

Testen der Sprachkenntnisse  
bei Schülern mit A1- bzw. A2-Deutsch

Pro gradu - Arbeit  
Germanistisches Institut  
Universität Jyväskylä  
Januar 2002

Leena Rantala

# JYVÄSKYLÄN YLIOPISTO

Tiedekunta  
Humanistinen tiedekunta

Laitos  
Saksan kielen laitos

Tekijä  
Leena Rantala

Työn nimi  
Testen der Sprachkenntnisse bei Schülern mit A1- bzw. A2-Deutsch

Oppiaine  
Germaaninen filologia

Työn laji  
Pro gradu

Aika  
Tammikuu 2002

Sivumäärä  
61 sivua ja liitteet

## Tiivistelmä

Saksa on suosituin valinnainen A2-kieli ala-asteella. Saksaa ensimmäisenä vieraana kielenä, A1-kielenä, lukevien oppilaiden määrä on puolestaan pieni. Kouluissa pohditaankin, voiko A1- ja A2-kielen lukijat yhdistää kielitaitonsa puolesta samaan ryhmään yläasteelle siirryttäessä. Tämän tutkimuksen tarkoituksena oli selvittää, kuinka paljon saksaa A1-kielenä ja saksaa A2-kielenä lukevien oppilaiden kielitaidoissa on eroja kuudennen luokan keväällä ala-asteen päättyessä.

Teoreettista tietoa kielitaidosta tarvitaan, jotta kielitaitoa voidaan testata monipuolisesti ja luotettavasti. Tutkimuksessa perehdytäänkin ensin kielitaidon olemukseen ja yksilöllisiin eroihin kielenoppimiskyvyssä. Myöhemmin pohditaan arvioinnin erilaisia tehtäviä sekä käsitellään arvioinnille yleisesti asetettuja vaatimuksia – reliabiliteettia ja validiteettia. Työssä vertaillaan myös erilaisia kielitaidon arvioinnissa käytettäviä tehtävätyyppejä.

A1- ja A2-saksaa lukevien oppilaiden tasoeroja selvitettiin Sukol-Palvelu Oy:n ja Suomen saksanopettajat ry:n tuottaman kirjallisen kokeen avulla. Tutkimuksen otos muodostui yhdeksästä ryhmästä, joissa oli yhteensä 119 oppilasta. Näistä A1-saksan lukijoita oli 48 ja A2-saksan lukijoita 71. Opettajille suunnatun kyselylomakkeen avulla kerättiin lisäksi tietoa taustatekijöistä ja kysyttiin opettajien mielipiteitä kokeesta sekä A1- ja A2-saksan lukijoiden kielitaidon eroista. Saatua tietoa käytettiin hyväksi oppilaiden koetuloksia arvioitaessa.

Tutkimus osoitti, että A1- ja A2-lukijoiden kielitaidoissa ei voi tämän tutkimuksen perusteella todeta olevan merkittäviä tasoeroja. Huolimatta siitä, että A1-saksan lukijat ovat opiskelleet kieltä huomattavasti enemmän kuin A2-lukijat, menestyivät he keskimäärin heikommin kokeessa. Kaikkein ryhmien välillä oli kuitenkin huomattavia eroja. Näihin eroihin pyrittiin löytämään selityksiä eri taustatekijöistä, mutta mitään johdonmukaisia eroja ei löydetty. Tämän perusteella todettiin, että erot oppilaiden ja ryhmien koemenestyksessä johtuvat todennäköisesti suurelta osin yksilöiden välisistä eroista kielenoppimiskyvyssä. Myös opettajan taidoilla ja persoonallisuudella pääteltiin olevan merkittävä vaikutus kielen oppimiseen.

Avainsanat: A1-saksa, A2-saksa, kielitaidon arviointi, kielitaito, validiteetti, reliabiliteetti

# Inhaltsverzeichnis

<b>1. Einleitung</b> .....	1
<b>2. Sprachkompetenz und Sprachlernfähigkeit</b> .....	3
<b>2.1 Sprachlernfähigkeit</b> .....	4
<b>2.2 Theorien über Sprachkompetenz</b> .....	5
2.2.1 <i>Das strukturalistische Modell</i> .....	6
2.2.2 <i>Das Modell der kommunikativen Kompetenz</i> .....	6
<b>3. Funktionen der Leistungsmessung</b> .....	10
<b>3.1 Die pädagogische Funktion</b> .....	10
<b>3.2 Die Beurteilung</b> .....	11
3.2.1 <i>Normorientierte und kriteriumsorientierte Leistungsmessung</i> .....	12
3.2.2 <i>Formative und summative Leistungsmessung</i> .....	13
<b>4. Gütekriterien für Tests</b> .....	14
<b>4.1 Reliabilität</b> .....	15
<b>4.2 Validität</b> .....	16
<b>4.3 Praktikabilität</b> .....	17
<b>5. Aufgabentypen beim Testen der Sprachkenntnisse</b> ....	18
<b>5.1 Testen der schriftlichen Produktionsfähigkeit</b> .....	20
5.1.1 <i>Traditionelle, freie Aufgaben</i> .....	20
5.1.2 <i>Gesteuerte Aufgaben</i> .....	22
<b>5.2 Testen der rezeptiven Sprachfähigkeit</b> .....	22
5.2.1 <i>Leseverständnis</i> .....	23
5.2.2 <i>Hörverständnis</i> .....	24
<b>5.3 Testen der mündlichen Sprachkompetenz</b> .....	25

<b>6. Methoden</b> .....	26
<b>6.1 A1- und A2-Sprachen</b> .....	27
<b>6.2 Die Untersuchungsmittel</b> .....	28
<b>6.3 Forschungsgegenstand</b> .....	29
<b>6.4 Verlauf der Untersuchung</b> .....	30
<b>7. Ergebnisse</b> .....	31
<b>7.1 Meinungen der Lehrer über die Testaufgaben</b> .....	31
7.1.1 <i>Testen der produktiven Sprachfähigkeit</i> .....	32
7.1.2 <i>Testen der rezeptiven Sprachfähigkeit</i> .....	33
<b>7.2 Hintergrundfaktoren</b> .....	34
7.2.1 <i>A1- bzw. A2-Deutsch</i> .....	35
7.2.2 <i>Größe der Schule</i> .....	37
7.2.3 <i>Größe der Lerngruppe</i> .....	37
7.2.4 <i>Die demografischen Faktoren – die Lage und Einwohnerzahl der Gemeinde</i> .....	38
7.2.5 <i>Geschlecht des Schülers</i> .....	39
7.2.6 <i>Qualifikation der Lehrer</i> .....	40
7.2.7 <i>Zahl der Unterrichtsstunden</i> .....	41
<b>8. Interpretation der Testergebnisse</b> .....	43
<b>8.1 Umfang der gelernten Sprache und Zahl der Unterrichtsstunden</b> .....	43
<b>8.2 Größe der Schule und der Lerngruppe</b> .....	45
<b>8.3 Bedeutung der demografischen Faktoren</b> .....	46
<b>8.4 Geschlecht des Schülers</b> .....	47
<b>8.5 Rolle des Lehrers</b> .....	48

<b>9. Validität der Untersuchung</b> .....	49
<b>9.1 Reliabilität der Untersuchungsprüfung</b> .....	50
9.1.1 <i>Testen der schriftlichen Produktionsfähigkeit</i> .....	50
9.1.2 <i>Testen der rezeptiven Sprachfähigkeit</i> .....	52
<b>9.2 Fragebogen als Informationsquelle</b> .....	53
<b>9.3 Forschungsgegenstand</b> .....	54
<b>10. Schlussbetrachtung</b> .....	55
<b>Literaturverzeichnis</b> .....	58
ANHANG 1. Untersuchungsprüfung. ....	62
ANHANG 2. Der erste Brief. ....	72
ANHANG 3. Der Brief von Sukol ry.....	73
ANHANG 4. Fragebogen. ....	74
ANHANG 5. Formular für die Punkte. ....	76
ANHANG 6. Der zweite Brief. ....	77

## 1. Einleitung

In der vorliegenden Arbeit werden die Grundlagen der Forschung der Leistungsmessung vorgestellt und Leistungsmessung einer Fremdsprache behandelt. Unter Leistungsmessung werden hier solche Verfahren verstanden, mit denen nach Beweisen von den Sprachkenntnissen eines Lerners für einen bestimmten Zweck gestrebt wird.

Das Ziel dieser Untersuchung ist herauszufinden, ob die Sprachkenntnisse bei Schülern mit A1- und A2-Deutsch sich am Ende der Unterstufe unterscheiden. Der Ausgangspunkt ist das pragmatische Problem an den Schulen: können die Schüler mit A1- und A2-Deutsch am Anfang der Oberstufe in die selbe Gruppe platziert werden oder nicht. Die Frage ist, ob die Deutschkenntnisse dieser Gruppen auf demselben Niveau schon am Ende der Unterstufe oder erst am Ende der Oberstufe sind. In einigen Gemeinden werden die A1- und A2-Gruppen in der Oberstufe vereinigt, in den meisten Gemeinden nicht.

Der Zweck dieser Untersuchung ist zu testen, ob die Schüler mit A1-Sprache besseren Erfolg in einem Test hatten als die Schüler, die Deutsch als A2-Sprache gelernt haben. Zusätzlich wird überprüft, ob unterschiedliche demografische und andere Hintergrundfaktoren die Testperformanz beeinflussen.

Die Hypothese ist, dass die Sprachkenntnisse bei Schülern mit A1-Deutsch besser als bei Schülern mit A2-Deutsch sind, weil die Schüler, die Deutsch als A1-Sprache lernen, bis Ende der Unterstufe mehr Unterrichtsstunden gehabt haben. Es wird auch vermutet, dass die Lernergebnisse besser in kleinen Lerngruppen und kleinen Schulen sind als in größeren.

Das Wissen über Sprachkompetenz ist bei der Leistungsmessung notwendig. Beim Ausarbeiten und Interpretieren der Tests muss klar sein, was eigentlich gemessen wird. Durch Bewertung werden die Leistungen der Schüler beurteilt und die Lernergebnisse getestet. Theoretisches Wissen ist Voraussetzung dafür, dass die durch Testen erbrachte Information zuverlässig und die anhand der Leistungsmessung gegebenen Noten gerecht sind.

Als theoretische Grundlage werden in dieser Arbeit die bekanntesten Auffassungen über Sprachkompetenz vorgestellt, die in den westlichen Ländern die Leistungsmessung am meisten beeinflusst haben. In diesem Zusammenhang wird auch die Sprachlernfähigkeit besprochen. Danach werden die Funktionen der Leistungsmessung behandelt: die Optimierung des Lehrens und Lernens, die Beurteilung im Rahmen eines Lehrprogramms und der Nutzen für die sprachwissenschaftliche Forschung. Die Gütekriterien für Tests – Reliabilität, Validität und Praktikabilität – werden auch näher besprochen und die verschiedenen Typen von Testaufgaben vorgestellt.

Im empirischen Teil werden die Methoden und der Verlauf der Untersuchung erläutert und die Ergebnisse vorgestellt. Der Test, der in dieser Untersuchung verwendet wurde, wird auch näher besprochen. Die Testergebnisse der Schüler werden verglichen und der Einfluss der Hintergrundfaktoren wird geschätzt. Es wird versucht, die Ergebnisse zu interpretieren und Gründe für die Unterschiede zwischen den Schülern finden. Die Validität der Untersuchung – die Reliabilität des Testes, das Untersuchungsverfahren und die Interpretation der Ergebnisse – wird auch besprochen.

## 2. Sprachkompetenz und Sprachlernfähigkeit

Um die sprachlichen Fähigkeiten testen zu können, muss das Wesen der Sprachkompetenz für den Verfasser des Testes bekannt sein. Die Theorien über Sprach- und Kommunikationsfähigkeit versuchen, die Fähigkeiten aus mehreren Gesichtswinkeln zu beschreiben. Kognitive Theorien betrachten u.a. die Struktur der Sprachkompetenz und kognitive Prozesse bei der Spracherlernung und dem Sprachgebrauch. Mit kognitiven Prozessen werden hier solche Prozesse verstanden, die mit dem Denken, dem Gedächtnis und der Behandlung der Information zusammenhängen. Soziale Theorien dagegen untersuchen, wie Menschen die Sprache beim Kommunizieren anwenden und ob die Sprache passend für die jeweilige Situation ist und was mit bestimmten sprachlichen Äußerungen gemeint wird. Die Auffassungen, die die Leistungsmessung am meisten beeinflusst haben, sind Mischungen von kognitiven und sozialen Auffassungen. (Huhta & Takala 1999, 183.)

Theoretisches Wissen ist nötig, um die Sprachkenntnisse vielfältig bewerten zu können. Das Wesen der Sprachkompetenz muss für den Tester bekannt sein, damit die Testergebnisse auf eine relevante und zweckmäßige Weise interpretiert werden können. In der Leistungsmessung geht es um einen Versuch, einen Teil der abstrakten Sprachkompetenz in eine konkrete Form zu verwandeln. Was und wie ein Test die Sprachkompetenz testet, hängt davon ab, was der Verfasser wesentlich findet und wie er das Wesen der Sprachkompetenz auffasst. (Huhta & Takala 1999, 181 – 182.)

Die linguistischen Theorien beschreiben die Sprachkompetenz als eine multidimensionale Fähigkeit. Dies hat selbstverständlich Einfluss auf die Leistungsmessung. Gemeinsam für die neuesten Modelle ist das Hervorheben der soziolinguistischen Kompetenz. Die Fertigkeit, die

Sprache je nach dem Kontext passend anzuwenden, wird als ein wesentlicher Teil der Sprachkompetenz betrachtet. (Huhta & Takala 1999, 181 – 182; 188 – 189.)

Die Theorien helfen, die großen Linien zu konzipieren, wenn entschieden wird, welche Teilbereiche der Sprachkompetenz im Fremdsprachenunterricht am wichtigsten sind (Huhta & Takala 1999, 181 – 182). Wenn z.B. gerade die soziolinguistische Fertigkeit als eine wichtige Fähigkeit angesehen wird, kann der Schwerpunkt des Unterrichts in diesem Teilbereich liegen.

Kognitive Stile und andere individuelle Fähigkeiten beeinflussen Entwicklung der Sprachkompetenz eines Individuums. Diese Sprachlernfähigkeiten werden näher im Folgenden besprochen.

## **2.1 Sprachlernfähigkeit**

Der eine lernt fremde Sprachen schnell und leicht, dem anderen ist es viel schwieriger. Wenn die Fähigkeit der Fremdsprachenerlernung (engl. *language aptitude*) bewertet wird, geht es nicht um Testen der Sprachkompetenz, sondern um die Fähigkeiten, die dem Lerner beim Fremdsprachenerwerb helfen. (Huhta & Takala 1999, 193.) Unterschiede zwischen den Lernern und wie diese Unterschiede den Fremdsprachenerwerb beeinflussen, sind ein viel untersuchtes Gebiet (siehe Jaakkola 1997, 68).

In einer gesteuerten Fremdsprachenlernsituation ist der Unterricht vielleicht der wichtigste, nicht aber der einzige Faktor, der den Fremdsprachenerwerb beeinflusst (Jaakkola 1997, 68). Der Erwerb einer fremden Sprache wird auch durch viele andere Faktoren beeinflusst. Sowohl die Muttersprache als auch die Zielsprache sind entscheidend.

Andere Faktoren sind u.a. das Alter, die Bildung und die soziale und kulturelle Herkunft des Lerners. Auch das Lernvermögen und die Lernbereitschaft spielen eine Rolle. (Merten 1997, 68.) Zu diesen können auch die Verhaltensweisen, die Persönlichkeit, die kognitiven Stile und die Lernstrategien des Lerners hinzugefügt werden (Larsen-Freeman & Long 1991, 175 ff). Forscher haben festgestellt, dass Erwachsene im gesteuerten Unterricht die Sprache viel schneller als Kinder lernen. Kinder aber kommen besser im ungesteuerten Lernprozess zurecht und können ein höheres Niveau in Hinsicht der Zweisprachigkeit erreichen. (Siehe Larsen-Freeman & Long 1991, 154ff.)

Das Alter, die sprachliche Kompetenz und die kognitiven Stile sind wahrscheinlich die wichtigsten Faktoren, die zumindest den Grammatikunterricht und die Wirkung des Grammatikunterrichts beeinflussen. Im Unterricht kann man auch die Einstellung der Lerner gegenüber der Grammatik beeinflussen. (Jaakkola 1997, 68.) Es ist zu vermuten, dass dieses sich auch im allgemeinen auf den ganzen Sprachlernprozess bezieht und auch in anderen Bereichen des Sprachunterrichts als nur Grammatik stimmt. Laut dieser Auffassung kann der Lehrer im Unterricht zu einer positiven Einstellung der Lerner gegenüber der Spracherlernung beitragen.

## **2.2 Theorien über Sprachkompetenz**

Im Folgenden wird die Sprachkompetenz mit theoretischen Begriffen erklärt. Zuerst wird eine traditionelle Auffassung der strukturalistischen Linguistik vorgestellt, weil sie laut Huhta und Takala (1999, 183) immer noch üblich ist. Diese Auffassung wird kurz im Kapitel 2.2.1 behandelt. Das Modell von Bachman wird dagegen näher im Kapitel 2.2.2 besprochen.

Die Theorie über kommunikative Kompetenz (engl. *communicative language ability*) von Bachman ist eine allgemein anerkannte Theorie, die auf einer der bekanntesten Theorien über Sprachkompetenz, nämlich der von Canale und Swain, basiert. Das Modell von Canale und Swain konzentriert sich auf die Struktur der Sprachkompetenz, aber Bachman nimmt auch Rücksicht darauf, wie die Sprachkompetenz als Performanz sich verwirklicht. (Huhta & Takala 1999, 184 – 186.) Das Modell von Bachman gilt hier als ein Beispiel für die Vielfältigkeit der Sprachkompetenz. Es ist nicht die einzige Möglichkeit, die Kompetenz zu beschreiben.

### *2.2.1 Das strukturalistische Modell*

Die traditionelle Auffassung über Sprachkompetenz stammt aus der strukturalistischen Linguistik der 60er Jahre. Die Strukturalisten teilten die Sprache in einzelne Elemente – Aussprache, Grammatik, Lexik und kulturelle Bedeutungen – ein. Diese Einteilung ist immer noch üblich. (Huhta & Takala 1999, 183.) Laut der strukturalistischen Auffassung können die vier Elemente getrennt gelernt und gemessen werden, obwohl sie nie selbständig in der Sprache vorkommen, sondern mit den Sprachfertigkeiten – Sprechen, Hören, Lesen, Schreiben und eventuell auch der Fähigkeit zu übersetzen – verknüpft sind (Lado 1961, 25 – 26).

### *2.2.2 Das Modell der kommunikativen Kompetenz*

Am Anfang der 80er Jahre fassten Canale und Swain die Theorien über kommunikative Kompetenz zusammen und sind zu einem Modell gekommen, das vermutlich eine der bekanntesten Auffassungen über Sprachkompetenz in der angewandten Linguistik darstellt (Huhta &

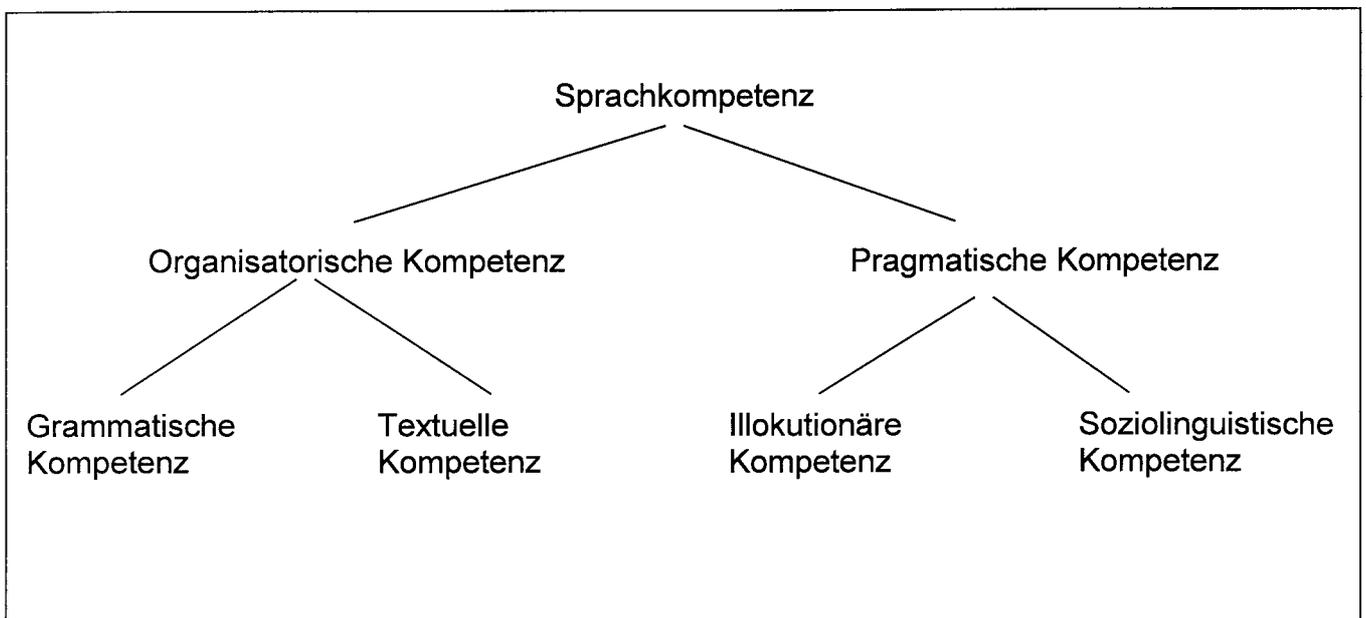
Takala 1999, 184). Die Theorie enthält drei Hauptkategorien: grammatische Kompetenz (engl. *grammatical competence*), soziolinguistische Kompetenz (engl. *sociolinguistic competence*) und strategische Kompetenz (engl. *strategic competence*). Grammatische Kompetenz beinhaltet Kenntnis über die lexikalischen Einheiten und Regeln der Sprache. Soziolinguistische Kompetenz enthält das Wissen über die sozialen Regeln. Die strategische Kompetenz besteht aus verbalen und nicht-verbalen Strategien in der sprachlichen Performanz. (Canale & Swain 1980, 28 – 30.)

Weil die Theorie die Sprachkompetenz auf einem allgemeinen Niveau beschreibt, ist ein Modell entwickelt worden, die besser zu der Leistungsmessung passt. Heute wird in der Leistungsmessung am meisten die Theorie von Bachman und Palmer herangezogen. Das Modell basiert auf dem von Canale und Swain, ist aber komplexer und detaillierter. (Huhta & Takala 1999, 185.) Bachman und Palmer haben die Theorie zusammen verfasst, aber später hat Bachman sie weiterentwickelt. Bachman versucht nicht nur die Struktur der Sprache darzustellen, sondern nimmt Rücksicht darauf, dass die sprachliche Kompetenz in einer echten Anwendungssituation als Performanz verwirklicht wird (Huhta & Takala 1999, 185 – 186).

Sprachkompetenz (engl. *language competence*) umfasst nur die Teile des Wissens, die typisch für die Anwendung der Sprache sind. Der Sprecher braucht aber auch andere Fertigkeiten, um die Sprache in einer Kommunikationssituation zu verwenden. Strategische Kompetenz (engl. *strategic competence*) ist eine wichtige Komponente in Sprachanwendung. (Bachman 1990, 84 – 85.) Die strategische Kompetenz im Modell von Bachman kann als metakognitive Strategien verstanden werden. Metakognitive Strategien sind allgemeine kognitive Prozesse, die der Sprecher verwendet, um die sprachliche Kompetenz in Anwendung der Sprache zu verwandeln. (Huhta & Takala 1999, 186 – 188.)

Die Komponenten der Sprachkompetenz sind in der Sprachanwendung keine selbständigen Elemente, sondern interagieren ständig miteinander und mit dem Kontext. Bachman betont, dass das Diagramm eine visuelle Metapher ist, kein theoretisches Modell. (Bachman 1990, 86.)

GRAFIK 1. Komponenten der Sprachkompetenz (Bachman 1990, 87).



Die Sprachkompetenz beinhaltet zwei grosse Teilgebiete: die organisatorische Kompetenz (engl. *organizational competence*) und die pragmatische Kompetenz (engl. *pragmatic competence*). Der erste Teil der organisatorischen Kompetenz, die grammatische Kompetenz (eng. *grammatical competence*), umfasst Morphologie, Syntax, Lexik, Phonologie und Orthographie. Mit Hilfe der grammatischen Kompetenz drückt der Sprecher spezifische Bedeutungen aus, wählt die richtigen Wörter für seinen Ausdruck und formuliert die einzelnen Sätze. Wie diese Wörter und Sätze miteinander so verknüpft werden, dass sie Texte und Gespräche bilden, heißt textuelle Kompetenz (eng. *textual*

*competence*). Die textuelle Kompetenz bedeutet, dass der Sprecher die Konventionen der Satzverknüpfung kennt. (Bachman 1990, 87 – 88.)

Der andere Teil der Sprachkompetenz ist pragmatische Kompetenz. Die pragmatische Kompetenz hilft dem Sprecher, die Regeln der Sprachanwendung und Interpretation zu beherrschen. Die illokutionäre Kompetenz (eng. *illocutionary competence*) gibt dem Sprecher die Fähigkeit, die illokutionäre Kraft der Sprache zu interpretieren und selbst solche sprachliche Elemente auszudrücken, die eine bestimmte illokutionäre Rolle haben. (Bachman 1990, 89 – 92.)

Die soziolinguistische Kompetenz (engl. *sociolinguistic competence*) ist der andere Teil der pragmatischen Kompetenz. Die Fähigkeit, eine solche Sprache anzuwenden, die zur jeweiligen Situation passt, wird soziolinguistische Kompetenz genannt. Erstens bedeutet das die Fähigkeit, die verschiedenen Varianten der Sprache – soziale oder geographische – zu verstehen und zu verwenden. Der zweite Aspekt ist die Fähigkeit, unterschiedliche Register zu beherrschen. Die soziolinguistische Kompetenz bedeutet auch, dass der Sprecher die Sprache nicht nur grammatisch korrekt, sondern auch idiomatisch und natürlich anwendet. Das ist der dritte Aspekt. Der vierte Teilbereich ist laut Bachman das Vermögen, die kulturspezifischen Anweisungen und Metaphern interpretieren zu können. (Bachman 1990, 94 – 98.)

Die Sprachkompetenz kann also als ein sehr multidimensionales Phänomen angesehen werden. Was Bachman (1990) unter kommunikativer Sprachkompetenz versteht, ist eine Fähigkeit, sprachliches Wissen anzuwenden unter Voraussetzungen der jeweiligen Sprachanwendungssituation. Der Zweck ist, Bedeutungen selbst zu schaffen und Bedeutungen anderer zu interpretieren (Huhta & Takala 1999, 186).

### **3. Funktionen der Leistungsmessung**

Beim Ausarbeiten und Anwenden eines Tests müssen immer praktische Fragen gelöst werden. Sind die Tests gültig für den jeweiligen Gebrauch? Wie zuverlässig sollen die Tests sein? Wie können die Ergebnisse der Tests interpretiert werden? Um diese Fragen zu beantworten, müssen die spezifischen Anwendungsweisen, für wen der Test geplant ist, geklärt werden: wie die Ergebnisse interpretiert und verwendet werden sollen, unter welchen Umständen der Test geschrieben wird und was ist die Funktion des Tests. (Bachman 1990, 1.)

Dieses Kapitel befasst sich mit den Funktionen der Leistungsmessung. Es werden die pädagogische Funktion, die beurteilende Funktion und Wechselbeziehung zwischen Leistungsmessung und linguistischer Forschung besprochen.

#### **3.1 Die pädagogische Funktion**

Die wichtigste Funktion der Leistungsmessung ist die pädagogische (Doyé 1992, 5). Leistungsmessung gibt dem Lehrer Information über den Erfolg und die Effektivität seines Unterrichts. Durch die Leistungsmessung im Fremdsprachenunterricht bekommen sowohl der Lehrer wie der Lerner Information über erfolgte Lernprozesse und noch bestehende Schwierigkeiten, Stärken und Schwächen des Lerners. (Vollmer 1995, 273.) Leistungsmessung hilft dem Lehrer, die künftigen Schritte besser zu planen (Doyé 1995, 277).

Auswertung der Ergebnisse ermöglicht also die Optimierung des Lehrens und Lernens. Dadurch wissen sowohl der Lehrer als auch der

Lerner, welche Lernziele schon erreicht worden sind und welche nicht. Mit Hilfe dieses Aufschlusses, können die Kräfte im Unterricht auf das Schwierigste gerichtet werden, anstatt das schon Gelehrte zu wiederholen.

### **3.2 Die Beurteilung**

Zweitens dient die Leistungsmessung in unserem Bildungssystem oft auch als Grundlage für die Beurteilung der Schüler im Vergleich zu anderen mit Konsequenzen für die Notengebung. Diese bewertende Funktion steht häufig im Vordergrund. (Vollmer 1995, 273.) Doyé (1992, 5) findet diese Funktion zwar wichtig, aber betont, dass sie nicht überbewertet sein darf und die pädagogische Funktion nicht vernachlässigt werden soll.

Eine zu punktuelle Leistungsmessung wird kritisiert, weil die jeweilige Performanz eines Lerner je nach dem Zeitpunkt der Leistungsmessung, nach der Aufgabenstellung, dem Testformat und der Zusammensetzung der Lerngruppe stark variieren kann. Eine zu viele einzelne Punkte betreffende Leistungsmessung ist zu vermeiden – über einen Lerner soll möglichst viel Information gesammelt werden. (Vollmer 1995, 273.) Mit punktueller Leistungsmessung sind hier solche Testverfahren gemeint, in denen nur ein Teil der Sprachkompetenz – wie z.B. Schreibvermögen oder einzelne Wörter – getestet werden. Mit dieser Feststellung verweist Vollmer auf Reliabilität und Validität beim Testverfahren, die später in dieser Arbeit behandelt werden.

Es ist zu bemerken, dass bei der Leistungsmessung nicht alle Lernziele gemessen werden. Sprachbeherrschung ist nämlich meistens nicht das einzige Ziel des Sprachunterrichts. Die anderen Ziele sind z.B. ein positives Verhalten gegenüber der jeweiligen Zielsprache, deren

Sprechern und der fremden Kultur. Wie diese Ziele erreicht werden, ist schwierig festzustellen. Die Notengebung und Zeugnisse beschreiben also nur teilweise, wie die Lernziele erreicht worden sind. (Huhta & Takala 1999, 188 – 189.)

Die Benotung kann sich nach unterschiedlichen Verweisstandards richten (Vollmer 1995, 273). Man unterscheidet sowohl normorientierte und kriteriumsorientierte Leistungsmessung als auch formative und summative Bewertung.

### *3.2.1 Normorientierte und kriteriumsorientierte Leistungsmessung*

Bei normorientierter Bewertung gilt als Bezugspunkt das durchschnittliche Niveau einer Gruppe. Die Lerner werden dabei in bezug auf ihre Testleistungen miteinander verglichen und in eine Rangfolge geordnet. (Vollmer 1995, 273.) Weil Vergleich als Grundlage für die Ergebnisse und Benotung des Testes gilt, sind die Bemessungsgrundlage der Noten eines Lerners nicht seine Leistungen, sondern die Leistungen von anderen in der Gruppe. Besonders in einer kleinen Bezugsgruppe ist das Risiko einer Verzerrung der Ergebnisse groß. Die Leistungen eines durchschnittlichen Lerners kommen einem in einem ganz anderen Licht vor, wenn er in eine Gruppe platziert ist, in der die meisten entweder viel besser oder viel ungeschickter sind als er. Die normorientierte Bewertung ist also nur zuverlässig, wenn die Bezugsgruppe groß genug ist. (Huhta & Takala 1999, 219.)

Bei kriteriumsorientierter Leistungsmessung dagegen wird auf die gesetzten Lehrziele Bezug genommen und gefragt, ob und wie weit ein bestimmtes Lehrziel erreicht worden ist (Vollmer 1995, 273). Die Leistungen der anderen in der Gruppe spielen dabei keine Rolle.

In der Praxis vermischen sich die normorientierte und kriteriumsorientierte Leistungsmessung leicht. Daneben gibt es noch die auf eine individuelle Norm bezogene Leistungsmessung, wobei die Leistungen eines Schülers mit seinen vorausgegangenen Leistungen verglichen werden. (Vollmer 1995, 273.)

### 3.2.2 *Formative und summative Leistungsmessung*

Wenn die Leistungsmessung an die möglichst klar definierten Lehrziele eines bestimmten Lehrprogramms gebunden ist, spricht man von formativer Evaluation. Die Lehrziele können z.B. einzelne Teilgebiete der sprachlichen Kompetenz sein (z.B. der Schüler soll den Dativ beherrschen) oder kommunikative Fähigkeiten (z.B. der Schüler kann auf der Bank seine Angelegenheiten erledigen). Wenn dagegen am Ende eines Lehrgangs das sprachliche Niveau getestet wird und das Ziel eine Rangreihenbildung ist, spricht man von summativer Evaluation. (Vollmer 1995, 274.)

Die formative Bewertung ist wahrscheinlich die üblichste Form der Leistungsmessung. Formative Leistungsmessung stützt den Unterricht. Das Ziel ist, Information darüber zu bekommen, wie der Lerner das Gelehrte beherrscht und wo es noch Schwächen gibt. Dies geschieht in der Regel durch eine dauernde unformale Observierung. Dadurch bekommt der Lerner Feedback über sein Lernen. (Huhta & Takala 1999, 197 – 198.)

Der Zweck der summativen Leistungsmessung ist Notengebung meistens aufgrund schriftlicher Tests. Bei summativer Bewertung werden die sprachlichen Leistungen des Schülers benotet. (Huhta & Takala 1999, 199.)

Landesweite Prüfungen sind oft summative Tests. Wenn Schüler und Lehrer Feedback von den Testleistungen bekommen, ist der Test auch formativ. Die finnischen Abiturprüfungen sind ein gutes Beispiel. Die Prüfungen in sich sind summativ. Die Ergebnisse könnten aber guten Aufschluss über die Stärken und Schwächen des Unterrichts und der Kenntnisse der Schüler geben. Diese Information könnte bei der Lehrplanung verwendet werden. Wenn man solches Feedback aufgrund der Ergebnisse gäbe, könnte die Abiturprüfung auch als ein formatives Testverfahren betrachtet werden. (Siehe Huhta & Takala 1999, 198.)

#### **4. Gütekriterien für Tests**

Die wichtigsten Eigenschaften eines guten Tests sind Reliabilität, Validität und Praktikabilität (Huhta & Takala 1999, 211). Das Kriterium der Reliabilität (Zuverlässigkeit) bezieht sich auf den Grad der Genauigkeit des Messvorgangs und der Testergebnisse. Die Validität (Gültigkeit) des Tests heisst, dass ein Test gerade das misst, was er messen soll. (Vollmer 1995, 274.) Wenn die Ergebnisse eines Tests nicht reliabel und valid sind, können sie nicht als Indikator der Kompetenz eines Individuums interpretiert werden (Bachman 1990, 24). Praktikabilität bezieht sich auf das Praktische – die Ökonomie und Effektivität – im Testdesign und Testverfahren (Weir 1990, 1).

Was beim Testen der Sprachkenntnisse vorzugsweise bedacht werden soll, ist der Zweck, für den ein Test beabsichtigt ist. Je grössere Einflüsse die Bewertung auf die Individuen und Lehrprogramme hat, desto grössere Sicherheit muss es über die Reliabilität und Validität der Testpunkte geben. (Bachman 1990, 78.)

Sprachkompetenz sollte auf eine solche Weise definiert werden, dass die verwendeten Methoden zweifellos diese Fertigkeiten messen, die typisch und natürlich für die Sprachanwendung auch außerhalb einer Testsituation sind (Bachman 1990, 9). Dies bedeutet, dass auch Authentizität beim Testen eine wichtige Dimension ist.

Im Folgenden werden die primären Gütekriterien beim Testverfahren behandelt.

#### **4.1 Reliabilität**

Wenn die Kenntnisse eines Lerners mehrere Male getestet werden und das Ergebnis immer dasselbe ist, kann das Testverfahren als reliabel betrachtet werden (Huhta & Takala 1999, 211 – 212). Der Grad der Reliabilität misst, inwiefern die Performanz eines Lerners von einem Messungsfehler oder von anderen Faktoren als der Sprachkompetenz, die wir messen wollen, abhängt. Forderung nach Reliabilität bedeutet, dass die Effekte dieser Faktoren auf die Testergebnisse minimiert werden. (Bachman 1990, 160 – 161.)

Es gibt viele Faktoren, die sich auf die Performanz eines Individuums im Laufe eines Testvorgangs auswirken können. Die Testperformanz kann u.a. durch Testumstände, Erschöpfung oder Gespanntheit beeinflusst werden. Sie können die Testperformanz des Individuums beeinflussen, sind aber meistens nicht mit der Sprachkompetenz verknüpft. Die verschiedenen Fehlerquellen sollen identifiziert werden und ihr Effekt soll eingeschätzt und minimiert werden. (Bachman 1990, 24; 160.)

Einfluss der Fehlerquellen kann öfters reduziert werden so dass die Reliabilität des Testverfahrens immer verbessert werden kann. Schwankungen in diesen Faktoren verursachen, dass die Ergebnisse in

zwei verschiedenen Evaluierungen nicht gleich sind, obwohl es keine Schwankungen in den Sprachkenntnissen gibt. (Huhta & Takala 1999, 212.) Minimierung der Anzahl der Messungsfehler maximiert die Reliabilität.

## **4.2 Validität**

Die wichtigste Qualität in Testinterpretation und Testgebrauch ist die Validität. Wieviel die Performanz eines Sprechers von den Sprachfertigkeiten, die wir messen wollen, abhängt, ist der Grad der Validität. Die Effekte dieser Fertigkeiten auf die Testergebnisse sollen maximiert werden. Ein möglichst zuverlässiges und gültiges Testverfahren heisst also, dass die Effekte der Messungsfehler minimiert und die Effekte bestimmter Sprachfertigkeiten maximiert werden. Wenn die Testergebnisse stark von den Messungsfehlern beeinflusst sind, können sie den Grund für eine valide Interpretation oder einen validen Gebrauch nicht darstellen. Ein Testergebnis, das nicht zuverlässig ist, kann auch nicht valid sein. (Bachman 1990, 25; 161.)

Der Test muss die vorher bestimmten Fertigkeiten messen. Wenn die Testergebnisse von anderen als den gewünschten Fertigkeiten beeinflusst sind, können sie keine sinnvollen Indikatoren einer bestimmten Fertigkeit sein. Wenn zum Beispiel die Schüler zuerst einen Text zuhören und danach über den schreiben sollen, ist die Übung sowohl durch ihre produktiven und als auch ihre rezeptiven Fertigkeiten beeinflusst. (Bachman 1990, 25.) Die Übung kann also nicht als ein gültiger Indikator der schriftlichen Fertigkeit eines Schülers gelten.

Wenn Reliabilität Qualität der Testergebnisse bedeutet, ist Validität Qualität des Testgebrauchs und der Testinterpretation – wie die Ergebnisse eines Testes interpretiert werden. Was beispielsweise

bedeutet, wenn ein Schüler bestimmte Punktzahl auf einen Test bekommt oder eine bestimmte Übung nicht absolvieren kann. Was ist die Rolle der anderen Faktoren – Unterricht, Begabung, Lernschwierigkeiten, mangelnde Motivation usw. – außer Sprachkompetenz?

Ein Test kann also nicht als invalid oder valid betrachtet werden, sondern die Validität bezieht sich auf die Eigenschaften der Schlussfolgerungen, die aus den Ergebnissen gezogen werden. Nur die Schlussfolgerungen können valid sein, nicht der Test. Der Grad der Validität hängt damit zusammen, wie die Interpretation der Ergebnisse begründet wird. Die Begründungen sind vielfältig. Die Validität kann beispielsweise durch die inhaltliche Reliabilität des Testverfahrens motiviert werden oder mit empirischen Beobachtungen bekräftigt werden. (Huhta & Takala 1999, 213.)

Reliabilität und Validität sind also beide wesentlich für den Gebrauch eines Tests. Aber wie Bachman (1990, 26) feststellt, ist eine vollständig fehlerfreie Leistungsmessung vielleicht nie möglich, sondern es muss definiert werden, welcher Grad von Reliabilität oder Validität verlangt wird.

### **4.3 Praktikabilität**

Eine allgemeine Anforderung an die Qualität des Testverfahrens ist auch die Praktikabilität. Die Leistungsmessung soll nicht zu viel Zeit und Ressourcen verbrauchen. Was dieses in Praxis bedeutet, hängt von der Testsituation ab. (Huhta & Takala 1999, 215.) Wie Helvelahti (1996, 19) feststellt, sollen z.B. die landesweiten Prüfungen, die meistens in der neunten Klasse verwendet werden, einfach und schnell zu korrigieren sein. Trotzdem dürfen die Prüfungen keine Ratespiele sein. Deswegen

sollten die Prüfungen sowohl Produktion als auch Multiplechoice enthalten, damit die Arbeit der Lehrer nicht übermäßig groß wird. (Helvelahti 1996, 19.)

Wenn die Leistungsmessung von ihren Folgen her nicht besonders wichtig und bedeutsam ist, sind die Ressourcen meistens auch kleiner. Trotzdem sollte die Qualität nie aufgrund praktischer Gründe schlechter werden: wenn die Ergebnisse nicht zuverlässig sind oder nicht interpretiert werden können, ist die Bewertung umsonst gemacht. (Huhta & Takala 1999, 215.) Andererseits ist ein zuverlässiger und valider Test kaum zu verwenden, wenn er nicht praktisch ist (Weir 1990, 34). Das Wichtigste also ist, einen optimalen Grad der Validität, Reliabilität und Praktikabilität in jeder Testsituation zu finden.

## **5. Aufgabentypen beim Testen der Sprachkenntnisse**

Die heterogenen Gruppen an den heutigen Schulen erschweren das Verfassen der Tests. Einerseits soll es in einem Test anspruchsvolle Aufgabe für die Schüler sein, die im Sprachunterricht schnelle Fortschritte machen. Andererseits dürfen die schwierigsten Aufgaben nicht so kompliziert sein, dass niemand die lösen kann. Ein Test sollte auch immer etwas enthalten, die alle absolvieren können. (Helvelahti 1996, 19 – 20.)

Helvelahti (1996) findet es besonders schwer, solche Übungen zu verfassen, die auch die schwachen Lerner absolvieren können. Um die Übungen absolvieren zu können, müssen die Lerner zeigen, dass sie etwas wirklich beherrschen. Es soll beispielsweise Nummer in einer Hörübung verstanden werden oder Farben in einem Text erkannt

werden. Nach dem Test soll der Schüler fühlen, dass er etwas aufgrund seiner Sprachkenntnisse richtig absolviert hat und nicht nur geraten oder mit Hilfe von Allgemeinwissen die Übungen gelöst hat – ohne den Text zu lesen oder zu hören. (Helvehtti 1996, 19 – 20.) Die Übungstypen sollen in der Klasse vorher geübt werden, damit sie für die Schüler in der Testsituation bekannt sind (Weir 1993, 69).

Die verschiedenen Teilkomponenten der Sprachkompetenz werden auf unterschiedliche Weise bewertet. Verschiedene Typen von Testaufgaben werden nicht nach den vier sog. Hauptfertigkeiten – Sprechen, Schreiben, Hörverständnis und Leseverständnis – eingeteilt, sondern es wird zwischen Produktionsfähigkeit (Sprechen und Schreiben) und Verstehen einen Unterschied gemacht. Wenn Produktionsfähigkeit ziemlich explizit getestet werden kann, kann das Verstehen nur implizit geschätzt werden – wie der Lerner eine fremde Sprache versteht, kann nur aufgrund der Antworten auf das Gefragte beurteilt werden. (Huhta & Takala 1999, 204 – 206.)

Wenn das Schreibvermögen bewertet wird, sind die häufigsten Testtypen allgemeine, freie Aufsätze oder genauer definierte schriftliche Aufgaben. Die üblichsten Testtypen, mit dem sowohl das Verstehen als auch Wortschatz und Grammatik in einem schriftlichen Test getestet werden, sind Multiplechoiceverfahren, offene Fragen und Lückenfüllung. (Huhta & Takala 1999, 209.) Im folgenden werden die üblichsten Testtypen in einem schriftlichen Test besprochen. Das Testen der mündlichen Sprachkompetenz wird auch kurz behandelt, obwohl die mündliche Kompetenz in dieser Arbeit kein zentrales Thema ist.

## 5.1 Testen der schriftlichen Produktionsfähigkeit

Die schriftliche Produktionsfähigkeit wird meistens durch freie Aufsätze oder gesteuerte Aufgaben getestet. Weir (1993, 134) betont, dass es in einem Test immer mehr als nur eine Schreibaufgabe geben soll, weil die Performanz eines Individuums sogar in sehr ähnlichen Aufgaben variieren kann. Nur dann haben die Lernenden die Möglichkeit zu zeigen, was sie können. Der Grad der Reliabilität und Validität eines Testes wird wesentlich höher, wenn es mehrere Proben von dem Schreibvermögen gibt. Laut Weir ist kaum zu vermuten, dass nur eine Schreibaufgabe die allgemeine Schreibkompetenz eines Schülers testen könnte. Je mehrere Proben es gibt, desto zuverlässiger sind die Testergebnisse. (Weir 1993, 134.)

### 5.1.1 Traditionelle, freie Aufgaben

Mit traditionellen, freien Aufsätzen wird das Schreibvermögen explizit bewertet. Die Themen sind oft allgemein und stützen sich auf das kulturelle Wissen, Allgemeinwissen oder Phantasie des Schülers. Das Thema spielt in einer Schreibaufgabe eine wichtige Rolle. Das Thema soll bekannt für die Schüler sein, damit sie darüber schreiben können. Wenn die Schüler die Übung unrealistisch, unpassend, sinnlos oder unmöglich finden, schreiben sie vermutlich nicht so gut wie sie eigentlich könnten. Was, warum und wem die Schüler schreiben sollen, wird nicht erklärt. Die Aufgaben haben keine klaren Motivierungen und keine definierten Ziele. Die Schreiber werden nur kurz oder gar nicht darüber informiert, was sie schreiben sollen und wie der Text beurteilt wird. (Weir 1993, 135 – 143.)

Es ist aber zu bemerken, dass die Schüler oft solche Übungen sinnlos und schwer finden, die keinen Zweck haben. Die Leser des Textes sind

unidentifiziert und die Aufgabe in sich hat kein Ziel. (Huhta & Takala 1999, 209.) Die Schreiber produzieren sehr unterschiedliche Texte. Deshalb ist es schwer, die Texte exakt und zuverlässig zu beurteilen. Wenn es zusätzlich mehrere Alternativen in Themen gibt, ist es kaum möglich, verschiedene Texte zu vergleichen. Wie gut der Schreiber die Aufgabe absolviert, hängt von seinem Wissen, seiner Kreativität und Phantasie ab. Wenn der Verfasser eines Tests beurteilen will, wie die Schüler neutrale und informative Prosa in einer bestimmten Kommunikationssituation produzieren, ist ein freier Aufsatz nicht passend, um diese Kompetenz zu evaluieren. (Weir 1993, 144.)

Aus diesen Gründen findet Weir (1993, 144) einen freien Aufsatz invalid, um die Schreibkompetenz zu beurteilen. Eine solche Aufgabe hat jedoch ihre Vorteile. Eine freie Aufsatzübung ermöglicht, dass mehrere Merkmale des Schreibvermögens beurteilt werden können (Huhta & Takala 1999, 209). Ein Aufsatz bietet auch die Möglichkeit, die Fortschritte eines Lerners abzuschätzen. Es ist leicht, den Text mit früheren oder späteren Texten gleicher Art zu vergleichen. Es ist auch ziemlich einfach, eine solche Aufgabe zu verfassen. (Weir 1993, 143 – 144.)

Freier Aufsatz hat also Vor- und Nachteile. Obwohl der Testtyp viel kritisiert wird, wird er trotzdem verwendet. Das Verfassen einer solchen Übung ist leicht, obwohl die Korrigierung keine einfache Prozedur ist. Auf jeden Fall ist zu vermuten, dass ein freier Aufsatz wegen seiner etablierten Stellung in der Beurteilung einen festen Platz unter den Testaufgaben auch in der Zukunft hat.

### 5.1.2 Gesteuerte Aufgaben

Der neue kommunikative Unterricht hat solche Aufgaben mitgebracht, die in einen Kontext eingebettet sind und ausführliche Instruktionen haben. Es wird definiert, warum und für wen der jeweilige Text geschrieben wird und was für Informationen vermittelt werden sollten. In der Beurteilung wird versucht, die soziolinguistische Kompetenz zu schätzen, die die heutigen Auffassungen über Sprachkompetenz betonen. (Huhta & Takala 1999, 208 – 209.) Der Zweck solcher gesteuerten Aufgaben ist oft der Versuch abzuschätzen, wie gut der Lerner die fremde Sprache unter den Voraussetzungen jeweiliger Kommunikationssituation anwenden kann.

Es gibt viele verschiedene Typen der Produktionsaufgaben. Als Beispiele können Lückenfüllung, Formularfüllung, Beantwortung auf gegebene Information und Korrekturlesen genannt werden. (Siehe Weir 1993, 130 – 154.)

## 5.2 Testen der rezeptiven Sprachfähigkeit

Das Verstehen wird hauptsächlich explizit durch Lese- und Hörverständnisübungen getestet. In schriftlichen Tests wird das Leseverständnis auch in anderen Übungstypen (z.B. Lückenfüllung, die sowohl Produktionsfähigkeit als auch Leseverständnis testet) implizit getestet. Hier wird jedoch das explizite Testen dieser Fähigkeiten besprochen.

Helvelahti (1996, 19) findet, dass die Multiplechoiseaufgaben zu viel kritisiert werden. Sie können leicht oder schwer sein – ebenso gut wie andere Aufgaben. Wenn es genug Antwortalternativen gibt, können die Schüler die richtigen Antworten nicht raten.

### 5.2.1 Leseverständnis

Helvelahti (1996) meint, dass Schüler schon in der Grundschule lernen sollen, verschiedene Texte zu lesen. Eine wichtige Fertigkeit ist, das Wesentliche in einem Text zu finden, trotz fremden Wörtern und schwierigen Äußerungen, die es im Text geben kann. Diese Fähigkeit ist in der Weiterbildung und später im Arbeitsleben notwendig. Wortgetreue Interpretation ist auch eine wichtige Fähigkeit, aber als die einzige Weise, verschiedene Texte zu behandeln, ist sie einseitig und unzureichend. Die Texte sollen in der ersten Linie inhaltlich verstanden werden. (Helvelahti 1996, 19.)

Sprachlich soll der Text in einer Leseverständnisprüfung nicht zu schwer sein. Der Text sollte nicht zu viele fremde Wörter enthalten und keine solchen Strukturen, die die Schüler nicht beherrschen. Der Text soll nicht zu viel neue Information beinhalten. Es müssen Überlegungen darüber gemacht werden, wie lang der Text sein soll und wieviel Zeit die Schüler brauchen, um den Text zu lesen. Wenn der Text zu kurz ist, kann nicht getestet werden, wie die Schüler neue Information schnell erfassen. Kurze Texte testen nur intensives Lesen. (Weir 1993, 67 – 68.)

Weir (1993) stellt fest, dass im allgemeinen oft gewünscht wird, dass die Texte sowohl in den Leseverständnisprüfungen als auch in den Hörverständnisprüfungen interessant wären. Das Thema des Textes soll nicht zu fremd oder eigentümlich sein. Trotzdem soll es nicht zu bekannt sein, damit die Schüler auf die Fragen aufgrund ihres Allgemeinwissens nicht antworten können. In einem Idealfall behandelt der Text in einer Lesverständnisprüfung ein solches Thema, das alle Teilnehmer interessant finden und das mit dem früheren Wissen der Teilnehmer verknüpft werden kann. (Weir 1993, 67.)

Helvelahti (1996) will jedoch in Frage stellen, ob der Text interessant sein soll. Er meint, dass es in einer Prüfungssituation immer um unnatürliche Anwendung der Sprache geht, wo der Schüler die Aufgaben absolviert, um eine möglichst gute Note zu bekommen und nicht daran denkt, ob die Texte interessant sind oder nicht. Die Teilnehmer sollen auf jeden Fall gleichgestellt sein. Niemand sollte mehr über das Thema wissen als die anderen, sondern alle sollten die Texte lesen oder hören, um die gestellten Fragen antworten zu können. Weil im Sprachunterricht heute die Kulturkenntnisse betont werden, sollen die Prüfungstexte vorzugsweise die Länder behandeln, wo die jeweilige Sprache gesprochen wird oder die Menschen, die die Sprache sprechen. (Helvelahti 1996, 20.)

### *5.2.2 Hörverständnis*

Hörverständnisprüfungen testen die Sprachkenntnisse hauptsächlich auf dieselbe Weise wie die Leseverständnisprüfungen. In beiden geht es ja um das Testen einer rezeptiven Fähigkeit. (Weir 1993, 99.)

Die Aufgabentypen variieren sehr. Die Schüler werden z.B. gebeten, Sätze zu ergänzen oder einen Weg auf einer Karte nach Anweisungen zu zeichnen (siehe Weir 1993, 104 – 122).

Es gibt viele verschiedene Texttypen, die in einer Hörverständnisprüfung verwendet werden können. Die Texte können Monologe, Dialoge oder Gespräche sein. Sie können beispielsweise Geschichten, Gedichte, Vorlesungen, Anweisungen, Nachrichten, Telefongespräche oder Interviews sein. Die Texte können kurz oder lang, formell oder informell sein. Die Leser haben zusätzlich verschiedene Akzente, Dialekte und Sprechmelodien. Heute werden die Texte nicht mehr so oft in kleine

Abschnitten geteilt, sondern es wird versucht, die Authentizität zu erhalten. (Weir 1993, 99 ; 101 – 102.)

### **5.3 Testen der mündlichen Sprachkompetenz**

Vollmer (1995) meint, dass die Schüler in der Zukunft vor allem Fremdsprachenkenntnisse im mündlichen Bereich benötigen; diese sind deshalb in der Schule zu lehren. Die mündlichen Fertigkeiten und insbesondere die Kommunikationsfähigkeit sollte regelmäßig und zuverlässig beurteilt werden. (Vollmer 1995, 276 – 277.)

In der finnischen Grundschule haben die kommunikativen Fähigkeiten schon eine zentrale Rolle im Sprachunterricht. Es ist jedoch festgestellt worden, dass manche Schüler Schwierigkeiten im Wortschatz und in der Grammatik haben. Der Schwerpunkt in dem heutigen Sprachunterricht liegt in der Vermittlung der Information, nicht in der Präzision der sprachlichen Äusserung. Pasanen (1994) meint, dass es ein Gleichgewicht zwischen Forderungen nach der Geläufigkeit und der Korrektheit geben soll. (Pasanen 1994, 16.) Es ist zu vermuten, dass andere Fähigkeiten als nur die kommunikative Fertigkeit für die Schüler in der Zukunft ebenso wichtig sind und deshalb im Sprachunterricht zu lehren sind.

Die Testtypen beim Testen der mündlichen Kompetenz sind Face-to-Face Gespräche und Sprechübungen im Sprachlabor (Huhta & Takala 1999, 209). Die Hindernisse für das Testen der mündlichen Sprachkompetenz sind oft grosser Zeitaufwand, schwieriger Einsatz technischer Medien, Schwierigkeiten in den Auswertungsprozeduren und mangelndes Training in diesem Bereich (Vollmer 1995, 276).

## 6. Methoden

In diesem Kapitel werden die Untersuchungsmethoden besprochen. Die Untersuchung wurde unter Mitwirkung des Verbandes Suomen Kieltenopettajien liitto ry (Sukol) verwirklicht. In der Untersuchung wurde ein Test verwendet, der von Sukol-Palvelu Oy und Suomen Kieltenopettajien Liitto ry für Schüler mit A1- und A2-Deutsch in der sechsten Klasse für das Jahr 2000 erstellt worden ist (Anhang 1).

Mit Hilfe eines Fragebogens (Anhang 4), der an die Lehrer gerichtet war, wurde auch Information gesammelt. Die Vereinssekretärin Päivi Koivistoinen-Toivonen und die Geschäftsführerin Sari Saarinen von Sukol haben Hinweise beim Verfassen des Fragebogens gegeben.

Die Untersuchung kann für eine quantitative Survey-Untersuchung gehalten werden, wo eine Befragung laut Hirsjärvi, Remes und Sajavaara (1997) eine zentrale Methode ist. Die Adressaten bilden den Forschungsgegenstand oder stehen als eine Auswahl für den ganzen Forschungsgegenstand. Kennzeichnend für die Befragung in einer Survey-Untersuchung ist, dass meistens ein Fragebogen oder ein strukturiertes Interview verwendet wird. Die Fragen werden an die Adressaten genau auf dieselbe Weise gestellt. Das Material ist häufig relativ klein und wird quantitativ analysiert. Mit Hilfe der gesammelten Information wird versucht, die Phänomene zu beschreiben, zu vergleichen und zu erklären. (Hirsjärvi, Remes & Sajavaara 1997, 130; 189.)

In dieser Arbeit wird versucht, die Unterschiede der Schüler mit A1- und A2-Deutsch zu beschreiben und zu vergleichen. Was mit den Begriffen A1-Sprache und A2-Sprache im finnischen Bildungssystem verstanden wird, wird näher im Folgenden erklärt. Später in diesem Kapitel werden

die Untersuchungsmittel, der Forschungsgegenstand und der Verlauf der Untersuchung besprochen.

### **6.1 A1- und A2-Sprachen**

In der finnischen Grundschule lernen alle Schüler mindestens zwei Fremdsprachen (Kimmoke-indikaattorit 1999). In der Unterstufe wird zusätzlich zu der Muttersprache mindestens eine zweite, obligatorische [fremde] Sprache gelernt. In einem Teil der Schulen wird auch eine dritte, fakultative [fremde] Sprache unterrichtet. In der Oberstufe lernen alle Schüler Muttersprache und zwei fremde Sprachen. Auch mehrere, fakultative Sprachen können unterrichtet werden. (Peruskoulun opetussuunnitelman perusteet 1994, 17.)

In *Peruskoulun opetussuunnitelman perusteet* (1994, 18) wird definiert, in welchem Umfang die [fremden] Sprachen unterrichtet werden. A1-Sprache ist die erste [fremde] Sprache in der Unterstufe, die für alle gemeinsam ist. A2-Sprache ist die in der Unterstufe angefangene fakultative Sprache. Das Ziel ist, dass die sprachliche Kompetenz der Schüler mit einer A1-Sprache und der Schüler mit einer A2-Sprache in der jeweiligen fremden Sprache am Ende der Grundschule auf dem selbem Niveau ist. In der Oberstufe werden obligatorische B1-Sprache und fakultative B2-Sprache unterrichtet und in der gymnasialen Oberstufe eine fakultative B3-Sprache unterrichtet. (Peruskoulun opetussuunnitelman perusteet 1994, 18.)

Die heutige Stundenverteilung in der finnischen Grundschule stammt vom Jahr 1993. Laut diesem Beschluss soll die in der Unterstufe anfangende Fremdsprache (A-Sprache) mindestens 8 Stunden pro Woche unterrichtet werden. (Koulusäädökset 1999, 21.) Die wöchentlichen Mindeststunden der zweiten, fakultativen Fremdsprache

[A2-Sprache] sind 4 Stunden. Jede Schule hat einen eigenen Lehrplan. Im Lehrplan wird festgelegt, welche fremde Sprachen unterrichtet werden, in welcher Klasse der Fremdsprachenunterricht beginnt und wieviele Stunden die fremden Sprachen unterrichtet werden. (Peruskoulun opetussuunnitelman perusteet 1994, 17.)

In den meisten Unterstufen wird mit der Fremdsprachlernung in der 3. Klasse begonnen, aber in einem kleinen Teil der Schulen beginnt der Fremdsprachenunterricht schon in der 1. oder 2. Klasse. Die erste obligatorische Fremdsprache wird A1-Sprache genannt. Die üblichste A1-Sprache ist Englisch – fast 90% der Schüler eines Jahrgangs haben im Schuljahr 1998-1999 Englisch als erste Fremdsprache gewählt. Die zweitüblichste A1-Sprache ist Deutsch (3,7 % der Schüler im Schuljahr 1998-1999) und die dritte Schwedisch (2,4%). (Kimmoke-indikaattorit 1999.)

A2-Sprache ist die zweite, fakultative Fremdsprache. Der Unterricht der A2-Sprache beginnt am häufigsten in der 5. Klasse, aber in einigen Schulen schon in der 3. oder 4. Klasse. Im Schuljahr 1998-1999 haben fast zwei Drittel der Schüler in der 5. Klasse eine A2-Sprache studiert. Die häufigste A2-Sprache ist Deutsch (42,3 % von allen, die eine A2-Sprache in der 5. Klasse im Schuljahr 1998-1999 gewählt hatten). (Kimmoke-indikaattorit 1999.)

## **6.2 Die Untersuchungsmittel**

Das primäre Untersuchungsmittel in dieser Untersuchung war die von Sukol-Palvelu Oy und Suomen saksanopettajat ry erstellte Prüfung. Obwohl die neueste, für das Jahr 2001 erstellte Prüfung auch zur Verfügung gestanden hätte, wurde die Prüfung von Jahr 2000 gewählt,

weil sie die letzte von Sukol erstellte Prüfung für sowohl A1-Lernende als auch A2-Lernende war.

Durch eine Befragung wurde Information über Hintergrundfaktoren und Meinungen der Lehrer gesammelt. Die Lehrer wurden gebeten, einen Fragebogen auszufüllen. Der Fragebogen (Anhang 4) bestand aus drei Teilen. Im ersten Teil wurde nach Kommentaren über die Prüfung gefragt. Im zweiten Teil wurden die Lehrer gebeten, ihre Meinungen über die Unterschiede in den Sprachkenntnissen von A1- und A2-Lernenden zu äußern. Im dritten Teil wurde nach Hintergrundinformation über die Lage der Schule, die Größe der Lerngruppe, die Anzahl der Unterrichtsstunden usw. gefragt.

### **6.3 Forschungsgegenstand**

Der Forschungsgegenstand bestand aus 36 Schulen. Bei Auswahl des Gegenstandes war das Kriterium, dass die Schulen die Prüfung von Sukol sowohl im Jahr 2000 als auch im Jahr 2001 bestellt hatten. Aus praktischen Gründen war es notwendig, dass die Schulen die Prüfung vom Jahr 2000 schon hatten und dass sie für die Lehrer dieser Schulen bekannt war. Das zweite Kriterium war, dass die Schulen die Prüfung von Sukol auch im Jahr 2001 bestellt hatten. Es wurde vermutet, dass dieses zeigte, dass die Lehrer in diesen Schulen mit der früheren Prüfung zufrieden gewesen sind und dass die Prüfung ihrer Meinung nach die Deutschkenntnisse der Schüler auf eine zweckmäßige Weise testet.

Die Schulen waren aus großen und kleinen Gemeinden in verschiedenen Teilen Finnlands. In dieser Hinsicht kann die Auswahl geografisch als relativ repräsentativ angesehen werden.

#### 6.4 Verlauf der Untersuchung

Den gewählten 36 Schulen wurde Ende März 2001 ein Brief (Anhang 2) geschickt, in dem gefragt wurde, ob die Schulen an der Untersuchung teilnehmen wollten. Anbei wurde auch eine Fürsprache von Sukol (Anhang 3) geschickt.

Im Brief wurde das Untersuchungsverfahren erklärt und die Deutschlehrer wurden gefragt, ob sie die Untersuchungsprüfung in ihren Deutschgruppen in der sechsten Klasse schreiben lassen könnten. Sie wurden gebeten, die Teilnahme ihrer Schule zu bestätigen und die Größe der Deutschgruppe, die die Prüfung schreiben würde, mitzuteilen.

Für jede Schule wurde eine eigene Forschungsnummer gegeben. Die Fragebögen wurden nummeriert, um die Behandlung der Fragebogen zu erleichtern. Dadurch konnte auch verfolgt werden, welche Schulen ihre Beteiligung schon bestätigt hatten.

Anfang April hatten Lehrer von zehn Schulen ihre Beteiligung bestätigt. 11 Lehrer meldeten 12 Gruppen an, in denen es insgesamt 194 Schüler gab. Es wurde ein zweiter Brief (Anhang 6) an die anderen Schulen geschrieben, um nach ihrer Interesse nochmal zu fragen. Keine neuen Anmeldungen wurden empfangen.

Für die zehn Gruppen, die sich gemeldet hatten, wurden der Fragebogen und ein Briefumschlag für Rückantwort geschickt. Die Prüfung wurde in 12 Gruppen geschrieben und die Lehrer haben die Prüfungen korrigiert. Es wurden Testergebnisse von insgesamt 175 Schülern empfangen. Drei Gruppen mit 56 Schülern wurden wegen unvollständiger Antworten abgelehnt. Der endgültige Querschnitt bestand aus 9 Gruppen, in denen es 119 Schüler gab. In vier Gruppen mit 48 Schülern wurde Deutsch als A1-Sprache und in fünf Gruppen mit 71 Schülern als A2-Sprache gelernt.

Von fünf Gruppen (69 Schüler) standen die Prüfungspapiere zur Verfügung und von vier Gruppen (50 Schüler) wurden die Punkte einzelner Aufgaben im Formular (Anhang 4) mitgeteilt. Alle 11 Lehrer hatten den Fragebogen ausgefüllt.

Die Prüfungsergebnisse wurden statistisch erfasst. Die Hintergrundinformation der Schüler wurde mit statistischen Daten zusammengefügt. Diese Daten wurden mit dem Computerprogramm SPSS analysiert. Die Information, die mit den offenen Fragen und Meinungsfragen im Fragebogen gesammelt wurde, wurde mit dem Testserfolg verglichen.

## **7. Ergebnisse**

Die schriftliche Prüfung, die in dieser Untersuchung verwendet wurde, testete sowohl produktive als auch rezeptive Sprachkompetenz. Im diesem Kapitel werden zuerst die Meinungen der Lehrer über die Prüfung erläutert. Danach werden die Testergebnisse der Schüler besprochen.

### **7.1 Meinungen der Lehrer über die Testaufgaben**

Die Lehrer wurden im Fragebogen gefragt, ob der Test ihrer Meinung nach den zentralen Wortschatz und wichtige sprachliche Fertigkeiten getestet hat. Es wurde eine fünfstufige Skala verwendet, wo '1' volle Zustimmung ('täysin samaa mieltä') und '5' das Gegenteil ('täysin eri mieltä') präsentierten. Die Mittelwerte der Antworten waren 2,75 und 3,25. Aufgrund dieser Mittelwerte ist zu vermuten, dass die Lehrer

meinten, dass der zentrale Wortschatz ziemlich gelungen und die wichtigsten Fertigkeiten befriedigend getestet wurden.

Die Lehrer wurden nach ihren Meinungen gefragt, ob die Punkte, die die Schüler von einzelnen Aufgaben bekommen hatten, zeigten, wie gut die Schüler die Sprache beherrschen. Die Lehrer wurden also gefragt, wie gut die Aufgaben das messen, was sie messen sollten. Die Lehrer hatten sehr unterschiedliche Meinungen und haben beim Antworten die ganze Skala verwendet. Durchschnittlich waren die Aufgaben nach ihren Meinungen am meisten nur befriedigend oder misslungen. Im allgemeinen meinten die Lehrer, dass die Testergebnisse etwas über die sprachliche Kompetenz der Schüler zeigten, aber betonten, dass der Test nicht ganz gelungen war. Sie meinten, dass der Test nicht vorzugsweise die Fertigkeiten getestet hat, die nach dem Lehrplan (fi. *opetusuunnitelma*) zentral sind, obwohl sie zugaben, dass der zentrale Wortschatz getestet wurde.

Die Lehrer wurden im Formular gebeten, ihre Meinungen über die einzelnen Aufgaben und die Prüfung frei zu äußern.

### *7.1.1 Testen der produktiven Sprachfähigkeit*

Das Produktionsvermögen wurde in dieser Prüfung mit einer Lückenfüllungsaufgabe, Wortschatzübung und mit einem kleinen Aufsatz getestet.

Lückenfüllung wird hier eine solche Aufgabe genannt, in der Lücken in einem Text erfüllt werden sollen. Die Übung (*Rakennetehtävä*, auf der Seite 5 in der Prüfung) testete Wortschatz und grammatische Strukturen. Nach den Lehrern war die Aufgabe schwer und unklar und der Übungstyp neu für viele Schüler. Die Lehrer meinten, dass die Aufgabe

auf eine solche Weise verfasst war, dass die Schüler Schwierigkeiten hatten, den Text und die Aufgabe in allgemeinen zu verstehen. Im Text gab es viele fremde Wörter und Strukturen, die Unsicherheit verursachten. Laut Lehrer haben die Schüler deswegen teilweise falsch geantwortet, obwohl sie die richtige Antwort gewusst hätten. Wie ein Lehrer feststellte, ist es schwer für die Schüler, viele Strukturen und grammatische Einzelheiten gleichzeitig zu beherrschen. Sie erinnern sich nur an die zuletzt gelernten grammatischen Einzelheiten, das früher Gelernte haben sie schon vergessen.

Wortschatz wurde im Bereich Lebensmittel gefragt (*TV-tarjotin*, S. 4). Die Schüler wurden gebeten, zehn verschiedene Lebensmittel zu nennen. Die Lehrer meinten, dass diese alltäglichen Wörter wichtig sind und die Übung deshalb gut war. Die Lehrer haben bemerkt, dass die Schüler das Genus der Wörter nicht kennen oder sich an das Genus nicht erinnern.

Das Schreibvermögen wurde mit Hilfe eines kurzen Aufsatzes beurteilt (*Kirjoitelma*, S. 6). Die Schüler wurden gebeten, mit sieben Sätzen über ein Bild zu schreiben. Die Lehrer meinten, dass das Bild zu arm war und dass es schwer war, die verschiedenen Texte der Schüler zu beurteilen.

### *7.1.2 Testen der rezeptiven Sprachfähigkeit*

Der Text in der Leseverständnisaufgabe (*Luetun ymmärtäminen*, S. 2) erzählte über Snowboarden. Die Lehrer meinten, dass die Fragen über den Text zu umfangreich und teilweise auch komisch waren. Im Text gab es fremde Strukturen und zu lange Sätze. Ein Lehrer kritisierte, dass um die Aufgabe zu absolvieren, die Schüler nicht nur Deutschkenntnisse sondern auch andere Fähigkeiten brauchen: die Schüler müssen verstehen, was eigentlich gefragt wird und danach müssen sie die wesentliche Information im Text finden.

Die Aufgabe, in der die Schüler zehn unterschiedliche Behauptungen als richtig oder falsch markieren sollten (*Richtig - falsch / oikein - väärin O/V*, S. 4), war nach den Lehrern misslungen. Die meisten von den Sätzen waren solche Behauptungen, dass die Schüler nicht unbedingt wussten, ob sie richtig oder falsch waren, unabhängig davon, ob die Schüler sie verstanden oder nicht. Beispielsweise *Turku ist älter als Helsinki* (Satz 4), *Him ist populär in Deutschland* (Satz 9) oder *Tiktak ist eine schwedische Band* (Satz 10) sind keine selbstverständlichen Fakten für Schüler der sechsten Klasse.

In einer Aufgabe (*Yhdistä kysymys ja vastaus*, S. 4) sollten die Schüler Fragen mit richtigen Antworten kombinieren. Die Lehrer meinten, dass diese Übung gelungen war und wesentliche Fertigkeiten testete.

Die Hörverständnisprüfung (*Kuullun ymmärtäminen*, S. 3) wurde im allgemeinen für schwer gehalten. Die Lehrer meinten, dass der Text passend war, aber die Fragen, auf die die Schüler antworten sollten, zu kompliziert waren und die Zeit zu knapp war. Das Thema war auch nicht besonders interessant. Die Qualität des Tonbandes war schlecht. Ein Grund für den schlechten Erfolg in dieser Aufgabe war nach den Lehrern, dass die deutsche Sprache fremd für die Schüler klingt. Manche sagten, dass dieselbe Aufgabe viel leichter auf Englisch gewesen wäre. Das haben besonders die Lehrer der A2-Gruppen betont.

## 7.2 Hintergrundfaktoren

Der Testserfolg einzelner Schüler und Gruppen wurde mit den Variablen verglichen, die im Fragebogen gefragt wurden. Der primäre Faktor war der Umfang der gelernten Sprache (A1-/A2-Sprache). Die anderen Hintergrundfaktoren waren die Lage der Schule (der Bezirk), die Einwohnerzahl der Gemeinde, die Anzahl von Schülern in der Schule,

die Anzahl von Schülern in der Lerngruppe, die Dienststellung der Lehrer und der akademische Grad der Lehrer im Fach Germanistik und das Geschlecht des Schülers. Zusätzlich wurde im Fragebogen gefragt, in welcher Klasse der Unterricht begonnen hatte und wie hoch die durchschnittliche Zahl von Unterrichtsstunden war.

Mit statistischen Erhebungen wurden Mittelwerte der Gruppen und Schüler von den Gesamtpunkten und einzelnen Aufgaben gerechnet. Die Maximalpunkte im Test waren 100. Die Mittelwerte wurden aufgrund verschiedener Faktoren klassifiziert.

### *7.2.1 A1- bzw. A2-Deutsch*

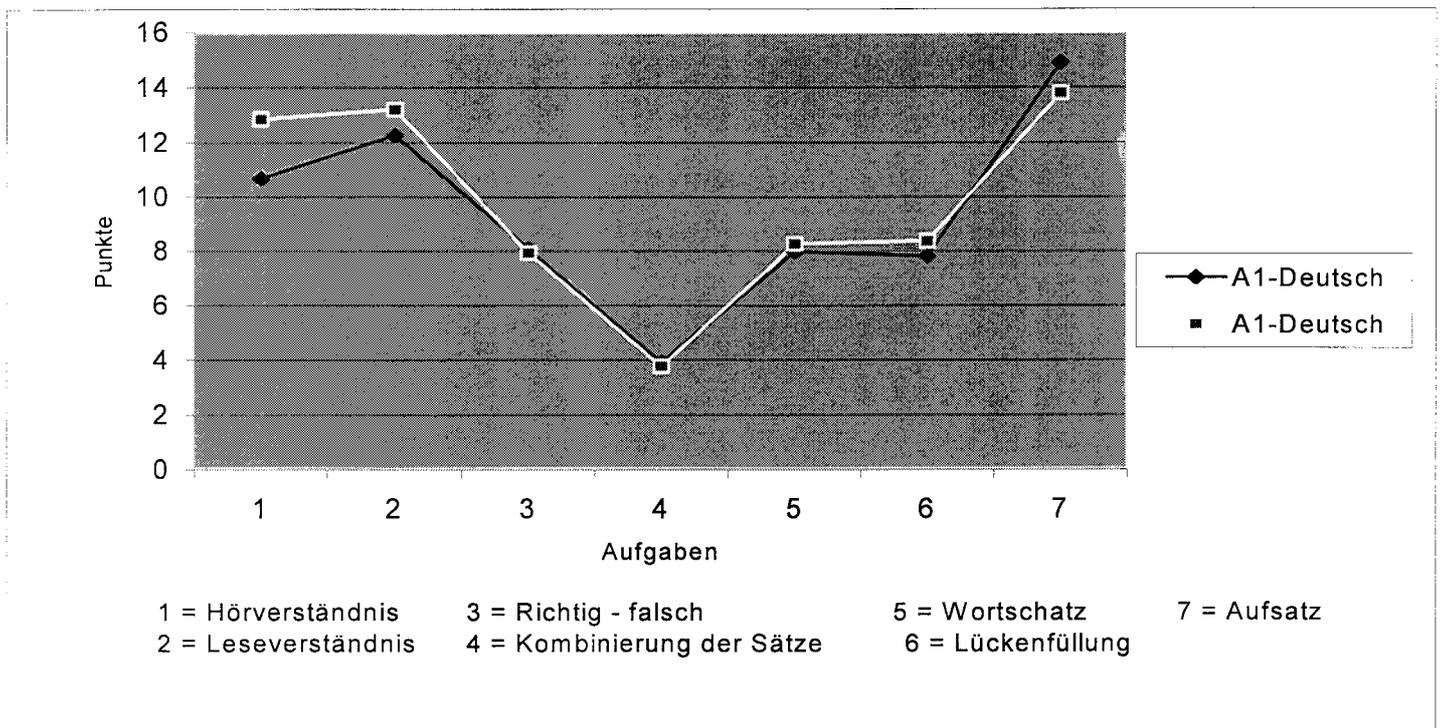
Von den 119 Schülern, die an dieser Untersuchung teilnahmen, haben 48 Deutsch als erste Fremdsprache und 71 als zweite Fremdsprache, A2-Sprache. Der Mittelwert der Testergebnisse der A1-Schüler war 64,90 und der Mittelwert der A2-Schüler 67,69.

Unter den Gruppen, die Deutsch als A1-Sprache lernten, gab es keine großen Unterschiede zwischen den Gruppen. Die durchschnittlichen Punkte der Gruppen waren zwischen 61,73 und 68,08.

Unter den Gruppen, die Deutsch als A2-Sprache lernten, unterschieden sich die durchschnittlichen Punkte voneinander viel mehr. Sowohl die Gruppe, die den Test am besten absolviert hat, als auch die schwächste Gruppe, waren beide A2-Gruppen. Der niedrigste Durchschnitt war 54,17, was viel niedriger ist als die durchschnittlichen Ergebnisse der A1-Gruppen. Der höchste Durchschnitt hatte dagegen eine andere A2-Gruppe mit einem Durchschnitt von 77,40. Das ist bedeutend höher als der Mittelwert von der besten A1-Gruppe (68,08).

Wenn die Durchschnitte einzelner Aufgaben betrachtet werden, kann festgestellt werden, dass es nur kleine Unterschiede zwischen A1- und A2-Lernern gibt.

GRAFIK 2. Die durchschnittlichen Punkte für einzelne Aufgaben bei Schülern mit A1- bzw. A2-Deutsch.



Wie in der Grafik 2 zu sehen ist, haben die Schüler mit A2-Deutsch durchschnittlich bessere Punkte in den Hör- und Leseverständnisaufgaben bekommen. Die Lerner der A1-Deutsch haben dagegen etwas besseren Erfolg im Aufsatz.

### 7.2.2 Größe der Schule

Die Schulen wurden aufgrund der Größe in drei Kategorien eingeordnet. Es gab eine Schule mit unter 300 Schülern und eine mit über 500 Schülern. Die Zahl von Schülern in anderen Schulen war 300 - 350.

GRAFIK 3. Die Größe der Schule.

Zahl der Schüler in die Schule	Durschnittswert	Zahl der Schüler
Unter 300	66,50	12
300 - 500	65,45	97
Über 500	77,40	10
Insgesamter Durchschnittswert	66,56	119

Wenn die durchschnittlichen Ergebnisse verglichen wurden, wurde festgestellt, dass die Gruppe in der größten Schule (über 500 Schüler) den höchsten Mitterwert (77,40) hatte. Wie die Grafik 3 zeigt, ist der Unterschied in Mittelwerten zwischen den anderen Kategorien klein (66,50 und 65,45).

### 7.2.3 Größe der Lerngruppe

Die Gruppen wurden aufgrund ihrer Größe in drei Kategorien eingeordnet: in Gruppen von unter 12 Schülern, von 12 - 15 Schülern bzw. über 15 Schülern.

GRAFIK 4. Größe der Lerngruppe.

Zahl der Schüler in der Gruppe	Durchschnittswert	Zahl der Schüler
Unter 12	68,59	32
12 - 15	62,84	50
Über 15	69,84	37
Insgesamter Durchschnittswert	66,56	119

Die meisten Gruppen hatten 12 - 15 Schüler. In dieser Mittelklasse waren die Durchschnitte einigermaßen niedriger als in größeren und kleineren Gruppen. Wenn die Mittelwerte in der Grafik 4 betrachtet werden, kann festgestellt werden, dass die Ergebnisse am besten in Gruppen mit weniger als 12 Schülern oder mehr als 15 Schülern waren.

#### 7.2.4 Die demografischen Faktoren – die Lage und Einwohnerzahl der Gemeinde

Von den Schulen, die an dieser Untersuchung teilnahmen, befanden sich sechs in Süd-Finnland in *Etelä-Suomen lääni*, eine in Ost-Finnland in *Itä-Suomen lääni* und zwei in Mittelfinnland. Die Schulen in Süd-Finnland hatten die besten Ergebnisse.

Die Gemeinden wurden aufgrund ihrer Einwohnerzahl in vier Kategorien eingeordnet. Die meisten Schüler (47 Schüler) wohnten in großen Gemeinden mit über 100 000 Einwohnern, der andere große Teil (34 Schüler) in mittelgroßen Gemeinden (12 000 - 45 000). Die zwei anderen Kategorien waren eine mit einer Einwohnerzahl von 45 000 - 100 000 und eine andere mit weniger als 12 000 Einwohnern.

GRAFIK 5. Größe der Gemeinde.

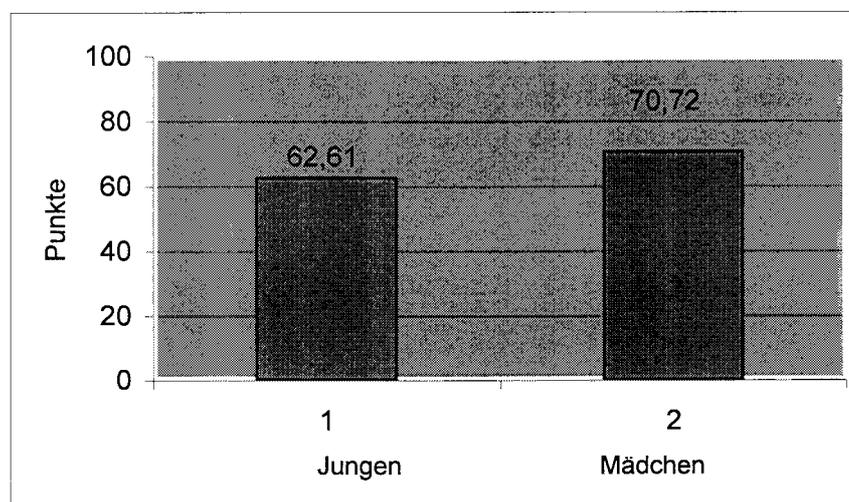
Einwohnerzahl der Gemeinde	Durchschnittswert	Zahl der Schüler
Unter 12 000	60,33	24
12 000 - 44 999	65,82	34
45 000 - 100 000	62,64	14
Über 100 000	71,45	47
Insgesamter Durchschnittswert	66,56	119

Zusammenfassend kann festgestellt werden, dass den besten Erfolg die Schüler in großen Gemeinden (über 100 000 Einwohner) hatten, die alle in Süd-Finnland (*Etelä-Suomen lääni*) lagen.

### 7.2.5 Geschlecht des Schülers

Die Jungen hatten in dieser Untersuchung eine kleine Mehrheit. In den Gruppen, die in dieser Untersuchung teilnahmen, gab es 61 Jungen und 58 Mädchen.

GRAFIK 6. Die Unterschiede zwischen Jungen und Mädchen.



Die Mädchen haben im Durchschnitt einigermaßen höhere Punkte fast in jeder Aufgabe und insgesamt im ganzen Test bekommen. Die Mittelwerte von Mädchen war 70,91 weil die Jungen durchschnittlich 62,61 Punkte bekamen. Wenn die Punkte einzelner Aufgaben betrachtet werden, haben die Mädchen durchschnittlich einen Punkt mehr für jede Aufgabe bekommen. Nur in der Aufgabe, in der die Schüler gefragt wurden, Behauptungen als 'falsch' oder 'richtig' zu markieren, haben die Jungen durchschnittlich etwas höhere Punkte bekommen.

#### *7.2.6 Qualifikation der Lehrer*

Alle Lehrer der Gruppen, die in dieser Untersuchung teilnahmen, waren qualifiziert. Alle hatten mindestens cum laude- Studien in Deutsch absolviert. Die meisten waren Stundenlehrer und die anderen waren Deutschlehrer, die ein festes Amt hatten.

Wenn die durchschnittlichen Gesamtpunkte der Gruppen verglichen werden, kann festgestellt werden, dass die Gruppen, deren Lehrer ein Stundenlehrer ist, etwas besseren Erfolg hatten (67,05) als die, deren Lehrer ein festes Amt hatten (65,35).

Die Gruppen, die von einem solchen Lehrer unterrichtet wurden, der laudatur in Deutsch absolviert hat, haben höhere Punkte (67,92) bekommen als die, deren Lehrer cum laude-Studien absolviert hatten (64,56).

### 7.2.7 Zahl der Unterrichtsstunden

In jeder A1-Gruppe, die an dieser Untersuchung teilnahm, wurde mit dem Deutschunterricht in der dritten Klasse angefangen. Die Lehrer dieser Gruppen haben mitgeteilt, dass diese Schüler 8 Stunden pro Woche jedes Jahr von der dritten bis in die sechste Klasse Deutsch gelernt haben.

Wie schon erwähnt, können die Schulen selbst entscheiden, wieviele Stunden die A2-Sprachen unterrichtet werden. In dieser Untersuchung gab es zwei A2-Gruppen aus solchen Schulen, in denen der Unterricht in der vierten Klasse begonnen hatte. In drei Schulen hat der Deutschunterricht in der fünften Klasse angefangen. In einer Schule, in der A2-Deutsch in der fünften Klasse begonnen hatte, wurde Deutsch 5-6 Stunden pro Woche unterrichtet. Die Schüler in anderen A2-Gruppen lernten Deutsch weniger als 4 Stunden pro Woche.

Um den Testserfolg der Gruppen zu vergleichen, wurde die Anzahl von wöchentlichen Unterrichtsstunden in jährliche Gesamtstunden verwandelt. Dies geschieht dadurch, dass die Anzahl von wöchentlichen Unterrichtsstunden mit den jährlichen Schulwochen (38 Wochen) multipliziert wird (Lahtinen, Lankinen, Penttilä & Sulonen 1999, 106). Die jährlichen Gesamtstunden wurden mit der Zahl der Jahre multipliziert.

Die Schüler, die Deutsch als A1-Sprache lernten, hatten Deutsch vier Jahre 8 Stunden pro Woche studiert. Am Ende der sechsten Klasse hatten alle A1-Lernende Deutsch ungefähr 1200 Stunden gelernt.

Die Unterschiede in Unterrichtsstunden unter A2-Lernenden waren viel größer. Ein Teil dieser Schüler wurde nur insgesamt 200-300 Stunden in Deutsch unterrichtet. Die anderen Gruppen wurden dagegen bis Ende der sechsten Klasse zwischen 450 und 600 Stunden unterrichtet.

Es stellte sich voraus, dass die durchschnittlichen Punkte der Schüler, die zwischen 450 und 600 Stunden unterrichtet wurden, durchschnittlich höher (69,84) waren. Dagegen die Punkte von diejenigen, die am wenigsten unterrichtet wurden (weniger als 300 Stunden in vier Jahren) und den Schülern die am meisten (ungefähr 1200 Stunden) unterrichtet wurden, fast auf dem selben, etwas niedrigeren Niveau waren (65,35 und 64,90). Diese Unterschiede sind in der Grafik 7 zu sehen.

GRAFIK 7. Einfluss der Unterrichtsstunden auf die Durchschnittswerte.

Zahl der Unterrichtsstunden	Durchschnittswert	Zahl der Schüler
Unter 300	65,35	34
300 - 600	69,84	37
Über 1200	64,90	48
Durchschnittswert aller Schüler	66,56	119

Zusammenfassend kann festgestellt werden, dass alle A1-Lerner und die A2-Lerner, die nur wenig Unterrichtsstunden hatten, den schlechtesten Erfolg hatten. Die A2-Lerner, die insgesamt 450-600 Stunden Deutsch gelernt hatten, bekamen im Durchschnitt die besten Ergebnisse. Es ist auch zu bemerken, dass unter den A2-Lernern, die am wenigsten (weniger als 300 Stunden) unterrichtet worden waren, die Gruppe gab, die den besten Erfolg von allen Gruppen hatte und die am wenigsten Deutschunterricht gehabt hatte.

## **8. Interpretation der Testergebnisse**

In diesem Kapitel wird versucht, die Unterschiede der Testergebnisse zu erklären und zu interpretieren. Die Mittelwerte der Schüler werden hier verglichen. Dabei ist zu erinnern, dass die Mittelwerte nicht die ganze Wahrheit aufdecken. In jeder Gruppe, die aufgrund des Geschlechts, der Größe usw. in Kategorien eingeordnet wurde, gibt es gute und schwache Ergebnisse.

### **8.1 Umfang der gelernten Sprache und Zahl der Unterrichtsstunden**

Keine großen Unterschiede zwischen Schülern mit A1- bzw. A2-Deutsch waren festzustellen. Die Hypothese, dass die A1-Lerner besseren Erfolg im Test hätten, hat sich nicht bestätigt. Wie oben angeführt wurde, haben die Schüler mit A1-Deutsch in diesem Test nicht höhere Punkte bekommen, sondern die durchschnittlichen Punkte der Schüler mit A2-Deutsch waren etwas höher. Die einzelnen Aufgaben haben auch keine systematischen Unterschiede herausgebracht. Obwohl die Lehrer meinten, dass die A1-Lerner mehr die deutsche Sprache gehört haben und dadurch die gesprochene Sprache ihrer Meinung nach besser verstehen, haben die A1-Lerner keinen besseren Erfolg in der Hörverständnisaufgabe.

Wie schon früher erwähnt wurde, hat ein Lehrer festgestellt, dass die Schüler mit A2-Deutsch unmotiviert sind, die Sprache zu studieren, weil die A2-Sprache fakultativ ist. Die Noten haben keine Bedeutung in Hinsicht der Weiterbildung und die Schüler können mit dem Lernen einer fakultativen Sprache zu jeder beliebigen Zeit abbrechen. Die Schüler, die Deutsch als A1-Sprache studieren, müssen dagegen den ganzen Lehrgang absolvieren und die, die in der gymnasialen Oberstufe

weiterstudieren, müssen in den Abprüfungen die obligatorische A1-Sprache absolvieren. Aus diesen Gründen wäre zu vermuten, dass die Lerner der A1-Deutsch vielleicht mehr motiviert sind als die Schüler mit A2-Deutsch.

In dieser Untersuchung haben die Lerner der A2-Deutsch jedoch besseren Erfolg gehabt. Es ist zu behaupten, dass die Tatsache, dass die A2-Sprache fakultativ ist, auch positive Einwirkungen auf den Testserfolg haben könnte. Die Schüler, die schon in der Unterstufe eine zweite, fakultative Fremdsprache wählen, sind vermutlich an Fremdsprachen interessiert, sprachlich begabt und dadurch motiviert, die Sprache zu lernen und deswegen guten Erfolg in den Sprachstudien haben.

Die Unterschiede zwischen A1-Lernern und A2-Lernern war jedoch sehr klein. Die Mittelwerte der Gesamtpunkte bei A1-Lerner war 64,90 und bei A2-Lerner 67,69. Es fragt sich, ob der Unterschied als statistisch relevant betrachtet werden kann.

In jeder A1-Gruppe wurde mit Deutschunterricht in der dritten Klasse angefangen und die Zahl der Unterrichtsstunden insgesamt war gleich. Die Unterschiede in Unterrichtstundenzahl unter den A2-Lernenden waren viel größer.

Es stellte sich heraus, dass die durchschnittlichen Punkte der Schüler, die in der Mittelkategorie waren, durchschnittlich höher waren als die Punkte anderer Schüler. Dagegen die Punkte der Schüler, die am wenigsten unterrichtet wurden und der Schüler, die am meisten unterrichtet wurden, auf einem niedrigerem Niveau waren.

Zusammenfassend könnte behauptet werden, dass je weniger Unterrichtsstunden es gibt, desto effizienter die Stunden verwendet

werden. Der Lehrer und die Schüler sind daran gewöhnt, dass im Sprachunterricht schnelle Fortschritte gemacht werden müssen. Dies ist vermutlich ein Grund dafür, dass die A2-Schüler mit 450-600 Unterrichtsstunden besseren Erfolg als die A1-Schüler hatten, die viel mehr unterrichtet wurden. Es ist auch zu bemerken, dass das Lernen der A1-Sprache gar nicht so intensiv ist, sondern teilt sich in einer längeren Periode als bei A2-Lernern. Besonders in den Unterstufen, in denen der Unterricht kursförmig ist, kann es eine längere Zeit zwischen Sprachkursen geben und die Schüler das schon Gelernte leicht vergessen.

Zu wenig Unterrichtsstunden sind jedoch zu wenig. Es ist zu vermuten, dass die A2-Lerner, die nur sehr wenig unterrichtet worden waren, nicht genug Zeit gehabt hatten, sich die Elementarkenntnisse in Fach Deutsch anzueignen. Der Mittelwert dieser Schüler wäre sogar niedriger gewesen, wenn die Gruppe, die den besten Erfolg von allen Gruppen hatte, nicht in dieser Kategorie gewesen wäre.

## **8.2 Größe der Schule und der Lerngruppe**

Wenn die statistischen Ergebnisse betrachtet werden, könnte behauptet werden, dass die Schüler in großen Schulen Deutsch besser lernen. Weil der Querschnitt relativ klein war und es im Querschnitt zusätzlich zu den mittelgroßen Schulen nur eine größere und eine kleinere Schule gab, können die Ergebnisse jedoch nicht generalisiert werden. Den Testserfolg haben vermutlich auch andere Faktoren als die Größe der Schule beeinflusst.

Im allgemeinen wird vermutet, dass Schüler in einer kleinen Gruppe effizienter lernen als in einer größeren Gruppe. Deswegen war es

erwartungsgemäß, dass der Testerfolg in kleinen Gruppen mit unter 12 Schülern besser wäre als in mittelgroßen Gruppen mit 12-15 Schülern. Somit ist überraschend, dass die Ergebnisse am besten in Gruppen mit über 15 Schülern waren.

Der Querschnitt war jedoch zu klein, um valide Schlussfolgerungen zu ziehen. Die Unterschiede in den Größen der Lerngruppe zwischen den drei Kategorien waren auch nicht groß – der Zweck dieser Klassifizierung war, große und kleine Gruppen von allen Gruppen zu unterscheiden. Deshalb ist festzustellen, dass in dieser Untersuchung die Größe der Lerngruppe kein entscheidender Faktor war.

### **8.3 Bedeutung der demografischen Faktoren**

Den besten Erfolg im Test hatten die Schüler in großen Gemeinden in Süd-Finnland. Dass es eine Korrelation zwischen diesen Variablen gibt, kann nicht gewiesen werden, weil es im Querschnitt keine großen Gemeinden von anderen Teilen Finnlands gab.

Die Sprachkenntnisse der Schüler variieren regional. Pasanen (1994, 16 – 17) hat festgestellt, dass es in den Testergebnissen von fremden Sprachen deutliche regionale Unterschiede gibt. Wenn Finnland in vier Teile eingeteilt wird, sind die Lernergebnisse durchschnittlich schwächer in den östlichen und nördlichen Teilen als in anderen. Der Unterschied zu der besten Region ist relativ groß.

Laut Pasanen (1994) sind die Schüler also nicht gleichgestellt, wenn es um Fremdsprachlernung geht. Die Ressourcen einer Schule hängen von der Ressourcen der Gemeinde ab und in den östlichen und nördlichen Teilen Finnlands gibt es viele minder bemittelte Gemeinden. Die

ausgeschriebenen Fremdsprachenlehrer sind in den ländlichen Gemeinden nicht verlockend. Wegen längerer Entfernungen nehmen die Lehrer vielleicht nicht gern an Ausbildungsveranstaltungen teil. Die Eltern von Schülern finden die Fremdsprachenkenntnisse vielleicht nicht wichtig oder haben keine Zeit oder Möglichkeiten, die Fremdsprachlernung ihrer Kinder zu stützen. Die Schüler haben nicht viele Möglichkeiten, die fremden Sprachen außer dem Unterricht anzuwenden und sind deshalb vielleicht nicht motiviert, fremde Sprachen zu lernen. Auch ein langer Schulweg kann Erschöpfung verursachen. (Pasanen 1994, 17.)

Was in dieser Untersuchung nicht gemessen wurde, ist der Unterschied zwischen städtischen und ländlichen Schulen. Pasanen (1994) bringt nämlich auch eine andere demografische Ungleichheit in Fremdsprachlernung vor. In Städten sind die Lernergebnisse besser als in ländlichen Schulen. Die Gründe sind vermutlich dieselben als die, die schon oben erwähnt wurden. Die Städte bieten viel mehr Anregungen und Interaktion für die Sprachlernung und Sprachverwendung in Praxis. (Pasanen 1994, 17.)

#### **8.4 Geschlecht des Schülers**

Die Tatsache, dass die Mädchen im Test etwas besseren Erfolg als die Jungen hatten, ist keine Überraschung, sondern das Phänomen ist laut Karppinen (1996) bekannt von vielen Untersuchungen, in denen Fremdsprachenkompetenz bewertet worden ist. Der Unterschied hängt vielleicht damit zusammen, dass die Mädchen fleissiger, gewissenhafter und mehr Interesse an Fremdsprachen haben als die Jungen. (Karppinen 1996, 19.)

Pasanen (1994) meint, dass beim Planen der Lehrpläne die Jungen besonders berücksichtigt werden sollen. Die Jungen sollten die Fremdsprachlernung nützlich und zweckmäßig finden. Dabei könnten die neuen technischen Medien helfen. Pasanen schlägt vor, dass andere Schulfächer, für die die Jungen sich interessieren, in Fremdsprachen unterrichtet werden könnten. (Pasanen 1994, 17.) Der neue kursförmige Unterricht bietet sicherlich neue Möglichkeiten an.

### **8.5 Rolle des Lehrers**

Alle Lehrer der Gruppen, die an dieser Untersuchung teilnahmen, waren qualifiziert und hatten mindestens cum laude-Studien im Fach Germanistik absolviert. Deswegen kann kaum behauptet werden, dass die Unterschiede zwischen den Schülern von der Stellung oder Studien der Lehrer abhingen. Die Unterschiede waren auch nicht statistisch relevant. Es war jedoch wichtig, nach der Qualifikation der Lehrer im Fragebogen zu fragen. In manchen kleinen Unterstufen werden A2-Sprachen von unqualifizierten Lehrern unterrichtet. Wenn es unter den Lehrern unqualifizierte Lehrer gegeben hätte, hätte überprüft werden können, ob sich dies auf den Erfolg der Schüler ausgewirkt hätte.

Wie schon früher in dieser Arbeit festgestellt wurde, hat der Lehrer oft eine große Wirkung darauf, wie die Schüler den Sprachunterricht finden. Dies beeinflusst natürlich die Sprachlernung. Beispielsweise die Persönlichkeit und die pädagogischen Fähigkeiten der Lehrer wirken sehr darauf, wie die Schüler die Sprache lernen.

## 9. Validität der Untersuchung

In diesem Kapitel werden die Validität der Untersuchung und die Reliabilität der Untersuchungsmittel abgeschätzt.

Wenn die Validität beim Untersuchungsverfahren betrachtet wird, sollen die vorher besprochenen Definitionen berücksichtigt werden. Es soll also die Validität der Testinterpretation geschätzt werden. Nicht nur die Validität der Untersuchungsmittel bzw. Testverfahren ist relevant, sondern vorzugsweise die Validität der Schlussfolgerungen, die aufgrund der Testergebnisse gezogen werden.

Die Validität dieser Untersuchung hängt davon ab, wie zuverlässig die Testergebnisse der Untersuchungsprüfung waren und wie die Ergebnisse interpretiert wurden. Es ist aber schwer zu schätzen, ein wie gelungener und zuverlässiger Indikator der Sprachkompetenz dieser Test war.

Die Schwächen des Testes wirken auf den Grad der Reliabilität und auf diese Weise auf die Reliabilität des ganzen Testes und dadurch auf die Validität der Testergebnisse. Es kann nur geschätzt werden, woraus die Unterschiede in den Testergebnissen resultieren. Die Punkte einzelner Schüler und Gruppen sind kein direkter Indikator der Sprachkompetenz, sondern werden von vielen Faktoren beeinflusst. Die Testergebnisse einzelner Schüler sind auch deswegen nicht ganz vergleichbar, weil es in der von Sukol erstellte Prüfung keine Instruktionen darüber gab, wie die Prüfung korrigiert werden soll, sondern die Lehrer haben die Prüfungen nach eigenem Ermessen benotet. Das war eine der größten Schwächen in der Prüfung von Sukol. Die Reliabilität der Prüfung wird im Folgenden besprochen.

## 9.1 Reliabilität der Untersuchungsprüfung

Die Lehrer wurden nach ihrer Meinung gefragt, ob die Testergebnisse der Schüler in einzelnen Aufgaben die Sprachfertigkeiten der Schüler entsprachen. Die Lehrer wurden gebeten, mit einer fünfstufigen Skala zu antworten. Die Meinungen waren sehr unterschiedlich und beim Antworten haben die Lehrer die ganze Skala verwendet. Keine Aufgabe wurde besonders misslungen oder gelungen gefunden. Von den Antworten dieser Frage können also keine Schlussfolgerungen gezogen werden.

Im allgemeinen meinten die Lehrer, dass die Aufgaben nicht ganz gelungen waren und deshalb die Sprachkompetenz nicht auf eine zuverlässige Weise testeten. Die Lehrer haben ihre Meinungen über die einzelnen Aufgaben frei geäußert.

### 9.1.1 Testen der schriftlichen Produktionsfähigkeit

Die Lückenfüllungsaufgabe wurde im allgemeinen für schwer gehalten. Die Aufgabe testet nicht nur das Schreibvermögen, sondern auch grammatische Kenntnisse und Wortschatz. Gleichzeitig ist die Aufgabe eine Leseverständnisübung. Die Aufgabe kann trotzdem als ein Beispiel für eine gesteuerte Produktionsaufgabe angesehen werden. Was hier allerdings fehlt, ist ein Kontext und Motivierung. Der Übungstyp war neu für die Schüler. Dies könnte Schwierigkeiten verursachen. Wie früher erwähnt, sollten die Übungstypen in einer Prüfung für die Schüler schon vorher bekannt sein.

Das Bild in der Aufsatzaufgabe war nicht besonders vielfältig und die Schüler hatten Schwierigkeiten, auszudenken, was sie schreiben könnten. Die Aufgabe war vermutlich leichter für diejenigen mit einer

reichen Phantasie. In dieser Prüfung wurde die Schreibfähigkeit mit einer Aufgabe getestet, was nach Weir (siehe Kapitel 5.1) nicht genügend ist, um die Schreibfähigkeit zuverlässig zu schätzen. Das Thema in der Übung war sehr allgemein und alltäglich. Deshalb ist es vorstellbar, dass die Schüler in diesem Sinn jedoch gleichgestellt waren.

Die Aufgabenbeschreibung war so allgemein, dass die Schüler sehr unterschiedliche Texte geschrieben haben. Der eine hat kurze, fehlerfreie Sätze geschrieben und gute Punkte bekommen. Der andere hat versucht, eine richtige längere Geschichte zu schreiben und hat deshalb mehrere Fehler gemacht, weil der Text länger war und hat dadurch schlechte Punkte bekommen. Wenn z.B. die Adjektivdeklination nicht gelehrt worden ist, macht der Schreiber, der Adjektive in seinem Text verwendet, viele Fehler. In der Aufgabenbeschreibung hat der Verfasser nicht versucht, die Schüler zu motivieren oder den Text in einen Kontext einzubetten. Manche Schüler haben nicht gewusst, worüber und was sie schreiben sollten. Mit guten und deutlichen Instruktionen könnten die Schüler motiviert werden und beraten werden, was sie schreiben sollten. Die Aufgabe war also kein typischer freier Aufsatz, weil die Schüler über ein Bild schrieben und weil es viele Beispielsätze gab.

Nach Huhta und Takala (siehe Kapitel 5.1.1) ermöglicht ein freier Aufsatz oft eine vielseitige Abschätzung der Schreibfähigkeit. In diesem Fall war der Aufsatz jedoch so kurz, dass die Schreibfähigkeit vermutlich nicht zuverlässig zu bewerten war, weil das sprachliche Niveau des Textes von der Phantasie des Schreibers abhängt, nicht nur von den sprachlichen Fähigkeiten.

### 9.1.2 Testen der rezeptiven Sprachfähigkeit

Ein Lehrer kritisierte, dass eine Leseverständnisaufgabe auch andere Fähigkeiten als nur das Verstehen testet. Es ist jedoch wichtig, dass die Schüler, um in der jeweiligen fremden Sprache kommunizieren zu können, auch andere Fähigkeiten als nur Fremdsprachenkenntnisse ausnützen können. Sprachkompetenz bedeutet ja nicht nur Beherrschung der Lexik und grammatischer Strukturen sondern die Fähigkeit, die Sprache passend in der jeweiligen Kommunikationssituation zu benutzen.

Die Lehrer meinten, dass die Aufgabe mit richtigen und falschen Behauptungen nicht nur die Sprachkompetenz testete, sondern auch Allgemeinwissen. Es ist zu vermuten, dass die Aufgabe deshalb kein zuverlässiger Indikator der Sprachkenntnisse war. Einige Lehrer stellten fest, dass einige schwache Schüler für diese Aufgabe höhere Punkte bekommen haben als ihre Gruppe durchschnittlich – in einer Aufgabe auf diese Art können die richtigen Antworten geraten werden.

Der Text in der Hörverständnisprüfung war ein traditioneller, narrativer Text. Die Schüler haben erst den ganzen Text einmal gehört und danach mit Pausen. Die Aufgabe versuchte also keine authentische Kommunikationssituation nachzuahmen. Das Thema Snowboarden war modern und für alle bekannt. Einige können darüber mehr wissen als andere, aber es ist kaum möglich, die Fragen aufgrund des Allgemeinwissens zu beantworten, ohne den Text zu hören.

## 9.2 Fragebogen als Informationsquelle

Das Verfassen einer guten Umfrage ist zeitraubend und anspruchsvoll (Hirsjärvi, Remes & Sajavaara 1997, 191). Eine Überraschung beim Sammeln des Materials war, wieviele Ressourcen eine Umfrage braucht. Es ist nicht einfach, eine vernünftige Umfrage durchzuführen und das Verfassen eines Fragebogens nimmt viel Zeit. Eine Umfrage durch die Post ist auch teuer.

Der Fragebogen und die Umfrage hat sich jedoch als ziemlich gelungen erwiesen. Mit Hilfe der Fragebögen wurden wichtige und wesentliche Informationen gesammelt, einschließlich eigene Erfahrungen und Meinungen der Lehrer. Weil die Zahl der eingefüllten Fragebögen nicht besonders hoch war, konnten sie sorgfältig ausgewertet werden.

Es ist aber misslungen, statistisch wertvolle Informationen herauszufinden. Wie Hirsjärvi, Remes und Sajavaara (1997) feststellen, bringen die Umfragen nur selten theoretisch wichtige Informationen. Wie sorgfältig und ehrlich die Adressaten die Fragen beantwortet haben, kann nur geschätzt werden. Es ist schwer zu wissen, wie gelungen die Fragen oder die gegebenen Antwortsalternativen sind oder ob die Adressaten überhaupt den Bereich beherrschen, über den die Fragen gestellt werden. (Hirsjärvi, Remes & Sajavaara 1997, 191.)

Es ist üblich, dass bei einer Umfrage – besonders wenn die Umfrage durch die Post abgewickelt wird – der größte Teil der Adressaten nicht antwortet. Wenn die Fragebögen durch eine Institution, Organisation oder Gemeinschaft verteilt werden, ist die Anzahl der Teilnehmer etwas höher. In einer kontrollierten Umfrage verteilt der Forscher die Fragebogen persönlich bzw. holt sie persönlich ab. (Hirsjärvi, Remes & Sajavaara 1997, 192 – 193.) In dieser Untersuchung war es jedoch nicht möglich, einen persönlichen Kontakt aufzunehmen.

### 9.3 Forschungsgegenstand

Als die Untersuchung geplant wurde, wurde ein möglichst repräsentatives Untersuchungsmaterial angestrebt. Das Ziel war, eine Probe mit 3–5 % von allen Schülern, die Deutsch als A1- bzw. A2-Sprache lernen, zu sammeln. Dies hätte ungefähr 100 A1-Lerner und 200 A2-Lerner bedeutet (Kimmoke-indikatorit 1999). Als die eigentlichen Vorbereitungen für die Befragung angingen, wurde klar, dass diese Zielsetzung unrealistisch war, weil die Ressourcen begrenzt waren. Die endgültige Probe stand aus 48 A1-Lerner und 71 A2-Lerner. Das war knapp die Hälfte von der ursprünglichen Zielsetzung.

Der Umfang des Untersuchungsmaterials ist befriedigend. Die geografische Repräsentanz litt jedoch darunter, dass der größte Teil der Schulen sich in den großen Gemeinden in Süd-Finnland befanden.

Von den 36 Schulen des ursprünglichen Forschungsgegenstandes nahm ein Viertel teil. Es ist zu vermuten, dass der primäre Grund für das Ausbleiben der Zeitpunkt dieser Untersuchung war. Der ursprüngliche Zweck war, das Untersuchungsmaterial Ende März zu sammeln. Die Lehrer wurden aber wegen des Zeitdrucks der Kontaktpersonen in Sukol ry erst Ende März informiert und gebeten, die Prüfung in ihren Gruppen im April - der ein sehr hektischer Monat in Schulen ist - schreiben zu lassen. Um eine größere Probe zu erreichen, hätte die Umfrage und Materialsammlung mindestens einen Monat früher im Frühjahr durchgeführt werden müssen.

Es muss aber berücksichtigt werden, dass besonders für die A2-Lernende sogar ein Monat eine längere Zeit ist, in der sie viel Neues lernen können. A2-Sprache wird ja in der Unterstufe nur zwei oder drei Jahre gelernt und ein Schuljahr ist in Praxis etw. 9 ½ Monate. Davon sind ein paar Monate eine lange Zeit. Wenn die Sprachkompetenz schon

im Spätwinter getestet würde, würde der Test nicht die Deutschkenntnisse gerade am Ende der Unterstufe messen.

Ein zweiter bedeutsamer Grund für das Ausbleiben der Schulen war vermutlich die Weise, auf die die Schulen der Zielgruppe gewählt wurden. Wie früher erwähnt, wurde die Zielgruppe unter diesen Schulen gewählt, die die Prüfung von Sukol im Jahr 2001 bestellt hatten. Die Deutschlehrer dieser Schulen hatten also vor, die neue Prüfung in ihren Deutschgruppen in der sechsten Klasse schreiben zu lassen. Die Lehrer, die an der Untersuchung teilnahmen, mussten zwei Prüfungen in einer kurzen Zeit schreiben lassen, welches viel Zeit nimmt. Diese Sache wurde aber schon vorher berücksichtigt. Im ersten Brief, der den Lehrern geschickt wurde, wurde empfohlen, dass diese Untersuchungsprüfung eine Übung für die eigentliche Prüfung des Jahres 2001 wäre.

## **10. Schlussbetrachtung**

Sprachkompetenz und Leistungsmessung sind vielfältige Forschungsgebiete der Linguistik. In dieser Arbeit wurden einige Theorien über Sprachkompetenz vorgestellt und die wichtigsten Funktionen und Gütekriterien der Leistungsmessung behandelt.

Sprachkenntnisse sind eine abstrakte Fähigkeit, die nur teilweise eingeschätzt werden kann. Deshalb ist das Ausarbeiten eines gültigen, passenden und zuverlässigen Tests nicht einfach. Es scheint unwahrscheinlich, dass ein zuverlässiges und umfassendes Bild über die Sprachkompetenz zu bekommen ist, wenn nur eine Seite der Sprachkompetenz gemessen wird. Deswegen ist notwendig, dass der Test möglichst vielfältig alle Teile der Sprachkompetenz testet. Die

Testaufgaben sollen sorgfältig ausgearbeitet werden und passend für das jeweilige Niveau des Lernalers sein. Zusätzlich soll immer ein Kompromiss zwischen Forderungen an die Qualität und praktischen Beschränkungen gefunden werden.

In dieser Untersuchung wurde die Sprachkompetenz nur mit Hilfe eines schriftlichen Tests geprüft. Deshalb ist zu bemerken, dass die Sprachkenntnisse nicht vollständig gemessen wurden. Leider konnte die mündliche Sprachkompetenz in dieser Untersuchung wegen kleiner Ressourcen nicht beurteilt werden.

Der Querschnitt in dieser Untersuchung war nicht groß genug, damit die Ergebnisse generalisiert werden könnten. Einige Tendenzen können jedoch festgestellt werden. Die Ergebnisse der Untersuchung waren überraschend und die Hypothesen haben sich nicht als richtig erwiesen. Es stellte sich heraus, dass die Lernaler der A1-Deutsch nicht unbedingt bessere Kenntnisse in Deutsch am Ende der Unterstufe haben, obwohl sie viel mehr Deutschunterricht haben. Die Schüler mit A1-Deutsch hatten in dem Test, der in dieser Untersuchung verwendet wurde, keinen besseren Erfolg als die Schüler, die Deutsch als A2-Sprache lernen. Der Einfluss verschiedener Hintergrundfaktoren war auch nicht so eindeutig, wie gedacht wurde.

Zusammenfassend könnte festgestellt werden, dass die Testperformanz einzelner Schüler und Gruppen nicht unbedingt von den Faktoren beeinflusst wird, die in dieser Untersuchung geprüft wurden. Es ist zu vermuten, dass der Erfolg mehr mit individuellen Unterschieden in der sprachlichen Kompetenz und mit den pädagogischen Fähigkeiten des Sprachlehrers zusammenhängen.

Um die Sprachkenntnisse bei Schülern mit A1- und A2-Deutsch zuverlässig vergleichen und die Ergebnisse generalisieren zu können,

sollte der Querschnitt größer sein. Am praktischsten wäre es im Zusammenhang mit einer jährlichen landesweiten Prüfung. Um das Material für diese Untersuchung zu sammeln, war es nötig, eine extra Prüfung nur für den Zweck dieser Untersuchung zu verwenden, weil es keine andere passende Prüfung gab. Die Prüfung müsste sorgfältig abgefasst werden, damit es keine Unklarheiten oder Missverständnisse gäbe. Es wäre gut, wenn eine Pionierprüfung mit einigen Gruppen getestet werden und die Lehrer nach ihren Meinungen gefragt werden könnten. Die Schwächen und Problemfälle würden auf diese Weise entdeckt und der Grad der Reliabilität könnte erhöht werden.

## Literaturverzeichnis

- Bachman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Butzkamm, Wolfgang 1989. *Psycholinguistik des Fremdsprachenunterrichts: natürliche Künstlichkeit: von der Muttersprache zur Fremdsprache*. Tübingen: Francke.
- Canale, Michael & Swain, Merrill 1980. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. In: *Applied Linguistics*, 1,1, S. 1 – 47.
- Carroll, John B. 1962. The Prediction of Success in Intensive Foreign Language Training. In: *Training Research and Education*. Hg. Robert Glaser. Pittsburg: University of Pittsburgh Press. S. 87 – 136.
- Doyé, Peter 1992. *Typologie der Testaufgaben für den Unterricht Deutsch als Fremdsprache*. 5. Aufl. Berlin: Langenscheidt.
- Doyé, Peter 1995. Funktionen und Formen der Leistungsmessung. In: *Handbuch Fremdsprachenunterricht*. Hg. Karl-Richard Bausch, Herbert Christ und Hans-Jürgen Krumm. 3., überarb. und erw. Aufl. Tübingen: A. Franke Verlag. S. 277 – 282.
- Halme, Anna 2000. Oppilaan arviointi perusopetuksessa. In: *Tempus* 36, 1, S. 18 – 19.
- Helvelahti, Juha 1996. 9. Luokan valtakunnallinen koe 1995: laatijan palautetta. In: *Tempus* 31, 1, S.19 – 20.

Hirsjärvi, Sirkka & Remes, Pirkko & Sajavaara Paula 1997. *Tutki ja kirjoita*. Helsinki: Kirjayhtymä Oy.

Huhta, Ari & Takala, Sauli 1999. Kielitaidon arviointi. In: *Kielenoppimisen kysymyksiä*. Hg. K. Sajavaara ja A. Piirainen-Marsh. Soveltavan kielentutkimuksen keskus, Jyväskylän yliopisto. S. 179 – 228.

Hüllen, Werner & Lothar, Jung 1979. *Sprachstruktur und Spracherwerb*. Düsseldorf: August Bagel Verlag.

Jaakkola, Hanna 1997. *Kielitieto kielitaitoon pyrittäessä*. Vieraiden kielten opettajien käsityksiä kielioopin oppimisesta ja opettamisesta. Jyväskylä Studies in Education, Psychology and Social Research 128. Jyväskylä University Printing House, Jyväskylä. Diss.

Karlsson, Fred 1994. *Yleinen kielitiede*. Helsinki: yliopistopaino.

Karppinen, Marjatta 1996. Ruotsia osataan mukavasti. *Tempus* 31, 2. S. 18 – 19.

Kelly, Louis G. 1969. *25 Centuries of Language Teaching*. 2. Aufl. Rowley, Massachusetts: Newbury House.

Kimmoke-indikaattorit, 1999. *Tilastotietoa kieltenopetuksesta perusopetuksessa ja lukiossa*. Moniste 31/1999. Opetushallitus. <http://www.sukol.fi/yleis.html> 24.1.2001.

Koulusäädökset 1999. Lakikokoelma. Helsinki: Edita.

Lado, Robert 1961. *Language Testing: the Construction and Use of Foreign Language Tests*. London: Longmans Green and co Ltd.

- Lahtinen, Matti, Lankinen, Timo, Penttilä, Antero & Sulonen Arto, 1999. *Koulutuksen lainsäädäntö käytännössä*. Juva: WSOY.
- Larsen-Freeman, Diane & Long, Michael H. 1991. *An Introduction to Second Language Acquisition Research*. New York: Longman.
- Merten, Stephan 1997. *Wie man Sprache(n) lernt: eine Einführung in die Grundlagen der Erst- und Zweitsprachenerwerbsforschung mit Beispielen für das Unterrichtsfach Deutsch*. Frankfurt am Main: Peter Lang.
- Pasanen, Ulla-Maija 1994. 9. Luokan yhteinen koe. *Tempus* 29, 1. S. 16 – 18.
- Peruskoulun opetussuunnitelman perusteet 1994. Helsinki: Painatuskeskus. Opetushallitus.
- Tuokko, Eeva 2000. A1-englannin kansallinen arviointi. *Tempus* 35, 4. S. 17.
- Vollmer, Helmut J. 1982. *Spracherwerb und Sprachbeherrschung: Untersuchungen zur Struktur von Fremdsprachenfähigkeit*. Tübingen: Gunter Narr Verlag.
- Vollmer, Helmut J. 1995. Leistungsmessung: Überblick. In: *Handbuch Fremdsprachenunterricht*. Hg. Karl-Richard Bausch, Herbert Christ und Hans-Jürgen Krumm. 3., überarb. und erw. Aufl. Tübingen: A. Franke Verlag. S. 273 – 277.
- Weir, Cyril 1990. *Communicative Language Testing*. New York: Prentice Hall.

Weir, Cyril 1993. *Understanding and Developing Language Tests*. New York: Prentice Hall.

**Peruskoulun  
6. luokan  
koe 2000**

**A2/A1-saksa**

---

**Sukol-Palvelu Oy  
Suomen saksanopettajat ry**

## Peruskoulun 6. luokkien A2/A1-saksan koe 2000

*Koe suositellaan pidettäväksi 1.5. alkavalla viikolla 18.*

Koeaika on 60-90 minuuttia opetusryhmästä ja siitä riippuen, missä laajuudessa opettaja haluaa kokeen pitää.

Kokeen osat:

Kuullun ymmärtäminen	20p	(Äänitteen pituus 12')
Luetun ymmärtäminen	20p	
Oikein-väärin -tehtävä	10p	
Yhdistä kysymys ja vastaus	5p	
TV-tarjotin	10p	
Rakennetehtävä	15p	
Kirjoitelma	20p	/ Koko koe yhteensä 100p

### Koemateriaali

Oppilaan materiaali käsittää luetun ymmärtämisen tekstin, kuullun ja luetun vastauslomakkeet sekä muut tehtävät. Opettajan materiaali käsittää opettajan ohjeet, kuullun ymmärtämisen tekstin (ja äänitteen), ratkaisut ja palautesivun.

### Kokeen suoritusohjeet

Koeaika on 60-90 minuuttia. Ennen koetta oppilaille selvitetään kokeen suoritustapa. Opettajan on syytä tutustua kokeeseen ennakolta. Hän voi antaa suomeksi joitakin sanoja, jos katsoo sen tarpeelliseksi.

### Kuullun ymmärtämisen koe

Kuullun ymmärtämisen koe tehdään joko äänitteen kanssa (A) tai ilman äänitettä (B).

*Oppilaat lukevat kysymykset läpi ennen ensimmäistä kuuntelukertaa.*

- A) Sitten kuullun ymmärtämisen teksti kuunnellaan kasetilta seuraavasti: ensin koko teksti kuullaan ilman taukoja, sen jälkeen teksti tulee tauotettuna ja taukojen aikana oppilaat vastaavat kysymyksiin. Kasetti on käytettävissä kaksipuolisesti eli sen molemmilla puolilla on sama kuullun ymmärtämisen teksti.
- B) Vaihtoehtoisesti opettaja lukee tekstin ensin kokonaan ja toisen riittävästi tauotetun luennan aikana oppilaat vastaavat kysymyksiin.

*Jaetaan loput tehtävät suoritettaviksi.*

## Oikeat vastaukset

### Kuullun ymmärtäminen (yksi asia riittää)

1. 21. 2. Tytöt. 3. Punatukkainen / pitää farkkuja ja villapaitaa. 4. Rap. 5. 1998.  
6. Lentokoneita. 7. Keväällä. 8. Paljon rahaa. 9. Hyviä aikoja, huonoja aikoja. 10. 5 päivänä viikossa. 11. Ei. 12. Kreikasta. 13. Neljä vuotta vanhempi. 14. Helmikuussa. 15. Luminen.  
16. Matkalla sairaalasta kotiin. 17. Auringonpaistetta. 18. Kauniit silmät. 19. Terveyttä ihmisille. 20. Enemmän aikaa / lepoa / vapaa-aikaa.

20x1p=20p

### Luetun ymmärtäminen ( yksi asia riittää)

1. 10 vuodessa. 2. Koko maailman laajuista. 3. Itävallassa. 4. Tanssia / kuunnella musiikkia.  
5. Muotivaatteet. 6. Myssyjä, aurinkolaseja, kenkiä, käsineitä, housuja, takkeja. 7. Ei kirjavaa, rauhalliset värit. 8. Sveitsissä. 9. Myös aloittelijoille. 10. Nopeasti. 11. Toukokuussa lunta.  
12. Uskonto. 13. Vanhemmat. 14. 1965. 15. Poppenin lapset. 16. Talon takana. 17. Lauta, vuori ja lunta. 18. Hyppäävät ja lentävät. 19. Myy parhaita lumilautoja ja ansaitsee paljon.  
20. Hiihto ja laskettelu.

20x1p=20p

### Oikein-väärin: O/V

1. O 2. O 3. V 4. O 5. O 6. V 7. O 8. V 9. O 10. V

10x1p=10p

### Yhdistä kysymys ja vastaus

1. F  
2. E  
3. G  
4. C  
5. D

5x1p=5p

### TV - tarjotin

10x1p=10p

### Rakennetehtävä

1. Siehst du 2. auf dem Tisch 3. kommt 4. aus 5. haben - gemacht 6. in 7. Meine  
8. unserem 9. hundert 10. den 11. In der Schule 12. Da sind 13. nach 14. schöner  
15. mit

15x1p=15p

### Kirjoitelma kuvasta

Oppilaat laativat seitsemän lauseen kirjoitelman annetusta kuvasta.

20p

## KUULLUN YMMÄRTÄMINEN

*Kuunnellaan kuullun ymmärtämisen teksti kasetilta TAI opettajan lukemana ensin kokonaan ja sitten tauotettuna.*

### Alle lieben Oli P.

Oliver Petszokat ist 21 Jahre alt. Besonders die Mädchen lieben den Rapper Oli P. sehr. Heute hat er rote Haare und trägt gerne Jeans und Pulli. (1-3)

Oli P. sagt, er kann nicht singen, er rappt, spricht mit Musik. Im Jahre 1998 war sein erster großer Hit "Flugzeuge im Bauch". (4-6)

Der neue Frühlingshit heißt "So bist du". Der Frühling bringt ihm viel Geld, denn mit seiner Rappmusik hat er schon drei Millionen Platten verkauft. (7-8)

Oli P. hat dazu in der populären Fernsehserie "Gute Zeiten, schlechte Zeiten" mitgespielt. Fünf Tage in der Woche sehen 6,5 Millionen Menschen "Gute Zeiten, schlechte Zeiten" im Fernsehen. Das hat jetzt ein Ende für Oli P., denn er will etwas Neues machen. (9-11)

Oli P. ist jung, aber er hat schon eine Frau und ein Kind. Seine Frau Tati kommt aus Griechenland und sie ist vier Jahre älter als Oli P. (12-13)

Das Baby wurde im Februar 1999 geboren. Es hat geschneit, und überall war Schnee. (14-15)

Aber auf dem Weg vom Krankenhaus nach Hause hat die Sonne geschienen. Darum heißt der Junge Ilias, denn der Name Ilias ist das griechische Wort für Sonnenschein. (16-17)

Von der Mutter hat das Baby die Augen und den Mund, vom Vater die Nase und die Ohren. Oli P. meint, alle Menschen haben schöne Augen. (18)

Oli P. hat auch Wünsche. Er möchte so gern, dass alle Menschen auf der Welt gesund bleiben. Es gibt zu viele Krankheiten. Dazu wünscht er sich, dass er mehr Zeit hätte. Seine Tage sollten 36 Stunden haben. Er braucht auch Ruhe und Freizeit. (19-20)

# Tehtävä- ja vastauslomake

Nimi: \_\_\_\_\_

## Kuullun ymmärtäminen

*Vastaa kysymyksiin suomeksi.*

1. Kuinka vanha Oliver Petszokat on? \_\_\_\_\_
2. Ketkä rakastavat häntä? \_\_\_\_\_
3. Mitä ulkonäöstä sanotaan? \_\_\_\_\_  
\*\*\*\*\*
4. Mitä musiikinlajeja Oli P. edustaa? \_\_\_\_\_
5. Minä vuonna syntyi ensimmäinen hitti? \_\_\_\_\_
6. Mitä hittilevyn ”vatsassa” oli? \_\_\_\_\_  
\*\*\*\*\*
7. Minä vuodenaikana syntyi ”So bist du” ? \_\_\_\_\_
8. Mitä levy tuo tullessaan? \_\_\_\_\_  
\*\*\*\*\*
9. Mikä on suosituksen TV-sarjan nimi? \_\_\_\_\_
10. Kuinka usein sitä esitetään? \_\_\_\_\_
11. Jatkaako Oli P. sarjassa? \_\_\_\_\_  
\*\*\*\*\*
12. Mistä vaimo Tati on kotoisin? \_\_\_\_\_
13. Mitä Tatin iästä kerrotaan? \_\_\_\_\_  
\*\*\*\*\*
14. Missä kuussa vauva syntyi? \_\_\_\_\_
15. Millainen sää oli? \_\_\_\_\_  
\*\*\*\*\*
16. Missä vauvan nimi keksittiin? \_\_\_\_\_
17. Mitä Ilias tarkoittaa? \_\_\_\_\_  
\*\*\*\*\*
18. Mitä Oli P.:n mielestä kaikilla ihmisillä on? \_\_\_\_\_
19. Mikä toive hänellä on? \_\_\_\_\_
20. Mitä Oli P. toivoo lopuksi? \_\_\_\_\_  
\*\*\*\*\*

\_\_\_\_\_ /20p

## LUETUN YMMÄRTÄMINEN

### Auf dem Schneebrett: Snowboarden!

Das machen sie alle: "Marienhof"-Star Annika Murjahn, 21, Prinz Charles, 51, und Formel-1-Pilot Eddie Irvine, 34: Snowboarden. In den letzten zehn Jahren ist Snowboarden in der ganzen Welt zum Kult geworden. (1-2)

Ischgl in Österreich ist das Boarderparadies in Europa, denn da gibt es alles, auch Popkonzerte und Modeschauen. Abends kann man tanzen und Musik hören. Dazu muss man natürlich modisch und elegant gekleidet sein, denn tolle Kleidung ist wichtig. Man trägt "richtige" Mützen, Sonnenbrillen, Schuhe, Handschuhe, Hosen und Jacken. Die Wintermode des Jahres 2000 ist nicht bunt, die Farben sind eher ruhig. (3-7)

Davos in der Schweiz ist auch ein Treffpunkt für Snowboard-Anfänger, wo man alles von A bis Z lernen kann. Der Sport ist nicht schwer, man lernt ihn schnell. (8-10)

Garmisch-Partenkirchen in Deutschland ist noch im Mai schneesicher. In der BRD gibt es in diesem Winter 250 000 Schneebrettler, denn Snowboarden ist fast zur Religion geworden. Zuerst haben fast nur junge Leute diesen Sport getrieben, aber jetzt kommen die Eltern, die den Modesport erlernen wollen. (11-13)

Genau vor fünfunddreißig Jahren ist Sherman Poppen in Michigan in den USA auf die Idee gekommen, seine Kinder hinter dem Haus auf einem Monoski fahren zu lassen. Er hat das Brett selbst gebaut, und es ist bestens gefahren. Ein Brett, ein Berg und Schnee, das ist alles, was Snowboarder brauchen. Sie springen und fliegen. (14-18)

Poppens Landsmann Jake Burton verkauft heute die besten Schneebretter der Welt und verdient große Summen Geld damit. (19)

Seit 1998 ist der Snowboard-Sport olympisch. Skilaufen und Skifahren, die bleiben, aber der Trendsport ist Snowboarden. (20)

# Tehtävä- ja vastauslomake

Nimi: \_\_\_\_\_

## Luetun ymmärtäminen

*Vastaa kysymyksiin suomeksi.*

1. Missä ajassa lumilautailusta on tullut pop? \_\_\_\_\_
2. Kuinka laajaa suosio on? \_\_\_\_\_
3. Missä maassa Ischgl sijaitsee? \_\_\_\_\_
4. Mitä illalla voi tehdä? \_\_\_\_\_
5. Mikä on urheilun ohella tärkeää? \_\_\_\_\_
6. Mainitse esimerkkejä ”oikeista” asioista. \_\_\_\_\_
7. Mikä on trendi vuonna 2000? \_\_\_\_\_
8. Missä maassa Davos sijaitsee? \_\_\_\_\_
9. Miten se eroaa joistain muista paikoista? \_\_\_\_\_
10. Miten lumilautailussa yleensä edistytään? \_\_\_\_\_
11. Miksi Garmisch-Partenkirchen on hyvä? \_\_\_\_\_
12. Mikä lumilautailusta on melkein tullut? \_\_\_\_\_
13. Ketkä ovat nyt alkaneet harjoittaa lautailua? \_\_\_\_\_
14. Laske, minä vuonna lumilautailu keksittiin. \_\_\_\_\_
15. Ketkä olivat ensimmäiset kokeilijat? \_\_\_\_\_
16. Missä he yrittivät lautoilla Michiganissa? \_\_\_\_\_
17. Mitä tarvitaan lumilautailuun? \_\_\_\_\_
18. Miten lumilautailijat liikkuvat rinteessä? \_\_\_\_\_
19. Mitä Jake Burtonista kerrotaan? \_\_\_\_\_
20. Mitkä asiat tulevat vielä säilymään? \_\_\_\_\_

\_\_\_\_\_ / 20p

## Richtig - falsch / Oikein - väärin O / V

*Merkitse vastauslomakkeelle.*

1. In Finnland gibt es Seen.
2. Finnland ist größer als die Schweiz.
3. Inari liegt in Südfinnland.
4. Turku ist älter als Helsinki.
5. In Kuusamo gibt es viel Wald.
6. In Finnland geht man mit fünf Jahren in die Schule.
7. Im Februar oder März haben die finnischen Schüler Ferien.
8. Alle Finnen haben ein Sommerhaus.
9. Him ist populär in Deutschland.
10. Tiktak ist eine schwedische Band.

\_\_\_\_\_/10p

Yhdistä kysymys ja vastaus.

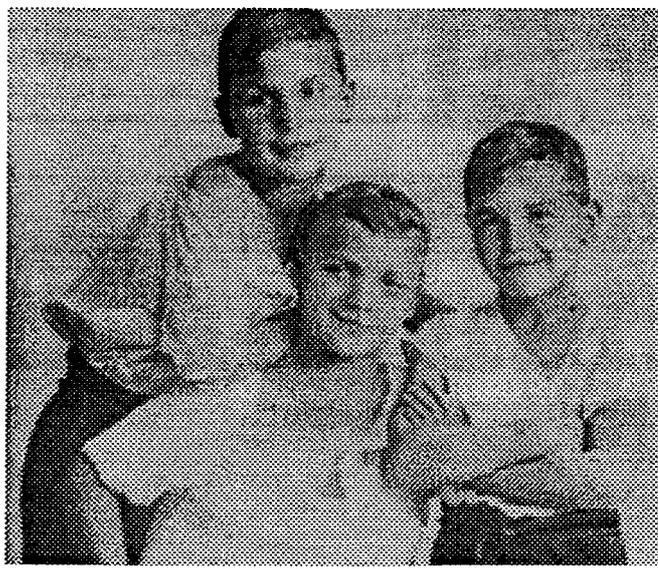
*Merkitse oikea kirjain vastauslomakkeelle ruutuun. 3 kohtaa jää yli.*

- |                             |                                     |
|-----------------------------|-------------------------------------|
| 1. Was möchtest du?         | A. Es ist ja eiskalt.               |
| 2. Woher kommt Johan?       | B. Der Tee schmeckt nicht.          |
| 3. Wie spät ist es?         | C. Am Freitag.                      |
| 4. Wann gehst du schwimmen? | D. Ich war in einem tollen Konzert. |
| 5. Warum bist du so froh?   | E. Aus Holland.                     |
|                             | F. Cola, bitte.                     |
|                             | G. Viertel nach elf.                |
|                             | H. Ich habe keine Zeit.             |

\_\_\_\_\_/5p

TV-tarjotin

*Aiot katsoa televisiota ja keräilet evästä. Luettele vastauslomakkeelle 2 juotavaa, 1 makeislaji, 3 hedelmää ja 4 voileipään liittyvää ruoka-ainetta.* \_\_\_\_\_/10p



## RAKENNETEHTÄVÄ

\_\_\_(1)\_\_\_ , hier ist ein altes Foto von den Nivea-Jungen \_\_\_(2)\_\_\_ ? Nivea \_\_\_(3)\_\_\_ wirklich \_\_\_(4)\_\_\_ Deutschland. Dr. Oscar Troplowitz hat 1911 die weiße Nivea-Creme erfunden: Aus Wasser, Vaseline, Glycerin, Rosen- und Maiglöckchenöl und Eucerit. Drei Jungen, Peter, Rolf-Robert und Wilhelm Wiethüchter \_\_\_(5)\_\_\_ Reklame für Nivea \_\_\_(5)\_\_\_ . Peter war damals zehn Jahre alt und wohnte \_\_\_(6)\_\_\_ Hamburg. Er erzählt: " \_\_\_(7)\_\_\_ Mutter hat uns Jungen fotografieren lassen und \_\_\_(8)\_\_\_ Vater das Bild zu Weihnachten gegeben. Das Bild war im Schaufenster des Fotografen und so hat man uns gefunden. Wir haben je \_\_\_(9)\_\_\_ Mark bekommen und eine Sportjacke. Das war alles für \_\_\_(10)\_\_\_ Job. \_\_\_(11)\_\_\_ haben viele gerufen: "\_\_\_(12)\_\_\_ die Nivea-Jungs!" Rolf-Robert ist hier geblieben, aber Wilhelm ist später \_\_\_(13)\_\_\_ Bolivien gegangen. Viele Menschen möchten ein bisschen \_\_\_(14)\_\_\_ werden und das geht auch \_\_\_(15)\_\_\_ Nivea."

1. Näetkös

2. pöydällä 3. tulee

4. -sta

5. ovat ... tehneet / perf.

6. -ssa

7. Minun

8. unser / datiiivin!

9. sata

10. der / akkusatiiviin! 11. Koulussa

12. Tuolla ovat

13. -aan / mihin?

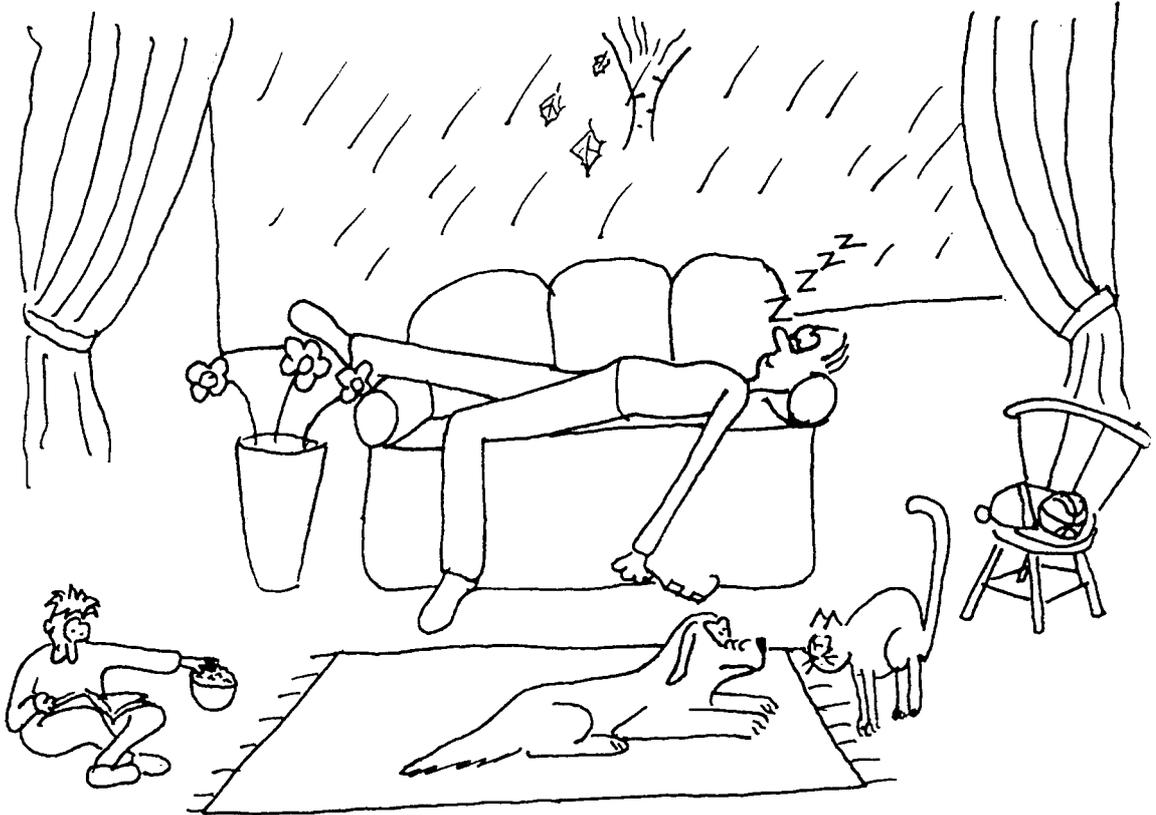
14. kauniimmiksi / komparatiivi

15. kanssa

Nimi: \_\_\_\_\_

## Kirjoitelma

Kirjoita 7 lausetta alla olevasta kuvasta. Esim.: On syksy ja sataa. On kylmä. Kotona on lämmintä ja mukavaa. Missä äiti on?! Jatka ...



ky: kalle

---



---



---



---



---



---



---



---

Kirjoitelma \_\_\_\_\_ / 20p

Nimi: \_\_\_\_\_

**Vastauslomake***Oikein - väärin: O / V*

1) \_\_\_\_\_ 2) \_\_\_\_\_ 3) \_\_\_\_\_ 4) \_\_\_\_\_ 5) \_\_\_\_\_ 6) \_\_\_\_\_ 7) \_\_\_\_\_ 8) \_\_\_\_\_ 9) \_\_\_\_\_ 10) \_\_\_\_\_

\_\_\_\_\_ /10p

*Yhdistä kysymys / vastaus:* 1.  2.  3.  4.  5. 

\_\_\_\_\_ / 5p

*TV-tarjotin:*

1. \_\_\_\_\_

6. \_\_\_\_\_

2. \_\_\_\_\_

7. \_\_\_\_\_

3. \_\_\_\_\_

8. \_\_\_\_\_

4. \_\_\_\_\_

9. \_\_\_\_\_

5. \_\_\_\_\_

10. \_\_\_\_\_

\_\_\_\_\_ /10p

*Rakennetehtävä:*

1. \_\_\_\_\_

9. \_\_\_\_\_

2. \_\_\_\_\_

10. \_\_\_\_\_

3. \_\_\_\_\_

11. \_\_\_\_\_

4. \_\_\_\_\_

12. \_\_\_\_\_

5. \_\_\_\_\_

13. \_\_\_\_\_

6. \_\_\_\_\_

14. \_\_\_\_\_

7. \_\_\_\_\_

15. \_\_\_\_\_

8. \_\_\_\_\_

\_\_\_\_\_ /15p

Koko koe yhteensä \_\_\_\_\_ / 100p

PRO GRADU – TUTKIMUS PERUSKOULUN 6. LUOKAN  
A1/A2-SAKSAN OPPIMISTULOKSISTA

Hyvä saksanopettaja,

Olen saksan kielen opiskelija Jyväskylän yliopistosta ja olen saanut pro gradu-työni aiheeksi tutkia A1- ja A2-saksan lukijoiden kielitaidon tasoeroja ala-asteen päätyessä. Tutkimuksen taustalla on opettajien käytännön työssään kohtaama ongelma: kuinka suuria eroja A1- ja A2-lukijoiden kielitaidossa on ja voiko oppilaat yhdistää samaan ryhmään seitsemännellä luokalla.

Tutkimuksessa käytän **Sukol-Palvelu Oy:n ja Suomen saksanopettajat ry:n kevääksi 2000 tuottamaa peruskoulun 6. luokan A2/A1-saksan koetta**. Koulullanne tätä koetta on käytetty keväällä 2000 ja siksi olen valinnut koulunne ryhmään, jolta tarvittava tutkimusaineisto tullaan keräämään. Kyselylomakkeet on numeroitu käsittelyn mahdollistamiseksi ja vastaukset käsitellään nimettöminä ehdottoman luottamuksellisesti.

Tutkimukseen osallistuaksenne pyydän teitä pitämään kyseisen kokeen koulunne tämän vuoden kuudennen luokan saksan ryhmille sekä täyttämään ja palauttamaan tutkimukseen liittyvän kyselyn, jonka oheistan tässä jo nähtäväksenne. Varsinaisen lomakkeen ja palautuskuoren lähetämme koulullenne vahvistettuanne osallistumisenne. Tutkimukseen liittyvä koe tulee pitää **27.04.2001 mennessä ja kysely palauttaa 04.05.2001 mennessä**.

Huomatkaa, että voitte pitää tutkimuskokeen oppilailenne heti tämän kirjeen saatuanne ja tutkimukseen osallistumisestanne päätettyänne, viimeistään 27.04.2001 mennessä. Tutkimukseen voitte osallistua myös jos olette tänä keväänä jo pitäneet viime kevään A2/A1 saksan kokeen nykyisille 6. luokan oppilailenne ja teillä on oppilasvastaukset tallella. Osiokohtaisia pisteytystietoja kysytään tutkimuskyselyn viimeisellä sivulla.

Kuten tiedätte, tämän keväisen 6. luokan saksan kokeen suositusviikko on viikko 17 eli 23.04.2001 alkava viikko. Voitte mainiosti pitää ensin tutkimuskokeen A-saksan ryhmällänne/ryhmillenne ja tämän kevään kokeen edellä mainitun suositusviikon jälkeekin, jolloin tutkimuskoe toimii oppilaille eräänlaisena harjoituskokeena tämän kevään koetta varten.

Toivon hartaasti, että voitte auttaa tässä kieltenopettajille tärkeässä asiassa. **Palauttamalla oheisen vastauslomakkeen 06.04.2001 mennessä vahvistatte osallistumisenne tutkimukseen.**

Voitte mielihyvin ottaa minuun yhteyttä, jos teillä on tutkimukseen liittyen kysymyksiä.

Lämmin kiitos avustanne jo etukäteen!

Kevätterveisin,

Leena Grönqvist / Jyväskylän yliopisto  
E-mail: [leenagro@st.jyu.fi](mailto:leenagro@st.jyu.fi)  
Tel. 050 581 5232  
Kotiosoite: Tellervonkatu 8 b 20  
40 100 JYVÄSKYLÄ

PRO GRADU -TUTKIMUS PERUSKOULUN 6. LUOKAN A1/A2-SAKSAN  
OPPIMISTULOKSISTA

Hyvä saksan opettaja,

Jyväskyläläinen opiskelija Leena Grönqvist tekee pro gradu -työnään vertailevan tutkimuksen peruskoulun 6. luokan A1/A2-saksaa lukevien oppilaiden oppimistuloksista.

Myös Suomen kieltenopettajien liitossa on usein herännyt kysymys, miten A1- ja A2-kieltä lukevien oppilaiden tulokset eroavat toisistaan. Mm. tähän kysymykseen haetaan Leena Grönqvistin pro gradu -työssä vastausta. Siksi toivomme, että kevätkiireiden keskellä voisitte osallistua tähän luottamukselliseen tutkimukseen, josta saatuja tietoja käytetään A1/A2-saksan opetuksen hyväksi.

Helsingissä, 28.3.2001

Lämpimin yhteistyöterveisin

Suomen kieltenopettajien liitto SUKOL ry

Sukol-Palvelu OY

  
Terttu Valojärvi  
puheenjohtaja

ANHANG 3. Der Brief von Sukol ry.

**PRO GRADU – TUTKIMUS PERUSKOULUN 6. LUOKAN  
A1/A2-SAKSAN OPPIMISTULOXSISTA**

**KYSELYLOMAKE**

Ympyröikää jokaisesta kohdasta vastaustanne vastaava numero. Avokysymyksissä kirjoittakaa vastaus riville, tarvittaessa myös kääntöpuolta voi käyttää vastaamiseen. Antamanne tiedot käsitellään luottamuksellisesti kullekin koululle annetun tutkimusnumeron avulla eikä koulunne nimi tule missään vaiheessa esiintymään tutkimuksessa.

Jos koulussanne opiskellaan saksaa sekä A1- että A2-kielenä tai koulussanne on useampi A1/A2-ryhmä, toivon, että pidätte kokeen kaikissa 6. luokan saksanryhmissä. Jos pidätte kokeen useammassa kuin yhdessä ryhmässä, kopioikaa kysely- ja pistekyselylomake ja vastatkaa kysymyksiin sekä täyttäkää pistekysely **kunkin ryhmän kohdalta erikseen**. Niitatkaa kunkin ryhmän kysely- ja vastauslomakkeet yhteen siten, etteivät eri ryhmien lomakkeet sekaannu keskenään. Kaikki kyselylomakkeet voi lähettää oheisessa palautuskuoressa.

Jos ilmoititte vahvistuslomakkeessa voivanne luovuttaa oppilaiden koevastauslomakkeet tutkijan käyttöön, ei pistekyselyä tarvitse täyttää. Oppilaiden vastauslomakkeet voi palauttaa nimettöminä, mutta niihin tulisi merkitä oppilaan sukupuoli. Oppilaiden vastauslomakkeet ja kirjoitelmat voi lähettää samassa palautuskuoressa kyselylomakkeen kanssa.

**I. Arvionne peruskoulun 6. luokan A2/A1 – saksan kokeesta:  
(Sukol-Palvelu Oy ja Suomen saksanopettajat ry 2000)**

Vastatkaa seuraaviin kysymyksiin asteikolla 1 – 5. (1 = täysin samaa mieltä – 5 = täysin eri mieltä)

- |  |           |        |
|--|-----------|--------|
| 1. Koe testasi mielestänne keskeistä sanastoa:   | 1 2 3 4 5 |        |
| 2. Koe testasi mielestänne keskeisiä taitoja:  | 1 2 3 4 5 |        |
| 3. Oppilaiden kokeesta saamat pisteet<br>vastaavat mielestänne hyvin heidän taitojaan: |           | Muuta: |
| • Kuullun ymmärtäminen   | 1 2 3 4 5 | _____  |
| • Luetun ymmärtäminen  | 1 2 3 4 5 | _____  |
| • Oikein-väärin –tehtävä   | 1 2 3 4 5 | _____  |
| • Yhdistä kysymys ja vastaus   | 1 2 3 4 5 | _____  |
| • TV-tarjotin  | 1 2 3 4 5 | _____  |
| • Rakennetehtävä   | 1 2 3 4 5 | _____  |
| • Kirjoitelma  | 1 2 3 4 5 | _____  |
| 4. Arvioikaa, oliko koe kokonaisuutena onnistunut:                                     |           |        |

**II. A1-ryhmän opettajille: (Jos ryhmä opiskelee saksaa A2-kielenä, siirtykää kysymykseen 7)**

5. A1-saksa alkaa koulussanne \_\_\_\_\_. luokalla
6. Kyseisen A1-saksan ryhmän todellinen vuosiviikkotuntimäärä on arvionne mukaan 3.-6. luokan ajan ollut keskimäärin:
- |               |   |
|---------------|---|
| Alle 8 tuntia | 1 |
| 8 tuntia      | 2 |
| 9 - 10 tuntia | 3 |
| Yli 10 tuntia | 4 |

(Koulun numero tutkimuksessa: \_\_\_\_\_)

⇒

III. **A2-ryhmän opettajille:** (Jos ryhmä opiskelee saksaa A1-kielenä, siirtykää kysymykseen 13)

7. A2-saksa alkaa koulussanne \_\_\_\_\_luokalla
8. Kyseisen A2-saksan ryhmän todellinen vuosiviikkotuntimäärä on arvionne mukaan 5.-6. luokilla ollut keskimäärin:
- |               |   |
|---------------|---|
| Alle 4 tuntia | 1 |
| 4 tuntia      | 2 |
| 5 - 6 tuntia  | 3 |
| Yli 6 tuntia  | 4 |
9. A2-kielen opetuksen tavoitteena koulussanne on saavuttaa A1-kielen taso ala-asteen päättyessä: Kyllä 1 Ei 2
10. Arvioikaa, missä määrin A2-saksassa saavutetaan A1-saksan taso. (Käyttäkää %-merkintää, esim. 80%): \_\_\_\_\_
11. Jos vastauksenne edelliseen kysymykseen on alle 100%, arvioikaa, missä taidoissa A2-saksan lukijoilla on puutteita tai mitkä tärkeät asiat jäivät käsittelemättä:
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
12. Arvioikaa, minkälaisia vaikeuksia A2-saksan lukijoilla olisi, jos heidät yhdistettäisiin samaan ryhmään A1-lukijoiden kanssa seitsemännellä luokalla:
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
13. Muuta kommentoitavaa tai lisätietoja :
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

IV. **Taustatietoja:**

14. Kieli: A1 A2
15. Koulun sijaitsee: \_\_\_\_\_läänissä
16. Kunnassanne on asukkaita: \_\_\_\_\_
17. Koulussanne on oppilaita: \_\_\_\_\_
18. Opetusryhmän koko: \_\_\_\_\_
19. Olette:
- |                                      |   |
|--------------------------------------|---|
| Perusopetuksen luokanopettaja        | 1 |
| Saksan kielen perusopetuksen lehtori | 2 |
| Tuntiopettaja                        | 3 |
| Kiertävä opettaja                    | 4 |
20. Yliopistoarvosananne saksan kielessä:
- |              |   |
|--------------|---|
| Approbatur   | 1 |
| Cum laude    | 2 |
| Laudatur     | 3 |
| Ei arvosanaa | 4 |

(Koulun numero tutkimuksessa: \_\_\_\_\_)

⇒

**PISTEKYSELY**

Kirjatkaa taulukoihin tiedot siitä, kuinka moni oppilas on saavuttanut taulukon osoittaman pistemäärän. Pyöristäkää pisteet lähimpään kokonaislukuun. Jokaisesta osiosta täytetään oma taulukkonsa.

Kuullun ymmärtäminen			Luetun ymmärtäminen			Oikein-väärin – tehtävä			Yhdistä kysymys ja vastaus		
pisteet	tytöt	pojat	pisteet	tytöt	pojat	Pisteet	tytöt	Pojat	pisteet	tytöt	pojat
0			0			0			0		
1			1			1			1		
2			2			2			2		
3			3			3			3		
4			4			4			4		
5			5			5			5		
6			6			6					
7			7			7					
8			8			8					
9			9			9					
10			10			10					
11			11								
12			12								
13			13								
14			14								
15			15								
16			16								
17			17								
18			18								
19			19								
20			20								

TV-tarjotin			Rakennetehtävä			Kirjoitelma kuvasta		
pisteet	tytöt	pojat	pisteet	Tytöt	pojat	Pisteet	tytöt	pojat
0			0			0		
1			1			1		
2			2			2		
3			3			3		
4			4			4		
5			5			5		
6			6			6		
7			7			7		
8			8			8		
9			9			9		
10			10			10		
			11			11		
			12			12		
			13			13		
			14			14		
			15			15		
						16		
						17		
						18		
						19		
						20		

(Koulun numero tutkimuksessa: \_\_\_\_\_)

**PRO GRADU – TUTKIMUS PERUSKOULUN 6. LUOKAN  
A1/A2-SAKSAN OPPIMISTULOKSISTA**

Hyvä saksanopettaja,

Toivon, että olette saanut kirjeeni, jossa kerroin pro gradu-tutkimuksestani, jonka tarkoituksena on tutkia A1- ja A2-saksan lukijoiden kielitaidon tasoeroja ala-asteen päättyessä. Muutamat koulut ovat jo ilmoittaneet halukkuudestaan osallistua tutkimukseen. Tarvitsisin kuitenkin vielä mukaan erityisesti niitä kouluja, joissa saksaa opiskellaan A1-kielenä. Toivon hartaasti, että kaikkien kevätkiireidenne keskelläkin voisitte osallistua tutkimukseen.

Tutkimuksessa käytetään siis **Sukol-Palvelu Oy:n ja Suomen saksanopettajat ry:n kevääksi 2000 tuottamaa peruskoulun 6. luokan A2/A1-saksan koetta**. Teidän tulisi pitää kyseinen koe koulunne tämän vuoden kuudennen luokan saksan ryhmille sekä täyttää ja palauttaa tutkimukseen liittyvä kyselyn, jonka lähetän koululenne vahvistettuanne osallistumisenne. Tutkimukseen liittyvä koe tulee pitää **27.04.2001 mennessä ja kysely palauttaa 04.05.2001 mennessä**.

Lämpimät kiitokset vaivannäöstänne!

Terveisin,

Leena Grönqvist / Jyväskylän yliopisto  
E-mail: [leenagro@st.jyu.fi](mailto:leenagro@st.jyu.fi)  
Tel. 050 581 5232  
Kotiosoite: Tellervonkatu 8 b 20  
40 100 JYVÄSKYLÄ

**VAHVISTUS OSALLISTUMISESTA**

- Koulumme haluaa osallistua tutkimukseen. Koulumme nro: \_\_\_\_\_
- Ryhmät, joilla pidämme kokeen:
- A1-saksa Ryhmiä: \_\_\_\_\_ Oppilaita ryhmissä yhteensä: \_\_\_\_\_
- A2-saksa Ryhmiä: \_\_\_\_\_ Oppilaita ryhmissä yhteensä: \_\_\_\_\_
- Voimme luovuttaa oppilaiden vastauslomakkeet tutkimusta varten.
- Emme voi luovuttaa oppilaiden vastauslomakkeita tutkimusta varten.
- Koulullamme ei ole kevään 2001 kokeeseen liittyvää äänitettä. Pyydämme toimittamaan koulullemme äänitteen.

Vastaava opettaja: \_\_\_\_\_

Koulun nimi: \_\_\_\_\_

Lähiosoite: \_\_\_\_\_

\_\_\_\_\_

Vahvistus osallistumisesta faksataan **18.04.2001 mennessä** numeroon:

**(014) 260 1421** (Jyväskylän yliopisto / saksan laitos) tai postitetaan osoitteeseen:

**Leena Grönqvist, Tellervonkatu 8 b 20, 40 100 JYVÄSKYLÄ**

(Koulun numero tutkimuksessa: \_\_\_\_\_)