azt a St. Jerome Publishing, fordításelméleti kutatással foglalkozó kiadó kategorizálása is bizonyítja: 27 kategóriába sorolja a fordításkutatással foglalkozó szakirodalom bibliográfiáját) - így kutatási célra bőven áll szakirodalom rendelkezésre.

A fordítás mint középfokú nyelvvizsgafeladat validása

magyar nyelvű tézisek

1. A jelen kutatás szükségessége, a fő téma és a disszertáció célkitűzései

Magyarországon az 1980-90-es években vált erősen vitatott kérdéssé, hogy a kétnyelvűséget alkalmazó nyelvvizsgarendszerekben a közvetítőkézség, ezen belül a fordítás mérése mennyire releváns a nyelvvizsgáztatásban. A kérdésre adott egymásnak ellentmondó válaszok mögött a magyarországi nyelvvizsgáztatás színterének átstrukturálási szándéka volt érezhető. A közvetítőkézség mérését, és elsősorban az idegen nyelvről magyarra fordítást, mint vizsgafeladatot ellenzők leginkább a validitás és megbízhatóság kérdéseire koncentrálnak bírálják azt, a bírálatok többsége a konstruktum validitás elméleti és módszertani aspektusait érintette.

A leggyakoribb kétségek a fordítás nyelvvizsgáztatás célú pedagógiai felhasználásával kapcsolatban: a) a fordítás illik-e egyáltalán a kommunikatív nyelvtanítás kontextusába, b) mér-e más is, mint az idegen nyelv tudását, c) lehet-e a fordításban nyújtott teljesítményt érvényes és megbízható módon mérni, és d) a fordítás mint nyelvvizsgafeladat teljesítéséhez olyan stratégiák tanítására is szükség van-e, amelyek a nyelvtudáson kívüli készségek fejlesztését igénylik. Összefoglalóan: kívánatos és módszertanilag védhető-e az anyanyelv használata az idegennyelvi tudás mérésében, különösen az idegen nyelvről magyarra fordítás alkalmazása igazolható-e középfokon.

A pedagógiai célú fordítás konstruktum validálása már csak azért is kívánatosnak tűnt, mert Magyarországon az államilag akkreditált nyelvvizsgarendszerek egy jogszabály előírásainak megfelelően megkezdtek nyelvi szintjeiket a Közös Európai Referenciakeret szintjeihez igazítani. Ennek a folyamatnak viszont elengedhetetlen része a belső és külső validálás. Így a vizsgaközpontok előtt álló kihívás többek között az volt, hogy fordítási vizsgafeladataik validálásának elméleti megalapozottságát bizonyítani tudják, és elsősorban a konstruktum validálás módszertani megoldásait megtalálják.

Jelen kutatás elsődleges célja így az volt, hogy a) megvizsgálja a fordításkutatással foglalkozó szakirodalomban a fordítás konstruktumának főbb aspektusait, b) áttekintse röviden a fordítás értékelésének a konstruktum validitást érintő főbb fogalmait, c) vizsgaadatok statisztikai elemzése segítségével választ keressen a konstruktum validálás releváns elméleti és módszertani kérdéseire, d) új kutatási módszerek (hangos gondolkodás és korpuszelemzés) alkalmazásának lehetőségét vizsgálja a konstruktum validálás területén.

Célkitűzése volt a disszertációnak továbbá az is, hogy nyitottabbá tegye a nyelvtanárokat és nyelvvizsgáztatókat a pedagógiai fordítás főbb elméleti kérdéseinek kutatásával kapcsolatban, hogy elméletileg megalapozottabb, a lényeges kérdéseket érintő vitát folytathassanak a fordításnak és a közvetítőkézségnek a nyelvvizsgáztatásban betöltött szerepéről.

Jelen kutatás várható főbb elméleti eredményei voltak a) a pedagógiai fordítás konstruktum validitás szempontjából meghatározó főbb aspektusainak beazonosítása a fordításkutatással foglalkozó szakirodalomban, b) a pedagógiai fordítás konstruktumának tentatív modellje, c) a fordítás értékelése szempontjából legfontosabb fogalmak vizsgálata, amelyek megkerülhetetlenek a konstruktum validitás szempontjából. Az empirikus kutatásban a fő hangsúly olyan módszerek vizsgálatán volt, amelyek konstruktum validálás céljából használhatóak az alábbi területeken: a) vizsgaeredmények statisztikai elemzése az érvényesség és megbízhatóság szempontjából, b) a hangos gondolkodás módszerének potenciális használata a válaszadási validitás (response validity) kérdésének vizsgálatában, és a c) a korpuszkutatás mint módszer potenciális használhatósága a vizsgaértékelésre vonatkozó validitás (scoring validity) vizsgálatában.

2. A disszertáció felépítése

1. táblázat: A kutatási terv – felhasznált módszerek

Módszer	Jellege	Adat jellege	Elemzési módszer	Eredmény
Szakirodalmi áttekintés	Kvalitatív	Validálási tanulmányok	Áttekintő Elemző olvasás	Validitás típusok (a jelen kutatás szempontjából használható)
Szakirodalmi áttekintés	Kvalitatív	Fordításkutatási tanulmányok	Feltáró jellegű elemző olvasás	A fordítás konstruktumának aspektusai (tentatív modell)
Szakirodalmi áttekintés	Kvalitatív	Fordításkutatási tanulmányok	Feltáró jellegű elemző olvasás	Kulcsfogalmak a fordítás értékelésben
Statisztikai elemzés	Kvantitatív empirikus	Vizsga adatok (éves adatok és egyedi alkalmak)	SPSS elemzés	Vizsgaadatok statisztikai értékelésének módszertana a konstruktum validáláshoz
Introspektív (Hangos gondolkodás)	Kvalitatív empirikus	Vizsgálók hangos gondolkodását rögzítő jegyzőkönyvek	Jegyzőkönyv elemzés – kísérleti	Fordítási stratégiák folyamat-alapú elemzése – analitikus keretrendszer, sikeres és sikertelen fordítási stratégiák: tentatív lista
Korpuszkutatás	Kvalitatív -kvantitatív aspektussal empirikus	Vizsgálók vizsgafordításai	Elemzés Konkordancia program segítségével – kísérleti	Potenciális módszerek az értékelés validitásának meghatározásához – fordítások produktum-alapú elemzése

1. ábra: A kutatás szerkezete

1. fejezet: Teszt validálás – elméleti háttér

