

Hörverstehen und Hör-Sehverstehen in der Fremdsprache – unterschiedliche Konstrukte?

Raphaela Porsch¹, Rüdiger Grotjahn² und Bernd Tesch³

According to the "National Educational Standards for the First Foreign Language" enacted by the Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in Germany (KMK 2003; 2004) listening comprehension and audio-visual comprehension in the first foreign language (in Germany: English or French) are competencies that need to be acquired until the end of lower secondary education. However, the Standards and the Common European Framework of Reference (Council of Europe 2001) – on the Standards are based – have not provided two separate construct descriptions for developing tasks for teaching and testing the understanding of audio and audio-visual material. The experimental study to be presented aims at answering the question of whether one can differentiate empirically between these constructs by assessing German ninth-graders ($N = 156$) with either a listening or an audio-visual comprehension test in French as a foreign language using the same items. Results show that if the pictures of the authentic videos in French provide help, learners are significantly better in the audiovisual condition. In addition, the challenges posed by the different tasks as well as the test-taker strategies are explored, and a discussion about consequences for the construct definitions along with recommendations for teacher training is provided.

1. Einleitung

Zu den Medien als "Lehr- und Lernmittel im Unterschied zu lehrenden Personen" (Weidenmann 2006: 425) bzw. jede Art "nicht personaler Informationsträger" (Brünken & Leutner 2008: 551) zählt im Fremdsprachenunterricht seit vielen Jahren der Gebrauch auditiver und audiovisueller Materialien. Sprachenlernen mit Hörtexten, Filmen und Fernsehprogrammen bzw. Fernseh- und Filmsequenzen wird in der (Fremdsprachen-)Didaktik als besonders geeignet für den Unterrichtseinsatz angesehen, wobei verschiedene Legimitationsargumente angeführt werden (vgl. z.B. Raabe 1997; 2003; Roche 2008; Schwerdtfeger 1989; Thaler 2007a; b). Besonders häufig werden die Authentizität der Inhalte, die Popularität und die motivierende Funktion von audiovisuellen Medien genannt, wobei im Hinblick auf das Lernen zu bedenken ist, dass audiovisuelle Medien im Alltag der

1 Korrespondenzadresse: Dr. Raphaela Porsch, Westfälische Wilhelms-Universität Münster, Institut für Psychologie, Abteilung Pädagogische Psychologie, Fliegerstr. 21, 48151 Münster, Telefon: 0251/1491233, raphaela.porsch@uni-muenster.de

2 Korrespondenzadresse: Prof. Dr. Rüdiger Grotjahn, Ruhr-Universität Bochum, Seminar für Sprachlehrforschung, 44780 Bochum, Telefon: 0234/32-28729, ruediger.grotjahn@ruhr-uni-bochum.de

3 Korrespondenzadresse: Dr. Bernd Tesch, Humboldt-Universität zu Berlin, Institut für Qualitätsentwicklung im Bildungswesen, Unter den Linden 6, 10099 Berlin, Telefon 030/2093-5505, bernd.tesch@iqb.hu-berlin.de

Schüler vor allem der Unterhaltung dienen. So weisen Niegemann, Domagk, Hessel, Hein, Hupfer & Zobel (2008: 265) darauf hin, dass Zuschauer Fernsehen vor allem als Unterhaltungsmedium ansehen und "sich die Assoziationen mit Unterhaltung negativ auf das Lernen auswirken können".

Auch im Bereich der Diagnose fremdsprachlicher Kompetenzen, der im vorliegenden Beitrag eine zentrale Rolle spielt, werden insbesondere im Kontext computerbasierten und webbasierten Testens immer häufiger auch audiovisuelle Medien eingesetzt (vgl. z.B. Ockey 2007; Wagner 2008). Allerdings handelt es sich dabei bisher zumeist um Standbilder – so z.B. beim neuen TOEFL iBT. Der Einsatz von Bildern wird u.a. folgendermaßen begründet:

It seems reasonable to conclude that some sort of visual stimulus should be employed since, among other reasons, most target-language-use situations include visual stimuli, and it seems unreasonable to have test takers stare at a blank computer screen while taking a computer-mediated test (Ockey 2007: 517).

Im vorliegenden Beitrag gehen wir der Frage nach, ob visuelle Informationen das Hörverstehen beeinflussen und inwieweit es sich beim fremdsprachlichen Hör- und Hör-Sehverstehen um empirisch unterscheidbare Kompetenzen und Konstrukte handelt. Dazu wird nach einer Darstellung relevanter theoretischer Aspekte (Abschnitt 2) eine empirische Studie vorgestellt (Abschnitt 3), welche 2009 im Kontext der Überprüfung der Bildungsstandards für die erste Fremdsprache (Französisch) für den Mittleren Schulabschluss (KMK 2003) durchgeführt wurde. Es folgen eine Diskussion und ein kurzer Ausblick auf weitere Forschungen in diesem Bereich (Abschnitt 4).

2. Theoretischer Hintergrund

2.1 Verarbeitung von Text- und Bildinformationen

Die pädagogisch-psychologische Medienforschung beschäftigt sich vor allem mit der Frage, wie effektiv das Lernen bestimmter Inhalte mit Hilfe unterschiedlicher Medien und bezogen auf unterschiedliche Lerner ist und welche Prozesse bei den Rezipienten als aktive Teilnehmer am Verarbeitungsprozess ablaufen. Bereits 1971 stellte Paivio die Theorie der dualen Kodierung vor, nach der Sprache und Bilder in zwei unabhängigen, jedoch miteinander in Verbindung stehenden kognitiven Systemen verarbeitet werden. An die Vorstellung von Paivio (1986) anknüpfend haben Forscher aus der Psychologie der Informationsverarbeitung wie Baddeley (1992; 2006; 2007), Mayer (2001) und Schnotz (2005) theoretische Modelle vorgestellt, die die kognitiven Prozesse bei der Verarbeitung von Text und Bild beschreiben.

Mayers Modell bzw. seine *Cognitive Theory of Multimedia Learning* gilt "heute als eine Art Standardtheorie des Lernens mit Medien" (Brünken & Leutner 2008: 553) und beschreibt zahlreiche Erkenntnisse der Kognitionspsychologie. Mayer (2005a: 32) betont diesen Anspruch für jede Theorie in diesem Bereich: "[A cognitive theory of learning must be] consistent with cognitive science principles of learning." Das so genannte SOI-Modell nach Mayer berücksichtigt drei zentrale Prozesse: *Selection*, *Organization* und *Integration* (daher SOI-Modell; vgl. Abbildung 1). Diese Vorstellungen basieren auf den folgenden theoretischen Ansätzen: (1) der Theorie der dualen Kodierung der Information nach Paivio (1986), (2) der Annahme einer begrenzten Arbeitsgedächtniskapazität und (3) der Vorstellung, dass der Informationsaufnehmende die Verarbeitungsprozesse aktiv steuert.

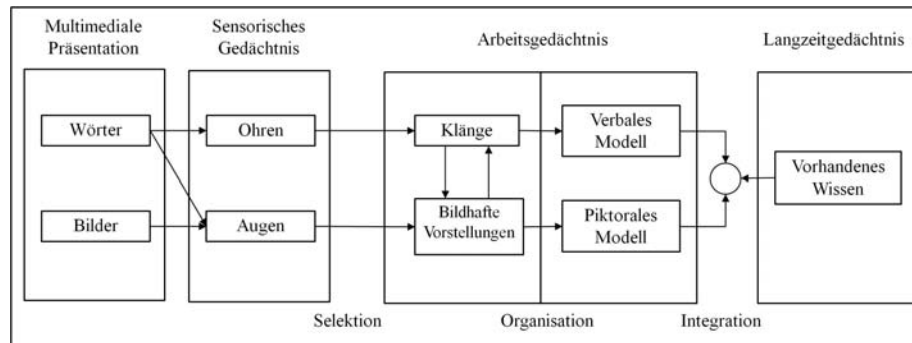


Abbildung 1: SOI-Modell nach Mayer (2001)

Die begrenzte Kapazität des Arbeitsgedächtnisses begründet sich in der *Cognitive Load Theory* (vgl. Swellers & Chandler 1991; vgl. auch die Ausführungen in Brünken, Plass & Leutner 2003; Niegemann et al. 2008; Tardieu & Gyselinck 2003). Dabei ist zu berücksichtigen, dass die Grenzen der Bearbeitungskapazität individuell verschieden sind und u.a. vom Vorwissen des Rezipienten abhängen. Als Unterstützung zur Entscheidung, welche der Informationen in welcher Reihenfolge bearbeitet werden sollten, um eine kognitive Überlastung beim Lernen zu vermeiden, dienen metakognitive Strategien (Mayer 2005a). Ihre Einzelfunktionen sind die Selbstbeobachtung, Selbstregulation und die Koordination der Lernaktivitäten. Die Betonung von aktiven Verarbeitungsprozessen im Rahmen der Verarbeitung macht u.a. den Bezug zum allgemeinen Lernbegriff in der pädagogischen Psychologie deutlich. Damit der Rezipient sich aktiv ein kohärentes mentales Modell des Gegenstands aufbauen kann, muss er relevante Infor-

mationen auswählen und Neues mit bereits erworbenem Wissen oder vorhandenen Kompetenzen in Verbindung bringen.

Welche Prozesse laufen bei den Empfängern von Text-Bild-Informationen nach Mayer ab? Mayer (2001) geht zunächst von einer grundsätzlich getrennten Verarbeitung von Text- und Bildinhalten aus – wobei jeder Kanal nur eine begrenzte Verarbeitungskapazität hat. Dazu werden die folgenden fünf kognitiven Prozesse unterschieden (nach Niegemann et al. 2008: 52):

- Auswahl (Selektion) von relevanten bzw. als wichtig erachteten Wörtern,
- Auswahl (Selektion) von relevanten Bildinhalten,
- Strukturierung (Organisation) der ausgewählten Textinhalte und Bildung eines kohärenten verbalen Modells,
- Strukturierung (Organisation) der als wichtig erachteten Bildinhalte zu einem kohärenten Modell,
- Verknüpfen (Integration) der Textrepräsentation mit der Bildrepräsentation und von neuem Wissen mit bereits erworbenen Wissensbeständen aus dem Langzeitgedächtnis.

Verarbeitung meint in den Ausführungen von Mayer das Lernen von Zusammenhängen oder Wissensinhalten. Nachfolgend soll in Anknüpfung an diese Modellvorstellungen auf die Besonderheiten der Verarbeitung von Text-Bild-Informationen in Testsituationen unter Verwendung von fremdsprachlichen Hörtexten und Videosequenzen eingegangen werden. Nach den Annahmen bzw. Prämissen des SOI-Modells kann an dieser Stelle bereits die Hypothese formuliert werden, dass die Informationsverarbeitung bzw. der Wissenserwerb mit Hilfe von audio-visuellen Medien behindert wird, wenn (1) zu viele Informationen gemeinsam über einen Kanal oder wenn (2) zu viele Informationen gleichzeitig auf beiden Kanälen geliefert werden. Dabei werden der zusätzlich genutzte Sinneskanal beim Hör-Sehverstehen und die damit verbundenen Prozesse als das primäre Unterscheidungsmerkmal zum Hörverstehen verstanden.

Clark & Mayer (2002) begründen in Bezug auf die Effektivität des Lernens die Vorteile einer Text-Bild-Kombination gegenüber der Darbietung von Texten ohne Bilder, solange bestimmte Prämissen gelten. Die in Tabelle 1 aufgeführten Prinzipien sollen den Autoren zufolge die Grundlage für die Erstellung medialer Lernangebote darstellen.

Multimediaprinzip	Vorteil von Text und Bild gegenüber nur Bild (insbesondere bei Lernern mit geringem Vorwissen)
Kontiguitätsprinzip	Zeitliche und räumliche Nähe bei der Darbietung von Bild und Text empfehlenswert
Modalitätsprinzip	Gesprochener Text in Kombination mit Bildern geeigneter als nur geschriebener Text
Redundanzprinzip	Vermeidung einer Darbietung von gleichen Informationen über mehrere Kanäle (z.B. gleichzeitige Darbietung von Inhalten in einem gesprochenen und geschriebenen Text)
Kohärenzprinzip	Verzicht auf die Präsentation irrelevanter Informationen wie z.B. von Hintergrundmusik, redundanten Graphiken und Wörtern)

Tabelle 1: Gestaltungsprinzipien für die Konzeption multimedialen Lernens (nach Clark & Mayer 2002)

Einige der genannten Prinzipien lassen sich jedoch deutlich einfacher bei der Gestaltung von multimedialen Lernangeboten umsetzen als andere. Dies gilt vor allem in Bezug auf die Art der Bilder und Texte. In diesem Zusammenhang sind u.a. folgende Fragen relevant: Handelt es sich um geschriebene oder gesprochene Texte? Werden bewegte Bilder (z.B. in Form von Videofilmen) oder statische Bilder (z.B. in Form von Standfotos) eingesetzt? Handelt es sich bei den bewegten Bildern z.B. um Animationen oder Simulationen? Und schließlich: Wurde auf didaktisch aufbereitetes oder auf authentisches Material zurückgegriffen? Im folgenden Abschnitt soll näher auf die Merkmale und Verwendung von Videofilmen eingegangen werden, da diese Art von Bewegtbildern in Kombination mit einem gesprochenen Text als Instrumente in der in Abschnitt 3 beschriebenen empirischen Studie eingesetzt wird.

2.2 Verstehensleistungen bei der Rezeption von Videosequenzen in einer Fremdsprache

Videos sind analoge oder digitale Aufnahmen der Realität und stellen eine Form von audio-visuellen Medien dar. Es sind Bewegtbilder mit oder ohne Ton bzw. Text, wobei Texte gesprochen oder geschrieben (z.B. als Untertitel, Beschriftungen) eingesetzt werden können. Die gleichzeitige Darbietung von Bildern und Text als Video impliziert in der Regel, dass die Informationen flüchtig sind. Die Geschwindigkeit kann zumeist nicht beeinflusst werden, denn wie beim "reinen" Hörverstehen von Tonaufnahmen hängt diese vom Sprecher bzw. von der Bildfolge ab (vgl. auch Grotjahn & Tesch 2010a). Eine Kontrolle des Informationseingangs zur Überprüfung und gegebenenfalls Korrektur des Verstehens

kann nur eingeschränkt stattfinden, u.a. dann, wenn die Möglichkeit besteht, das Video mehrmals anzusehen, es zu unterbrechen oder die Hörtexte in geschriebener Form als Untertitel⁴ mitlesen zu können. Neben dem auditiven Input bzw. dessen Merkmalen (wie Sprecheranzahl und -charakteristik) unterscheiden sich Videos durch ihre visuelle Gestaltung. Es gibt eine Fülle von Gestaltungsmitteln wie z.B. Kameraführung oder Schnitttechnik. Zudem sind paralinguistische Merkmale wie Gestik (z.B. deiktisch, ikonisch, metaphorisch), Mimik, artikulatorische Bewegungen,⁵ die Kleidung der Akteure der Videos oder auch Informationen zum Sprecherwechsel in Konversationen zugänglich, so dass im Vergleich zum reinen Hörverstehen weitere wichtige Informationen zur Verfügung stehen. Dabei können die paralinguistischen Informationen die verbalen Äußerungen begleiten, ergänzen oder ersetzen (vgl. z.B. Ankerstein 1972; Baltova 1999; Feak & Salehzadeh 2001; Hardison 2010; Hoven 1999; Kellerman 1992; Pisoni & Remez 2005; Riley 1979; Shin 1998; Skipper, Van Wassenhove, Nusbaum & Small 2007; Sueyoshi & Hardison 2005; Thaler 2007a; b).

Visuelle Informationen, die sich auf den Kontext unter Einschluss der jeweiligen Sprecher beziehen, werden von Ginther (2002: 134) und anderen Autoren als kontextbezogene Visualisierungen (engl. *context visuals*) bezeichnet. Eine inhaltsbezogene Visualisierung (engl. *content visual*) liegt dagegen vor, wenn sich die visuelle Information unmittelbar auf den Inhalt des verbalen Stimulus bezieht, wie z.B. bei der Erklärung einer statistischen Grafik in einer Vorlesung oder bei Bildern, die sich direkt auf die Inhalte einer Erzählung beziehen (vgl. auch Ginther 2001; Ockey 2007: 520).

Inhaltsbezogene Visualisierungen, die die auditive Information ergänzen, scheinen sich tendenziell durchgängig positiv auf die Verstehensleistungen auszuwirken. Nach Ginther (2001; 2002)⁶ scheinen dagegen kontextbezogene Visualisierungen vor allem dann einen Verständnis erleichternden Effekt zu haben, wenn sie, wie z.B. in akademischen Diskussionen, den Sprecherwechsel verdeutlichen (vgl. auch die entsprechenden Ausführungen in Feak & Salehzadeh 2001: 482). In anderen Situationen, wie z.B. in akademischen Vorträgen, können sie

4 In der Literatur wird über die Vor- und Nachteile von Untertitelung von bewegten Bildern zur Unterstützung beim Erwerb einer Zweit-/Fremdsprache kontrovers diskutiert. Empirische Studien zeigen konträre Ergebnisse, wobei unterschieden werden muss, ob die Untertitelung in der gleichen oder der zweiten Sprache erfolgt (vgl. den Forschungsüberblick in Vanderplank 2010 sowie Guichon & McLorinan 2008).

5 McGurk & MacDonald (1976) konnten in einer klassischen Studie zeigen, dass das Hörverstehen in hohem Maße von den Lippenbewegungen beeinflusst wird (sog. *McGurk-Effekt*). Hinweise zum McGurk-Effekt bei Fremdsprachenlernern geben Sueyoshi & Hardison (2005: 664f.). Die Autoren erwähnen zudem, dass weitere Merkmale, wie z.B. Veränderungen in der Gesichtsmuskulatur, das Hörresultat beeinflussen können.

6 Die Studie von Ginther (2001; 2002) bezieht sich auf den Einsatz von Standbildern beim TOEFL CBT (*Computer-based Test*).

sich jedoch sogar negativ auf das Verständnis auswirken (vgl. Ginther 2002: 162f. sowie auch Coniam 2001; Gruba 1999).

Ginther führt dabei den potenziell positiven Effekt sowohl von kontextbezogenen als auch von inhaltsbezogenen Visualisierungen auf das in Abschnitt 2.1 genannte medienpsychologische Kontiguitätsprinzip sowie auf die "Passung" von auditiver und visueller Information zurück:

Facilitative effects occur when the presentation of visuals is contiguous and, most importantly, when it is directly related to the content of the information presented in the audio portion of the stimuli or it marks a turn in the conversation. Thus, theoretical discussions with respect to contiguity are supported ... (Ginther 2002: 162)

Den Forschungsstand speziell zu kontextbezogenen Visualisierungen fasst Ockey (2007: 521) folgendermaßen zusammen:

The research to date seems to suggest that listening test input context-only visuals may not improve performance. There is also some evidence that context-only visuals may hinder the performance of test takers.

Zugleich kommt der Autor aufgrund einer eigenen Untersuchung zu dem Schluss, dass der Einsatz von kontextbezogenen Standbildern nur einen minimalen Einfluss auf das Testkonstrukt hat, dass jedoch kontextbezogene Videos das Testkonstrukt deutlich verändern können. Bezogen auf den akademischen Bereich charakterisiert Ockey (2007: 532) ein entsprechend erweitertes Testkonstrukt folgendermaßen:

The results suggest that an enlarged academic listening construct, one which includes the visuals present in the 'cognitive environment', would include the ability to gain information from lip movements, facial gestures, and other body language such as hand movements. This enlarged construct might also include the ability to simultaneously use visual cues, process audio input, and take notes.

Im Unterricht werden Videos zu verschiedenen Zwecken eingesetzt (vgl. Kittelberger & Freisleben 1994; Sherman 2003). In diesem Beitrag soll es vor allem um den Einsatz als Diagnoseinstrument für fremdsprachliche Verstehensleistungen gehen. Die in Abschnitt 2.1 genannten Gestaltungsprinzipien für das multimediale Lernen finden auch in diesem Zusammenhang Anwendung. Zusätzlich können das sog. Personalisierungsprinzip und das Prinzip der Stimmauswahl relevant sein. So kann ein personalisierter Sprachstil (z.B. Verwendung der zweiten Person) deswegen geeigneter als ein sachlicher Sprachstil sein, weil eine Personalisierung die Ähnlichkeit mit Realgesprächen erhöht. In Bezug auf die Stimmauswahl gelten menschliche Stimmen als geeigneter als computergenerierte. Schließlich werden für die Konzeption von Lernangeboten mit Videos Dialoge gegenüber Monologen präferiert und eine Informationspräsentation mit einem Sprecher im Off als ungünstig für das Verstehen angesehen. Im Kontext der Testung des Hör-Sehverstehens ist jedoch auch das zugrundeliegende Konstrukt zu

berücksichtigen (vgl. Abschnitt 2.6), welches beispielsweise von Fremdsprachlernern auf einem bestimmten Kompetenzniveau das Verstehen von Monologen beinhalten kann.

Wichtig für den Einsatz von Videos ist die Vermeidung unnötiger kognitiver Belastung bzw. eine nach Sprach- bzw. Lernstand geeignete Auswahl von Videos. Erleichternd kann eine Zerlegung des Inputs in mehrere Teile sein (Prinzip der Segmentierung; vgl. Mayer 2005b) oder die Darbietung von Ausschnitten eines Videos bzw. einer Film- oder Fernsehaufnahme. Weitere Hilfen für das bessere Verständnis, wie die Kontrolle über das Abspielen einer Sequenz (Interaktivitätsprinzip; vgl. Betrancourt 2005), sind im Kontext einer Testsituation, welche vergleichbare Bedingungen für alle Teilnehmer erfüllen soll, nur bedingt anwendbar. Einfach realisierbar ist z.B. die zwei- oder mehrmalige Darbietung von Videos, die allerdings mit einer Veränderung des Testkonstrukts einhergeht (im Vergleich zur einmaligen Rezeption). Eine mehrmalige Darbietung kann aber dennoch durchaus von Vorteil sein und wird in vielen Tests zur Überprüfung des Hörverstehens in der L2 angewendet (vgl. auch Abschnitt 2.4; Grotjahn & Tesch 2010a: 131f.).

Die Geltung der Gestaltungsprinzipien für die Konzeption multimedialen Lernens wird durch zahlreiche empirische Studien gestützt, wobei Wagner (2008) für den L2-Kontext darauf hinweist, dass die Ergebnisse der Studien nicht eindeutig seien. So vergleichen viele Studien in einem quasi-experimentellen Design die Hörverstehens- und die Hör-Sehverstehensleistungen von L2-Lernenden. Es zeigte sich, dass der zusätzliche visuelle Input nicht in allen Fällen zu einer erhöhten Testleistung der audio-visuellen Rezipienten führt (vgl. z.B. Brett 1997; 2001; Coniam 2001; Ginther 2001; 2002; Gruba 1993; 1997; 1999). Wenn Bild und Ton sich nicht hinreichend ergänzen, kann es leicht zu einer validitätsmindernden Informationsüberlastung auf Seiten des Rezipienten kommen. Zudem können Videos oder Standbilder im Hinblick auf den Hörtext eine Ablenkung darstellen. Auch können visuelle Informationen z.B. hinsichtlich der ethnischen Zugehörigkeit des Sprechers zu Erwartungen führen, die das Verstehen beeinflussen (vgl. Feak & Salehzadeh 2001: 490). Möglich ist auch, dass die visuellen Hilfen von den Zuschauern nicht oder ungenügend wahrgenommen werden (vgl. z.B. Bejar, Douglas, Jamieson, Nissan & Turner 2000; Ockey 2007; Wagner 2007, 2008). Auffallend ist im Vergleich der Studien ein erhebliches Maß an interindividueller Varianz. Es stellt sich jedoch die Frage, ob viele der Einwände bzw. nachgewiesenen Einflussfaktoren nicht auch für L1-Hör-Sehverstehensprozesse gelten.

Andere Studien fanden dagegen statistisch signifikante Unterschiede zugunsten der L2-Lerner, die Fragen zu Hör-Sehverstehensstimuli zu beantworten hatten (vgl. z.B. Chung 1994; Parry & Meredith 1984; Sueyoshi & Hardison 2005). Begründet werden die höheren Testleistungen der audio-visuellen Rezipienten

häufig mit der Annahme, dass diese während der Rezeption der Videos auch auf die nonverbale Kommunikation der Sprecher zurückgreifen können. Da jedoch die Gestik oder Mimik der Sprecher die Beantwortung nicht jeder erfragten Information erleichtert, ist zur Bestätigung dieser Schlussfolgerung die Analyse der Itemanforderungen in Interaktion mit dem Videomaterial unbedingt erforderlich.⁷ Zudem ist zu berücksichtigen, dass es eine Vielzahl von Einflussfaktoren gibt, z.B. Textart, Itemformat oder Sprachstand der Teilnehmer, welche als weitere Ursache für potenzielle Unterschiede in Frage kommen (vgl. Wagner 2008: 220).

Es dürfte klar geworden sein, dass für ein tieferes Verständnis der spezifischen Prozesse bei der Rezeption von audiovisuellen Medien⁸ sowohl eine Analyse des visuellen Inputs und der visuellen Verarbeitungsprozesse als auch der beteiligten Hörverstehensprozesse notwendig ist. Daher soll nachfolgend zunächst näher auf Prozesse beim fremdsprachlichen Hörverstehen eingegangen werden (Abschnitt 2.3), um schließlich Determinanten der Schwierigkeit von Hörverstehensaufgaben zu identifizieren (Abschnitt 2.4). Anschließend sollen Strategien von Testpersonen zur Überwindung von Schwierigkeiten, denen sie bei der Bearbeitung von Testaufgaben zum Hör- und Hör-Sehverstehen in einer Fremdsprache begegnen, beschrieben werden (Abschnitt 2.5).

2.3 Kognitive Prozesse beim fremdsprachlichen Hörverstehen⁹

Fremdsprachliches Hörverstehen wird wie das Leseverstehen in der Regel als ein "interdependentes Wechselspiel von datengeleiteten und von wissensgeleiteten (psycholinguistischen) Verarbeitungsprozessen (= *bottom-up* vs. *top-down processing*)" (Zydatiś 2005: 97; vgl. auch die weitergehenden Ausführungen in Grotjahn & Tesch 2010a: 126-132) dargestellt. Die datengeleiteten Prozesse charakterisiert Zydatiś (2005: 97) als "'Verstehensinseln' auf der phonologischen, syntaktischen oder lexikalischen Ebene, die in eine semantische Interpretation eingebunden werden". Die wissensgeleiteten *top-down*-Prozesse greifen dagegen auf hierarchiehöheres sprachliches Wissen sowie Weltwissen zurück (vgl. auch

7 Die Schlussfolgerung ist nachzuvollziehen, sofern in den Studien bzw. Videos vor allem "gespielte" Dialoge zwischen Personen gezeigt werden, so dass nonverbale Mittel tatsächlich auftreten und von den Testpersonen genutzt werden können. Laut Wagner (2008) nutzen Lerner zudem in sehr unterschiedlicher Weise nonverbale Informationen, um Testitems zu lösen (vgl. auch Ockey 2007). Für andere Formate, wie z.B. bestimmte Typen von Dokumentarfilmen, gilt damit die angeführte Begründung nur eingeschränkt.

8 Betrachtet werden in allen nachfolgenden Ausführungen lediglich Videos mit einem auditiven Gehalt (keine Stummfilme o.ä.).

9 Die Ausführungen der Abschnitte 2.3 und 2.4 beruhen in wesentlichen Teilen auf Grotjahn (2003, Kapitel 6), Grotjahn (2005) und Grotjahn (2010). Vgl. auch die entsprechende Darstellung in Abschnitt 2.4.1 in Grotjahn & Tesch (2010a).

Field 2004; 2008). Zudem handelt es sich um einen dynamischen Prozess, worauf u.a. Ockey (2007: 518) hinweist:

... comprehension is not constructed additively; rather, it is a dynamic process. Understanding is continually modified as incoming stimuli interact with previous input and other existing contextual information.

Außerdem scheint tendenziell zu gelten, dass vor allem Hörer mit Defiziten bei der *bottom-up*-Verarbeitung, und insbesondere mit Schwächen bei der schnellen Worterkennung zur Kompensation auf *top-down*-Prozesse zurückgreifen (vgl. z.B. Goh 2000; Field 2004; 2008; Poelmans 2003).

Ein gravierender Unterschied zwischen Hör- und Leseverstehen ist, dass Hörverstehen ein Echtzeit-Prozess ist (vgl. auch Field 2008: 27f.). Dies heißt u.a., dass das akustische Signal flüchtig ist und zudem die Geschwindigkeit der Informationsübermittlung in der Regel nicht vom Hörer beeinflussbar ist. Eine Einschränkung ergibt sich auch aus der im Vergleich zum visuellen Kanal deutlich geringeren kognitiven Kapazität beim Verarbeiten auditiver Signale. Außerdem müssen bereits verstandene Informationen für das weitere Verstehen im Arbeitsgedächtnis gehalten werden (vgl. z.B. Baddeley 2006; 2007; Leiser 2007). Sind dann noch parallel zum Hören Aufgaben zu lesen oder vorher gelesene oder gehörte Arbeitsanweisungen und Aufgaben im Gedächtnis zu behalten, kann es leicht zu einer Überlastung beim Rezipienten kommen.

Ein Hörtext ist für den Rezipienten zudem zunächst einmal ein mit Pausen unterbrochener Lautstrom, in dem insbesondere die Wortgrenzen nur sehr bedingt markiert sind. Die Segmentierung in Wörter stellt eine der Hauptschwierigkeiten beim Hören (fremdsprachiger) Texte dar – vor allem in einer "grenzsignalarmen Sprache" wie dem Französischen (vgl. Meißner 2006: 261). Die Schwierigkeiten werden noch verstärkt, wenn der Hörtext eine deutliche Tendenz zu Koartikulation, Assimilation, Reduktion und Elision aufweist (vgl. z.B. Field 2008: 27f.).

Weiterhin ist Hörverstehen eine zielgerichtete Aktivität, in dessen Verlauf der Hörer unterschiedliche und auch wechselnde Ziele verfolgen kann. Dabei kann der Fokus der Aufmerksamkeit eher global oder eher lokal sein. Zudem kann der Grad an Aufmerksamkeit eher gering oder eher hoch sein (vgl. Field 2008: 66). Bei einem Scanning z.B. ist die Aufmerksamkeit lokal gerichtet und insgesamt eher gering. Beim Verstehen der zentralen Aussagen richtet sich die Aufmerksamkeit dagegen eher global auf den gesamten Text und ist zudem eher hoch. Das jeweilige Hörziel steuert somit die Art und Weise, wie der Hörer den Text verarbeitet und damit die jeweilige spezifische Höraktivität/Hörstrategie. Entsprechend wird in der Literatur u.a. zwischen folgenden Höraktivitäten bzw. -strategien differenziert: detailliert, selektiv, global, orientierend, kursorisch, total (vgl. die Literaturbelege in Grotjahn 2005). Bei der Beschreibung der kognitiven Anforder-

rungen der eingesetzten Items in der empirischen Studie (vgl. Abschnitt 3.3) haben wir diese Unterscheidungen zumindest partiell berücksichtigt.

Weiterhin kann der Hörer in einer Kommunikationssituation eine Reihe unterschiedlicher Rollen einnehmen. Rost (2004) unterscheidet anhand des Ausmaßes an diskursiver Kooperation zwischen Sprecher und Hörer u.a. folgende Hörerrollen: a) Gesprächsteilnehmer; b) Adressat; c) Zuhörer (*auditor*); d) (zufälliger) Mithörer. In den im vorliegenden Beitrag verwendeten Testaufgaben hat der Lerner die Rolle eines Mithörers. Diese Rolle ist bei reinen Hörverstehensaufgaben nicht unproblematisch, da dem Mithörer nicht das Situations- und Sachwissen der Sprecher zur Verfügung steht (vgl. Brown 1995: 59f.). Befindet sich der Mithörer jedoch in einer Hör-Sehsituation, wie es bei der Rezeption von Videos zumeist der Fall ist, kann er zum einen die im Vergleich zum auditiven Kanal weit größere Kapazität des visuellen Kanals zusätzlich nutzen (vgl. z.B. Paschke 2001: 158f.). Zum anderen stehen über den visuellen Kanal weitere für das Verstehen wichtige Informationen zur Verfügung – z.B. in Bezug auf den situativen Rahmen oder auch bezüglich der durch Mimik und Gestik ausgedrückten Sprecherintentionen (vgl. auch Abschnitt 2.2). Inwieweit man vor diesem Hintergrund auch Items zulässt, die allein anhand der präsentierten visuellen Informationen gelöst werden können und damit eigentlich Sehverstehen testen, hängt von der Spezifikation des Testkonstrukts und den Zielen der Testung ab.

2.4 Determinanten der Schwierigkeit von Hörverstehensaufgaben

In den vorangehenden Abschnitten wurde bereits eine Reihe von potenziellen Determinanten der Schwierigkeit von Hörverstehensaufgaben genannt, ohne jedoch explizit zwischen bestimmten Merkmalsklassen zu unterscheiden.¹⁰ Mit Freedle & Kostin (1999) kann man zwischen folgenden drei Typen von Schwierigkeitsgenerierenden Merkmalen differenzieren: Reine Itemvariablen beziehen sich allein auf das jeweilige Item und nicht auf das Hördokument. Reine Dokumentvariablen charakterisieren allein den Inhalt und die Struktur des Hördokuments. Dokument-Item-Variablen beziehen sich dagegen sowohl auf das jeweilige Hördokument als auch auf die zugehörigen Items.

Beispiele für reine Itemvariablen sind das Itemformat (z.B. Mehrfachauswahl-Items vs. Kurzantwort-Items), die (syntaktische) Komplexität der Itemformulierung, die Zahl der unvertrauten Wörter in den Distraktoren oder auch der

¹⁰ Die folgenden Ausführungen zu Schwierigkeitsfaktoren sind weitgehend aus Grotjahn (2003: 122-124, 132-136) entnommen. Eine tiefergehende Darstellung findet sich in Grotjahn (2000). Weitere Hinweise zur Determination der Aufgabenschwierigkeit von Lese- und Hörverstehensaufgaben finden sich u.a. in Grotjahn (2005), Grotjahn & Tesch (2010a; 2010b), Kostin (2004), Nold & Rossa (2007a; b) sowie Rupp, Garcia & Jamieson (2001).

Zeitdruck bei der Beantwortung des Items. In Bezug auf das Itemformat gilt z.B., dass im Fall von Hörverstehensaufgaben Kurzantwort-Items tendenziell (deutlich) schwerer sind als analoge Mehrfachauswahl-Items (vgl. In'nami & Koizumi 2009). Beispiele für Dokumentvariablen sind die Komplexität der Satzstruktur, die Textkohärenz oder auch die Komplementarität von akustischer und visueller Information. Eine besonders wichtige Dokument-Item-Variable ist, ob die zur Lösung benötigte Information über das Hördokument verteilt ist. Ist Letzteres der Fall, ist das Item tendenziell schwieriger, als wenn sich die relevante Information z.B. am Anfang oder Ende des Dokuments befindet.

Brown (1995) benennt weitere wichtige Dokumentvariablen und weist darauf hin, dass transaktionale Hörtexte, d.h. Texte, die auf die Übermittlung von Sachinformationen zielen, dann vergleichsweise leicht zu verstehen sind, wenn

1. sie eine geringe Zahl von Referenten (Individuen, Objekten) aufweisen;
2. die Referenten klar voneinander unterschieden sind;
3. die Ortsrelationen wenig komplex sind;
4. die Erzählreihenfolge dem Ablauf der Ereignisse in der Realität entspricht;
5. für das Verständnis wenige Inferenzen notwendig sind;
6. die Inhalte präzise formuliert, untereinander konsistent und auf bereits vorhandene Informationen beziehbar sind.

Eine relativ umfangreiche, empirisch gestützte Liste von Item-, Dokument- und Dokument-Item-Variablen findet sich bei Grotjahn (2000: 47; 2005: 125f.). Es handelt sich allerdings um Merkmale, die sowohl für das Lese- als auch das Hörverstehen gelten. Eine nützliche Liste von Schwierigkeitsfaktoren, die speziell aus der akustischen Komponente des Hörvorgangs resultieren, hat Solmecke (2000) zusammengestellt. Diese ist in Tabelle 2 wiedergegeben. Weitere Hinweise zu potenziellen Schwierigkeitsfaktoren von Hördokumenten finden sich z.B. im Standardwerk von Buck (2001). Dort werden u.a. folgende Merkmale unterschieden (ebd., S. 31-60): Phonologische Modifikationen; Akzente und Dialekte; Prosodie; Sprechgeschwindigkeit; Verzögerungen und Pausen; Diskursstruktur; nonverbale Signale. Eine Liste von Selektionskriterien für "Hör-Seh-Texte" findet sich bei Thaler (2007b: 15).

Textschwierigkeit						
	← eher leicht					eher schwierig →
Akustische Komponente		1	2	3	4	
Störgeräusche	Leise					Laut
Raumakustik	Gut					Schlecht
Aufnahmequalität	Gut					Schlecht
Sprechgeschwindigkeit	Langsam					Schnell
Artikulation	Deutlich					Undeutlich
Intonation	Gliedernd					Monoton
Pausen zwischen Sätzen und Satzteilen	Eher länger					Eher kürzer
Dialekt-, Regiolekt-, Soziolektfärbung	Keine					Stark
Sonstige Abweichungen vom Standard	Keine					Starke
Zahl der Sprecher/innen	Eine(r)					Mehrere
Unterscheidbarkeit der Stimmen im Dialog	Problemlos					Schlecht
Sprecherwechsel überlappend	Nie					Häufig
Zwischenrufe/sonstige Unterbrechungen	Nie					Häufig
Präsentationshäufigkeit	Mehrmals					Einmal

Tabelle 2: Schwierigkeitsfaktoren der akustischen Komponente
(Solmecke 2000: 63)

Eine wichtige Schwierigkeitsdeterminante ist die Präsentationshäufigkeit des Hördokuments. Die meisten Dokumente werden in der Realität nur einmal gehört. Insbesondere beginnende Lerner sind jedoch bei nur einmaligem Hören nicht selten deutlich überfordert. Ein zweimaliges Hören hilft zudem, den potenziellen Einfluss von Störvariablen zu minimieren. So kann es vorkommen, dass Lerner ein Item nicht adäquat bearbeiten, weil sie noch dabei sind, ein vorangehendes Item zu lösen. Ein zweites Hören bedeutet in diesem Fall eine zweite Chance (vgl. auch Geranpayeh & Taylor 2008: 3f.; Paschke 2001: 159f.; Sakai 2009). Bei zweimaliger Präsentation des Hörtextes besteht zudem die Möglichkeit, dass sich die Lerner zunächst das Item anschauen, dann das Dokument hören und erst nach oder während des zweiten Hörens ihren Eintrag machen. Dadurch reduziert sich die potenzielle Beeinträchtigung des Hör- oder auch des Hör-Sehverstehens durch das Lesen und Schreiben, was für die Beantwortung der Testitems jedoch notwendig ist (z.B. Lesen der Optionen von Mehrfachwahlantworten, Kurzantworten schreiben).

2.5 Strategien bei der Bearbeitung von Testaufgaben zum Hörverstehen und zum Hör-Sehverstehen

Hör-Sehverstehensstrategien lassen sich horizontal – hinsichtlich der beiden beteiligten Kanäle sowie ihrer Überlappung – und vertikal – hinsichtlich der Hierarchietiefe der beteiligten Verarbeitungsprozesse – differenzieren. Mit Blick auf die beteiligten Wahrnehmungskanäle können Strategien des Hörverstehens und Strategien des Sehverstehens unterschieden werden. Bezogen auf das Hörverstehen lassen sich allgemeine bzw. globale Verstehensstrategien (z.B. Hörerwartungen auf Grund des Themas, der situativen Einbettung und der Textsorte aufbauen; auf Textstrukturwissen zurückgreifen; Intonation als Verstehenshilfe z.B. für Gemütsverfassungen nutzen) von Sinnerfassungsstrategien auf Laut-, Phonem- und Wort- bzw. Sequenzebene (z.B. Anwendung von Wortfeldwissen, Erfassen von Schlüsselwörtern, Kontextnutzung) unterscheiden.

Spezifische Sehstrategien sind dagegen die Nutzung der Bildinformation zum Verständnis der Situation und zur Identifikation von Personen, Objekten und Handlungen. Ein Sonderfall der Sehstrategien wäre die Nutzung von Leseinformationen im Bild (Vorspann, Lesetext im Bild, Titel und Aufschriften, Abspann). Aus dieser kurzen Zusammenstellung ergibt sich bereits das Forschungsinteresse an Informationen über Zusammenhänge bei der Nutzung des auditiven und visuellen Kanals sowie insbesondere an Befunden, die das mögliche Zusammenwirken von Strategien belegen könnten. Können beim Hör-Sehverstehen spezifische Strategien vermutet werden, Strategien also, die erst in der Kombination beider Kanäle eingesetzt werden, wie zum Beispiel der gegenseitige Abgleich von Hörerwartungen mit Bildinformationen im Sinne einer Deutungsvalidierung? Könnte auch der gegenteilige Effekt eintreten, nämlich einer Schwächung der Wirksamkeit von Hörstrategien durch eine Überlastung des Arbeitsspeichers durch gleichzeitiges Sehen bzw. von Sehstrategien durch die Konzentration auf das Hören?

Wie oben erwähnt lassen sich Strategien auch vertikal entsprechend einer angenommenen Hierarchie von Verstehensleistungen theoretisch klassifizieren. So verlangen handlungsarme und stark mit abstrakten bzw. impliziten Informationen besetzte Hör-Sehsequenzen sicherlich in wesentlich höherem Maße die Aktivierung von Vorwissen (z.B. kontext- und situationsbezogenes Weltwissen) als handlungsreiche Sequenzen, die bereits durch die visuelle Identifizierung von Personen, Objekten und Handlungen leicht zu verstehen sind.

Von besonderem Interesse im Hinblick auf die Fremdsprachenforschung könnten Studien sein, die Verstehensleistungen beim erstsprachlichen Hör-Sehverstehen in Beziehung zu Verstehensleistungen beim fremdsprachlichen Hör-Sehverstehen setzen. Es ist zu vermuten, dass bei Letzterem eine spezifische interkulturelle Verstehensleistung erbracht wird, in der die fremdsprachliche und

die fremdkulturelle Dekodierung zusammenwirken und in der Erwerbsperspektive zu einem zusätzlichen Strategienaufbau führen. Mit anderen Worten, es ist anzunehmen, dass sprachliche und visuelle Strategien beim fremdsprachlichen Hör-Sehverstehen im Sinne einer audio-visuellen fremdsprachlichen (Interims-) Strategie zusammenwirken.

Das Spezifikum von Strategien beim Lösen von Hör-Seh-Testaufgaben liegt, wie auch den in Abschnitt 2.4 genannten Ausführungen zu schwierigkeitsgenerierenden Merkmalen zu entnehmen ist, vor allem im Abgleich der gehörten Lexik in der Textvorgabe, gegebenenfalls auch gelesener Lexik in Lesetextelementen sowie der visuellen Identifikation von Personen, Objekten und Handlungen mit der gelesenen Lexik im Item. Auch allgemeine Testbearbeitungsstrategien wie etwa das Ausschließen bestimmter Distraktoren auf Grund kombinatorischer Überlegungen sind relevant (vgl. z.B. Kleppin & Reich 2010).

2.6 Hör-Sehverstehen in den Bildungsstandards für die erste Fremdsprache und im Gemeinsamen europäischen Referenzrahmen für Sprachen

In den vorangehenden Abschnitten wurden insbesondere folgende Merkmale als konstitutiv für das Hör-Sehverstehen in einer Fremdsprache herausgestellt: (1) Hör-Sehverstehen im Sinne der Verarbeitung bewegter audiovisueller Bilder unterscheidet sich vom reinen Hörverstehen durch die simultane Verarbeitung visuellen und auditiven Inputs. (2) Die Aufnahme von kombinierten Text-Bildinformationen kann die Verarbeitung bzw. das Lernen unter bestimmten Voraussetzungen erleichtern oder erschweren. Es ist dabei davon auszugehen, dass schwierigkeitsgenerierende Determinanten des Hörverstehens in einer Fremdsprache in der Regel auch für die Rezeption von fremdsprachlichen Videos gelten, wenn diese neben visuellem auch auditives Material anbieten. Ferner können die Bilder weitere Schwierigkeiten in der Interaktion mit dem L2-Rezipienten verursachen, z.B. aufgrund einer mangelhaften Vertrautheit mit dem Format oder der Schnitttechnik.

Das erfolgreiche Verstehen von Informationen in einer Fremdsprache, die auditiv oder audio-visuell dargeboten werden, hängt jedoch nicht allein von der Kodierungsform der Information ab, sondern vom Sprachstand und anderen Merkmalen der Rezipienten sowie von der Art der geforderten kognitiven Leistung (z.B. Detail- oder Globalverstehen) bzw. dem zugrunde gelegten Konstrukt. Um die standardbasierte Entwicklung von Testaufgaben zum Hörverstehen und Hör-Sehverstehen auf der Grundlage der Bildungsstandards für die erste Fremdsprache (KMK 2003) im Fach Französisch in der Sekundarstufe I nachvollziehen

zu können, sollen nachfolgend die relevanten Dokumente analysiert werden, um die curricularen Anforderungen an die Fremdsprachenlerner näher zu bestimmen.

Die Erwartungen an deutsche Schülerinnen und Schüler im Hinblick auf ihren Kompetenzstand am Ende der Sekundarstufe I finden sich in Form von Kann-Beschreibungen in den länderübergreifend geltenden Bildungsstandards für die erste Fremdsprache wieder. Diese stellen zudem in dieser Altersgruppe z.Zt. die Referenz für die Entwicklung von standardbasierten Test- und Lernaufgaben im Fach Englisch und Französisch dar. Die Standards für die erste Fremdsprache in der Sekundarstufe I liegen in zwei Versionen vor: für den Hauptschulabschluss (HSA; KMK 2004) und den Mittleren Schulabschluss (MSA; KMK 2003), wobei wir uns im Folgenden auf die Ausführungen zum MSA beziehen. Es soll der Frage nachgegangen werden, welche Anforderungen an die Fremdsprachenlerner zu diesem Zeitpunkt im Bereich des Hör-Sehverstehens seitens der KMK gestellt werden. Da die Bildungsstandards für die erste Fremdsprache an das Kompetenzstufensystem des Gemeinsamen europäischen Referenzrahmens für Sprachen (GeR; Europarat 2001) angebunden sind¹¹, werden die Beschreibungen und der Stellenwert des Hör-Sehverstehens im Vergleich zu den Ausführungen in den Bildungsstandards betrachtet.

Die Bildungsstandards benennen drei zentrale Kompetenzbereiche: funktionale kommunikative Kompetenzen, interkulturelle Kompetenzen und methodische Kompetenzen. Die Darstellungsweise impliziert jedoch nicht, dass die Bereiche als unabhängig voneinander anzusehen sind (vgl. KMK 2003: 8). Die Bildungsstandards unterscheiden zudem fünf kommunikative Teilkompetenzen: Leseverstehen, Sprechen (an Gesprächen teilnehmen, zusammenhängendes Sprechen), Schreiben, Sprachmittlung und Hör- und Hör-Sehverstehen. Die gemeinsame Nennung des Hör- und Hör-Sehverstehens in der Übersicht zu den Standards wird in den Deskriptoren fortgeführt. Die entsprechende Passage findet sich in der Tabelle 3. Die Deskriptoren beziehen sich auf die GeR-Niveaus B1 und B1+ und charakterisieren somit eine selbstständige Sprachverwendung. Als Stimuli werden Gespräche, Vorträge, Interviews, Ankündigungen und Mitteilungen, Filme, Nachrichtensendungen, Radionachrichten und Tonaufnahmen sowie Fernsehsendungen aufgeführt, wobei einige in Bezug auf die Länge (z.B. "längere Gespräche", "kurze Vorträge") eher ungenau spezifiziert werden. Das Verstehen bezieht sich auf "unkomplizierte Sachinformationen über gewöhnliche alltags- oder berufsbezogene Themen" oder "konkrete Themen". Französischlerner sollen auf dem Niveau B1/B1+ in der Fremdsprache Aufnahmen zu "vertrauten Themen", Englischlerner zu "Themen von persönlichem Interesse" verstehen können. Entsprechende Do-

11 Wir beschränken uns im Folgenden auf eine Darstellung thematisch relevanter Aspekte der Anbindung der Bildungsstandards an den GeR. Eine kritische Auseinandersetzung mit weiteren Aspekten findet sich z.B. bei Quetz & Vogt (2009).

kumente können verstanden werden, wenn die Sprecher "in deutlich artikulierter Standardsprache" und in "normaler Geschwindigkeit" sprechen. Die Französischler verstehen diese Tonaufnahmen, wenn "relativ langsam und deutlich gesprochen" wird.

Die Schülerinnen und Schüler können unkomplizierte Sachinformationen über gewöhnliche alltags- oder berufsbezogene Themen verstehen und dabei die Hauptaussagen und Einzelinformationen erkennen, wenn in deutlich artikulierter Standardsprache gesprochen wird (B1+).

Die Schülerinnen und Schüler können (Englisch und Französisch)

- im Allgemeinen den Hauptpunkten von längeren Gesprächen folgen, die in ihrer Gegenwart geführt werden (B1),
- Vorträge verstehen, wenn die Thematik vertraut und die Darstellung unkompliziert und klar strukturiert ist (B1+),
- Ankündigungen und Mitteilungen zu konkreten Themen verstehen, die in normaler Geschwindigkeit in Standardsprache gesprochen werden (B2),
- vielen Filmen folgen, deren Handlung im Wesentlichen durch Bild und Aktion getragen wird (B1).

Englisch:

Die Schülerinnen und Schüler können

- den Informationsgehalt der meisten Rundfunksendungen und Tonaufnahmen über Themen von persönlichem Interesse verstehen (B1+),
- das Wesentliche in vielen Fernsehsendungen zu Themen von persönlichem Interesse, z.B. Interviews, kurze Vorträge oder Nachrichtensendungen verstehen (B1+).

Französisch:

Die Schülerinnen und Schüler können

- in Radionachrichten und in einfacheren Tonaufnahmen über vertraute Themen die Hauptpunkte verstehen, wenn relativ langsam und deutlich gesprochen wird (B1),
- das Wesentliche von Fernsehsendungen zu vertrauten Themen verstehen, sofern darin relativ langsam und deutlich gesprochen wird (B1).

Tabelle 3: Hör- und Hör-/Sehverstehen in den Bildungsstandards für die erste Fremdsprache (Mittlerer Schulabschluss; KMK 2003: 11f.)

Im GeR sind den kommunikativen Sprachaktivitäten und -strategien produktive und rezeptive Aktivitäten und Strategien zugeordnet. Das Lese- und Hörverstehen wird im Referenzrahmen den rezeptiven Fertigkeiten zugeordnet. Hörverstehen wird wie folgt definiert: "Bei auditiven rezeptiven Aktivitäten (beim Hören) empfangen und verarbeiten Sprachverwendende als Hörer einen von einem oder mehreren Sprechern produzierten gesprochenen Input" (Europarat

2001: 71). Der GeR unterscheidet das Hören von Aufzeichnungen (Radio, Fernsehen, öffentliche Ansagen etc.) von Live-Gesprächen (als Zuhörer im Theater, bei Vorträgen etc. und als "zufälliger" Mithörer bei Gesprächen). Die aufgeführten Medien deuten bereits darauf hin, dass unter Hörverstehen auch das Hörsehverstehen gefasst wird, da u.a. Fernsehsendungen genannt sind. Auch die Skalenbeschreibungen zum "Hörverstehen allgemein" zeigen, dass nicht allein das Verstehen von auditivem Material gemeint ist. So beginnt die Beschreibung zum Niveau B2 wie folgt: "Kann im direkten Kontakt und in den Medien gesprochene Standardsprache verstehen ..." (Europarat 2001: 72).

Nach der Überschrift zum Abschnitt "Rezeptive Aktivitäten und Strategien" im GeR erfolgt der Hinweis "Rezeptive Aktivitäten umfassen Hören und Lesen" (Europarat 2001: 71). Der Beschreibung des Hör- und Leseverstehens folgen jedoch weitere Abschnitte (1) zur "audiovisuellen Rezeption" und (2) zu "Rezeptionsstrategien". Bei der "audiovisuellen Rezeption empfangen Sprachverwendende gleichzeitig einen auditiven und einen visuellen Input" (Europarat 2001: 77). Die folgenden Aktivitäten werden als beispielhaft genannt:

- einen vorgelesenen Text mitlesen
- eine Fernsehsendung oder Videoaufzeichnung oder einen Film mit Untertiteln ansehen
- neue Technologien anwenden (Multimedia, CD-ROM, usw.)

Die kognitiven Aktivitäten, die vom Rezipienten vollzogen werden müssen, werden nicht wie im Abschnitt zum Hörverstehen explizit vor den Skalen benannt (global verstehen bzw. erfahren, was insgesamt gemeint ist; selektiv verstehen bzw. eine ganz bestimmte Information erhalten; detailliert verstehen bzw. das Gesprochene in allen Einzelheiten verstehen; Schlussfolgerungen ziehen), können jedoch den Deskriptoren der Skala "Fernsehsendungen und Filme verstehen" (vgl. Tabelle 4) entnommen werden. Sie beschreiben die Fähigkeiten in diesem Bereich mit den Wendungen "einen großen Teil" und "das Wesentliche" verstehen (B1/B1+) sowie "die Hauptinformation erfassen" (A2+).

C2/C1
Kann Spielfilme verstehen, auch wenn viel saloppe Umgangssprache oder Gruppensprache und viel idiomatischer Sprachgebrauch darin vorkommt.
B2
Kann im Fernsehen die meisten Nachrichtensendungen und Reportagen verstehen.
Kann Fernsehreportagen, Live-Interviews, Talk-Shows, Fernsehspiele sowie die meisten Filme verstehen, sofern Standardsprache gesprochen wird.
B1
Kann in vielen Fernsehsendungen zu Themen von persönlichem Interesse einen großen Teil verstehen, z.B. in Interviews, kurzen Vorträgen oder Nachrichtensendungen, wenn relativ langsam und deutlich gesprochen wird.
Kann vielen Filmen folgen, deren Handlung im Wesentlichen durch Bild und Aktion getragen wird und deren Sprache klar und unkompliziert ist. Kann das Wesentliche von Fernsehprogrammen zu vertrauten Themen verstehen, sofern darin relativ langsam und deutlich gesprochen wird.
A2
Kann die Hauptinformation von Fernsehmeldungen über Ereignisse, Unglücksfälle usw. erfassen, wenn der Kommentar durch das Bild unterstützt wird.
Kann dem Themenwechsel bei TV-Nachrichten folgen und sich eine Vorstellung vom Hauptinhalt machen.
A1
Keine Deskriptoren vorhanden

Tabelle 4: Skala "Fernsehsendungen und Filme verstehen" des GeR
(Europarat 2001: 77)

Diese Skala wird explizit als ein *illustratives Beispiel* für die audiovisuelle Rezeption genannt und ist auch im Anhang der Bildungsstandards für die erste Fremdsprache (KMK 2003: 74f.) wiedergegeben. An den Deskriptoren, die für die Niveaus A2 bis C2 vorliegen, wird erkennbar, dass eine Differenzierung der Stufen vor allem über die folgenden Merkmale erfolgt: Sprachregister, Geschwindigkeit und Artikulation der Sprecher, Strukturierung und Thematik des Beitrags, (geforderte) kognitive Aktivitäten des Rezipienten sowie Art des Beitrags. In der Skala finden sich zwei Aussagen über das Verhältnis Bild zu Text wieder: Die Hauptinformation von Fernsehsendungen kann verstanden werden, "wenn der Kommentar durch das Bild unterstützt wird" (A2), und der Rezipient "kann vielen Filmen folgen, deren Handlung im Wesentlichen durch Bild und Aktion getragen wird" (B1).

Der GeR liefert einen zusätzlichen Abschnitt zur audio-visuellen Rezeption und definiert diese als die gleichzeitige Rezeption auditiven und visuellen Inputs. Medienpsychologische Erkenntnisse, in welcher Weise dieser Input auf welcher Kompetenzstufe beschaffen sein muss, um erfolgreich verarbeitet zu werden, werden jedoch auch im GeR nicht ausführlich beschrieben. Deskriptoren werden

lediglich für das Verstehen von Film- und Fernsehsendungen angeboten, auf die auch die Bildungsstandards zurückgreifen.

Die Analyse der Kann-Beschreibungen und der Verortung des Hör-Sehverstehens im GeR und den Bildungsstandards für die erste Fremdsprache geben Anlass zu folgenden Schlussfolgerungen:

- Das Verhältnis von Text- und Bildinformationen sowie die Berücksichtigung von (möglichem) zusätzlichem auditiven Input wie Hintergrundmusik, wie es beispielsweise in Fernsehfilmen oder anderem authentischen Material vorkommen kann, werden nicht hinlänglich in den Dokumenten in Bezug auf die Verarbeitung und das Verständnis der Rezipienten auf den einzelnen Niveaustufen beschrieben.
- Die Generierung von Standards, die Kompetenzen in einer Fremdsprache beschreiben sollen, stellt im Bereich des Hör-Sehverstehens eine besondere Herausforderung dar. Dies gilt insbesondere im Hinblick auf die Beschreibung von Hör-Sehverstehen in (möglicher) Abgrenzung zu Film-/Medienkompetenz, Hörverstehen sowie weiteren beteiligten Teilkompetenzen (z.B. Leseverstehen, Arbeitsgedächtnisleistungen). Notwendig ist, neben den bereits vorliegenden Kann-Beschreibungen eine Konstruktbeschreibung anzubieten, welche die genannten Aspekte berücksichtigt.
- Dieser Herausforderung begegnen die beiden Dokumente auf verschiedene Weise. Die Verfasser der Bildungsstandards für die erste Fremdsprache beziehen sich vor allem auf die Beschreibungen und Skalen des GeR zur auditiven Rezeption. Als Folge werden vor allem die sprachlichen (Hörverstehens-) Leistungen fokussiert. Die Ausführungen zum Hör-Sehverstehen erschöpfen sich in der Auflistung der GeR-Skala "Fernsehsendungen und Filme verstehen" in einem Anhang. Spezifische Einflüsse kombinierter Text- und Bildinformationen auf den Verstehensprozess werden nicht weiter thematisiert.
- Auch der GeR betont in seiner Darstellung die Gemeinsamkeiten von Hör- und Hör-Sehverstehensprozessen, indem im Rahmen der Beschreibung des Hörverstehens bildgestützte Medien wie das Fernsehen benannt werden.

3. Die empirische Studie

3.1 Fragestellungen und Hypothesen

Das Hör-Sehverstehen bzw. Verstehen von Videos unterscheidet sich vom Hörverstehen durch die (mögliche) Nutzung sowohl auditiven als auch visuellen Inputs. Dabei ist zu erwarten, dass Lern- bzw. Testleistungen tendenziell höher für das Hör-Sehverstehen als für das Hörverstehen sind, wenn die visuell präsen-

tierten Informationen eine Hilfe darstellen (vgl. Abschnitt 2.1 und 2.2). Im Rahmen einer auf den Bildungsstandards für die erste Fremdsprache (KMK 2003) basierenden Aufgabenentwicklung – wobei die Standards die beiden Kompetenzen gemeinsam beschreiben (vgl. Abschnitt 2.6) – stellt sich die Frage, ob sich die beiden Konstrukte empirisch in der Schülerpopulation der Bildungsstandards trennen lassen. Unsere erste allgemeine Forschungsfrage lautet daher:

1. Zeigen Fremdsprachenlerner in einem standardisierten Leistungstest im Fach Französisch höhere Verstehensleistungen, wenn sie begleitend zum Hörtext unterstützendes Bildmaterial präsentiert bekommen?

Weiterhin sollten Antworten auf folgende spezifische Forschungsfragen gefunden werden:

2. Welche Schwierigkeiten haben Fremdsprachenlerner beim Lösen von Hör- und Hör-Sehverstehens(test)aufgaben und welche Strategien wenden sie an?
3. Unterscheiden sich die Untersuchungsgruppen und die einzelnen Aufgaben in Bezug auf die Zahl der genannten Schwierigkeiten und Lösungsstrategien?

Zur Beantwortung dieser Fragen mit Hilfe einer Experimentalstudie wurden von uns aus dem zur Überprüfung der Bildungsstandards Französisch entwickelten Itempool drei unseres Erachtens geeignete Testaufgaben ausgewählt mit den Titeln *Alimentation*, *Eau de mer* und *Portables*. Unter Berücksichtigung der eingesetzten Instrumente wurden dann folgende aufgabenspezifische Hypothesen in Bezug auf das jeweilige Ausmaß an Hilfe durch den visuellen Kanal formuliert:

1. *Alimentation*

Der visuelle Kanal stellt keine Hilfe dar: Die Argumentationskette wird nicht durch eine entsprechende Bildabfolge gestützt; die Bilder liefern nur eine vordergründige Situierung; Hintergrundinformationen können nur durch Hören erschlossen werden.

2. *Portables*

Der visuelle Kanal stellt eine Hilfe dar: Die Handlungssequenz ist teilweise auf die Bildinformation hin konzipiert; der Bildtext erläutert teilweise den Redetext.

3. *Eau de mer*

Der visuelle Kanal stellt eine Hilfe dar: Das Fachvokabular wird durch Bilder illustriert und die Argumentationskette wird durch die Bildabfolge gestützt.

3.2 Stichprobe

Insgesamt nahmen 156 Gymnasiasten der Klassenstufe 9 aus vier Bundesländern (Berlin, Hessen, Niedersachsen, Rheinland-Pfalz) an der Untersuchung teil. Die Schulen waren Partner des Französischprojekts am Institut zur Qualitätsentwicklung im Bildungswesen (IQB). 90,8 Prozent der untersuchten Schüler erlernten Französisch als erste Fremdsprache und 98,7 Prozent sprachen laut eigenen Angaben normalerweise zu Hause Deutsch. Für die experimentelle Studie wurden per Zufall die intakten Lerngruppen bzw. Klassen den Untersuchungsbedingungen zugeteilt. Die schulorganisatorischen Rahmenbedingungen für die Testungen erlaubte keine individuelle Zuweisung der Schüler zu den Gruppen. Bewusst sollten die Testungen in dieser Studie unter vergleichbaren Bedingungen wie in den Schulleistungsstudien des IQB stattfinden, so dass Schlussfolgerungen für den Einsatz der Instrumente in weiteren Untersuchungen gezogen werden können. Die Tabelle 5 zeigt die jeweiligen Gruppengrößen und die verwendeten Bezeichnungen.

Gruppe 1	Hör-Sehverstehen (HSV)	$N_1 = 68$
Gruppe 2	Hörverstehen (HV)	$N_2 = 88$

Tabelle 5: Realisierte Stichprobe nach Gruppen

3.3 Instrumente und Ablauf der Untersuchung

Folgende Instrumente wurden in den einzelnen Untersuchungsphasen eingesetzt:

1. Prätest: Beide Gruppen bearbeiteten den gleichen Hörverstehenstest mit einer Gesamtzeit von 20 Minuten. Dazu wurde ein intakter Testblock von Hörverstehensaufgaben mit sechs Items aus der bereits durchgeführten Normierungsstudie im Fach Französisch (Sekundarstufe I) ausgewählt. Jeder Stimulus wurde zweimal vorgespielt (genaue Beschreibungen des Testkonstrukts zum Hörverstehen im Anhang von Porsch, Tesch & Köller 2010).
2. Testung in den zwei Gruppen: Die beiden Gruppen bearbeiteten einen weiteren Hör- bzw. Hör-Sehverstehenstest mit einer Gesamtzeit von 20 Minuten. Die Testhefte waren identisch, d.h. sie enthielten die gleichen Items und Instruktionen. Die beiden Gruppen unterschieden sich jedoch im angebotenen Stimulus bzw. im Hinblick auf das Medium. Der Gruppe

- 1 (HSV) wurden Videos mit Ton präsentiert, der Gruppe 2 (HV) lediglich der Ton. Jeder Stimulus wurde zweimal vorgespielt.¹²
3. Selbsteinschätzung der Schüler im Hinblick auf ein Item je Aufgabe: Allen Schüler wurden nach dem Einsammeln der Testhefte noch einmal die Input-Texte der drei Aufgaben (die reinen Hörtexte bzw. Videos) zusammen mit einem in Bezug auf die zu überprüfenden Hypothesen relevanten Item präsentiert (*stimulated recall*). Folgende beiden Fragen sollten dazu von den Schülern schriftlich beantwortet werden: (a) Welche Schwierigkeiten hattest du beim Lösen? (b) Wie bist du beim Bearbeiten vorgegangen?
 4. Beantwortung eines Fragebogens: Alle Lerner beantworteten Fragen zu ihrem Hintergrund (Geschlecht, Alter, Lernjahre in Französisch, Teilnahme am bilingualen Unterricht, gesprochene Sprache im Elternhaus bzw. weitere Sprachen). Die Informationen wurden für die Beschreibung der Stichprobe herangezogen und dienten zudem der Überprüfung der Vergleichbarkeit der beiden Gruppen.¹³

Die Hör- und Hör-Sehverstehensaufgaben stellen erprobte Aufgaben aus dem Itempool dar, der am IQB zur Testung der Schüler auf Grundlage der Bildungsstandards für die erste Fremdsprache sowie des GeR entwickelt wurde (vgl. Porsch, Tesch & Köller 2010). Die Videosequenzen entsprechen thematisch dem Interesse von Neuntklässlern. Die entwickelten Aufgaben und Items erlauben zwar alle das Verstehen auch ohne das Bild, bei einer genauen Analyse wird jedoch deutlich, dass für einige Items das Bild eine mehr oder minder deutliche Unterstützung darstellt. Diese Analysen bildeten die Grundlage für die in Abschnitt 3.1 aufgeführten drei aufgabenspezifischen Hypothesen in Bezug auf das jeweilige Ausmaß an Hilfe durch den visuellen Kanal. In Tabelle 6 werden Grundinformationen zu den ausgewählten Aufgaben bzw. Videos – alles Auschnitte aus dem französischen Fernsehen – und in Tabelle 7 die drei Items näher charakterisiert, auf die sich die genannten Hypothesen beziehen.

12 Aus den in Abschnitt 2.2 genannten Gründen werden die zur Überprüfung der Bildungsstandards entwickelten französischen Hörverstehens- und Hör-Sehverstehensaufgaben unter Einschluss der Items stets zweimal präsentiert.

13 Außerdem wurde eine adaptierte Version des *Perceptual Learning Style Preference Questionnaire* von Joy Reid eingesetzt (vgl. Reid 1995: 202-207). Mit Hilfe dieses Instruments konnte jedoch kein Einfluss von perzeptuellen Lernstilpräferenzen bei der Verarbeitung visueller und auditiver Informationen nachgewiesen werden (vgl. auch die Untersuchung von Chung 1994). Wir gehen deshalb nicht weiter auf diese Teilstudie ein.

	<i>Alimentation</i>	<i>Eau de mer</i>	<i>Portables</i>
Anzahl / Format der Items	1x Multiple-Choice 1x Kurzantwort	4x Kurzantwort	3x Multiple-Choice 1x Kurzantwort
Adressaten	Kinder	thematisch Interessierte	unspezifisch
Textart	Pädagogischer Dialog	Dokumentation	Nachrichten
Dauer des Videos	60 Sek.	1 Min. 36 Sek.	2 Min.
Thema der Aufgabe	Sammeln von Lebensmitteln für Hilfsorganisationen	Entsalzung von Meerwasser	Exzessive Mobil- funknutzung Jugendlicher
Hintergrund- geräusche	Rauschen, Musik im Vor-/ Abspann	musikalische Unterlegung	Umweltgeräusche
Alter und Anzahl der Sprecher	1 Frau (30-40 J.), 1 Kind + 2 gezeichnete Kinder	1 erwachsene Frau (off)	1 Nachrichten- sprecher, 1 Sprecher off, 2 Erwachsene, 1 Jugendlicher
Geschlecht der Sprecher	weiblich	weiblich	4 männlich, 1 weiblich
Akzente / Sprech- weise / Register	keine Akzente, standardsprachlich, Kinderstimmen	kein Akzent, standardsprachlich	keine Akzente, standardsprachlich
Fachvokabular	nein	ja	nein
Darstellungsweise	Realfiguren in ge- zeichneter Umge- bung (Supermarkt) + gezeichnete Figuren	Realaufnahmen	Realaufnahmen

Hinweis: Multiple-Choice-Items umfassen vier Optionen, davon eine richtige Lösung.

Tabelle 6: Merkmale der Aufgaben im 2. Untersuchungsteil

	<i>Alimentation</i>	<i>Eau de mer</i>	<i>Portables</i>
Itemnummer	Item 1	Item 1	Item 3
Itemformat	Multiple-Choice	Kurzantwort (Satzanfang beenden)	Multiple-Choice
GeR-Niveau	B1	B1	B2
kognitive Anforderung	Erfassen einer Hauptaussage: Grund für Spendenaktion soll verstanden wer- den	Identifikation einer Detailaussage: Ver- stehen, dass drei Viertel des trinkba- ren Wassers aus Eis sind	Erfassen einer Hauptaussage: was für einen Service <i>contrôle parental</i> bietet
Bild als Unterstützung zum Verstehen?	Antwort kann ledig- lich den Aussagen der Frau entnommen werden; erfragte Information bezieht sich nicht auf den Ort der Handlung	Antwort wird von off-Sprecherin genannt; während der Nennung sieht man Gletscher	Antwort wird von off-Sprecherin ge- nannt; während der Nennung sieht man Handy mit einer "verbotenen Internet- seite"

Tabelle 7: Merkmale des zu jeder Aufgabe ausgewählten Items (2. Testteil)

Generell standen Aufgaben zum Hör-Sehverstehen – im Folgenden auch als AV-Aufgaben (audiovisuelle Aufgaben) bezeichnet – ab Niveau B1 zur Verfügung. Diese Niveauzuordnung bezieht sich auf die Ergebnisse der Normierungsstudie zu den Aufgaben der Bildungsstandards und der Einordnung auf der Kompetenzstufenskala zum Hörverstehen (vgl. Porsch, Tesch & Köller 2010). Verstehensleistungen, die typisch für das Hör-Sehverstehen sind, wurden allerdings im Rahmen dieser Standardsetzung nicht berücksichtigt.

Die Aufgabe *Alimentation* ist in mehrfacher Hinsicht von eher geringer Komplexität. Die Sequenz besteht aus der Darbietung einer detailliert vorbereiteten, nicht spontan gesprochenen Sequenz in Dialog-Form. Thema der Sequenz sind die nationalen und staatlichen Hilfsorganisationen (*associations*) für Bedürftige bzw. Hungernde. Das Format entspricht Kinder- und Jugendsendungen, in denen häufig unterschiedliche Wissensgebiete angesprochen und kinder- bzw. jugendnah vermittelt werden, hier z.B. mittels der Verwendung eines Zeichentrickfilms als Hintergrund für den Dialog zweier "realer" Personen. Zusätzlich sprechen zwei weitere gezeichnete Personen. Es erscheint naheliegend, im Hinblick auf die hier verwendete Videosequenz von einem klaren Überhang an *context visuals* (vgl. Abschnitt 2.2) zu sprechen, während zum zentralen Thema der Hilfe für Hungernde bzw. Bedürftige kaum oder gar keine *content visuals* dargeboten werden. Die Dialoge vermitteln explizite Hintergrundinformationen zum besagten Thema, wobei das Kind Fragen stellt und die erwachsene Person antwortet bzw.

die Zusammenhänge erläutert. Das Thema wird linear entwickelt, von der Benennung konkreter Gegenstände für eine Lebensmittelsammlung bis zur Erläuterung größerer Zusammenhänge. Die visuelle Information besteht lediglich aus einem gezeichneten Supermarkt, in dem die erwachsene Frau Lebensmittel aus den Regalen nimmt und dem Kind erklärt, wofür sie bestimmt sind (*context visuals*). Visuelle Hinweise auf die im gesprochenen Text enthaltenen Themen "Armut", "Hunger" und "dritte Welt" (*content visuals*) fehlen. Der Beitrag der visuellen Informationen zum Verständnis des gesprochenen Textes ist damit eher minimal. Die Sprechweise ist standardsprachlich, die Sprechgeschwindigkeit normal bis eher langsam und pausiert – Merkmale, die die Validität der Zuordnung der Aufgabe zur GeR-Stufe B1 stützen.

Die Aufgabe umfasst zwei Items, eine Vierfachwahlaufgabe mit einem Attraktor und drei Distraktoren sowie ein Item, welches zwei Kurzantworten zu einer gemeinsamen Frage verlangt. Die richtige Antwort zum Item 1 enthält die beiden zentralen Informationen *famille pauvre* und *en France*, während die Distraktoren die nicht zielführenden Informationen *meurent de faim* und *dans les pays sous-développés* sowie *dans la rue dans les pays sous-développés* enthalten. Die distrahiende Information *meurent de faim* taucht zu Beginn der Sequenz als Frage auf, und die ebenfalls distrahiende Information *dans les pays sous-développés* wird am Ende der Sequenz sozusagen als Ausblick erwähnt. Die Aufgabe verlangt somit vom Hörenden eine einfache lexikalische Inferenz von der Aussage *trop de parents n'ont plus assez d'argent pour les nourrir correctement* zur – im Item synonym verwendeten – Wortgruppe *famille pauvre*. Es werden zwei kognitive Handlungen aktiviert, das Verstehen der Hauptaussagen und der Abgleich einer der Hauptaussagen mit einem Item, das genau nach dieser Hauptaussage fragt, also ein intensives Zuhören sowohl auf globaler als auch auf lokaler Ebene. Diese Anforderung gibt lösungsstrategisch betrachtet dem zweimaligen Hören eine besondere Bedeutung. Beim ersten Hören kann das globale Verstehen gesichert werden, das zweite Hören ermöglicht die sichere Identifikation der Hauptaussage(n) bzw. das Sortieren nach Haupt- und Nebenaussagen. Die Bilder liefern keinerlei Hinweise zur Lösung des Items, da wie bereits erwähnt weder das Thema der *familles pauvres* noch die Hungersituation in Bildern umgesetzt werden. Das Item, welches zwei Kurzantworten verlangt, fragt nach Einzelheiten des Redebeitrages der beiden zusätzlichen Sprecher (*du riz, un gâteau*). Hier wird selektives Zuhören auf lokaler Ebene abgerufen, das ebenfalls beim zweiten Hörvorgang erfolgt. Teststrategisch kann die Abfolge der Items entsprechend der Chronologie in der Hörsequenz genutzt werden. Ein Abgleich von Wortmaterial im Item mit Textmaterial ist im Gegensatz zum ersten Item nicht möglich. Da die Kinderstimmen nicht ganz leicht zu unterscheiden und für Fremdsprachlerner

generell schwerer zu verstehen sind als Stimmen Erwachsener, kann die Beantwortung des zweiten Items Schwierigkeiten hervorrufen.

Die Textvorgabe der Aufgabe *Eau de mer* ist ebenfalls linear aufgebaut. Die Sequenz passt zum Sendeformat kurzer Wissensdokumentationen, wie sie im Nachmittagsprogramm oder in Frankreich auch als Teil von Nachrichtensendungen üblich sind. Ein Sprecher erläutert die Bedeutung der Meerwasserentsalzung, spricht von der Problematik der Süßwasserknappheit auf der Erde, über die Option der Entsalzung von Meerwasser mit Hilfe entsprechender Anlagen bis zu den damit verbundenen Kosten, die für ärmere Länder problematisch sind, und der Alternative in Gestalt kleinerer, aber wirkungsvoller Anlagen. Die gesprochene Information wird durch Bildinformation – im Wesentlichen *content visuals* (vgl. Abschnitt 2.2) – gestützt, die zu jeder neuen Teilinformation einen entsprechenden Bildwechsel anbietet. So sieht man beispielsweise bei der gesprochenen Information *les trois quart de cette eau sont sous forme de glace* die Abbildung eines Gletschers. Eine deutliche stützende Funktion der visuellen Information zum Verständnis der gesprochenen bzw. gehörten Information wird deshalb vermutet. Neben der visuellen Information gibt es eine instrumentale Untermalung, die eine gewisse Dynamik und eine positive Gestimmtheit suggeriert. Auch die Vorstellung fließenden oder sprudelnden Wassers lässt sich damit verbinden. Der gesprochene Text wird von einer erwachsenen Frauenstimme, einer geübten Sprecherin ohne erkennbare akzentuale oder regionale Färbung, vorgetragen. Die Bedeutungsabschnitte sind durch Pausen abgegrenzt bzw. auch durch Stimm-Modulation variiert und markiert. Alle vier Items erfordern eine Kurzantwort.

Das erste Item ist jenes, zu dem die Schüler auch Lösungswege notieren sollten. Es enthält die Mengenangabe *les trois quarts* als expliziten Lösungshinweis. Die kognitive Tätigkeit besteht darin, die Wortoberfläche von *les trois quarts* auch hörend zu identifizieren und die Verbindung mit *sous forme de glace* herzustellen. Dabei könnte es zur Verwechslung mit der zuvor genannten Mengenangabe *moins de trois pourcent* kommen, die sich auf die Wortgruppe *propre à la consommation* bezieht. Die Schwierigkeit liegt in der anspruchsvollen Lexik, wobei sich als Lösungsstrategie einerseits der Abgleich von Text- und Itemmaterial und andererseits der Rückgriff auf Weltwissen anbietet. Diese Vermutung wäre anhand der Schülerkommentare zu verifizieren. Die Bildinformation enthält mehrere Lösungshinweise. Zunächst ist fließendes bzw. sprudelndes Trinkwasser zu erkennen, das anschließend mit verschiedenen Abbildungen von Gletschereis und anschließend mit Bildern von Meeresoberflächen kontrastiert wird. Die beiden nachfolgenden Items verlangen das Heraushören zweier spezifischer Informationen durch selektives Zuhören auf lokaler Ebene. Es sind die Namen der beiden Länder, die die größten Entsalzungsanlagen auf der Erde besitzen. Strategisch kann hier Weltwissen sowie die Konzentration auf Einzelwörter eingesetzt wer-

den. Das dritte Item fragt nach einer inferierenden Information, die im letzten Drittel der Sequenz genannt wird, auch wieder mittels selektiven Zuhörens auf lokaler Ebene. Die Antwort kann durch Rückschluss von *pays riches en pétrole* sowie *techniques moins coûteuses*, beide im Stimulus enthalten, auf *pays en développement* und *des problèmes* im Item gefunden werden, wobei strategisch wieder sehr stark das Weltwissen und der Abgleich von Stimulus- und Itemwortmaterial genutzt werden kann.

Die Aufgabe *Portables* (Langtitel: *Les ados et les portables*), die dem mittäglichen Nachrichtenprogramm des Senders TV1 entnommen wurde, ist die komplexeste der drei verwendeten Aufgaben. Es handelt sich um eine Reportage zu den mit dem exzessiven Handykonsum Jugendlicher verbundenen Kosten. Die Thematik hat seitdem allerdings auf Grund der immer günstiger werdenden Mobilfunktarife etwas an Aktualität verloren. Lediglich der mit dem Telefonieren verbundene Zeitaufwand, jugendspezifische kostenintensive Lockangebote im Internet und der Missbrauch des Internets auch per Handy, z.B. im Hinblick auf pornographische Seiten, sind unverändert aktuell. Die Sequenz wird eingeleitet durch die Ansage des Nachrichtensprechers. Die Reportage fokussiert zunächst den konkreten Fall Igors, der stundenlang SMS verschickt, um anschließend den Blick auf weitere Gefahren des Mobilfunks zu lenken, Strategien der Eindämmung seitens der Eltern Igors und abschließend das kostenlose Angebot des *contrôle parental*, mit dem bestimmte unerwünschte Angebote durch die Netzanbieter blockiert werden können.

Das Bildmaterial wechselt von der Großaufnahme der verschiedenen Sprecher zu illustrierenden Einstellungen auf Handyoberflächen, auf Igor, der sich in der elterlichen Wohnung bewegt, auf telefonierende Jugendliche und auf Zeitschriftenwerbungen für bestimmte Internetdienste, d.h. *context visuals* (vgl. Abschnitt 2.2) überwiegen. Dennoch ist eine stützende Funktion der visuellen Information zum Verständnis der gesprochenen und gehörten Information auf Grund der Gesamtinformationskomplexität zu vermuten. Die Informationsentnahme wird als schwieriger im Vergleich zu den Aufgaben *Alimentation* und *Eau de mer* angesehen, da mehrere unterschiedliche Fragen untersucht und verschiedene Perspektiven eingenommen werden. Auch die Sprecherzahl ist höher. Neben dem Nachrichtenmoderator ist ein nicht sichtbarer Sprecher zu hören, der die wichtigsten Kommentare bzw. Hintergrundinformationen zu den Bildern vermittelt, daneben ein Jugendlicher (Igor), seine Mutter sowie der Sprecher einer Verbraucherschutzorganisation. Schließlich wird Textinformation eingeblendet, die lediglich zu lesen, aber nicht zu hören ist. Die Sprecher verwenden zwar Standardsprache ohne erkennbaren Akzent, doch ist die Sprechgeschwindigkeit nicht absichtlich verringert. Die Vermutung liegt nahe, dass bereits der Stimulus höhere Verstehensanforderungen stellt als bei den beiden anderen Aufgaben. Das

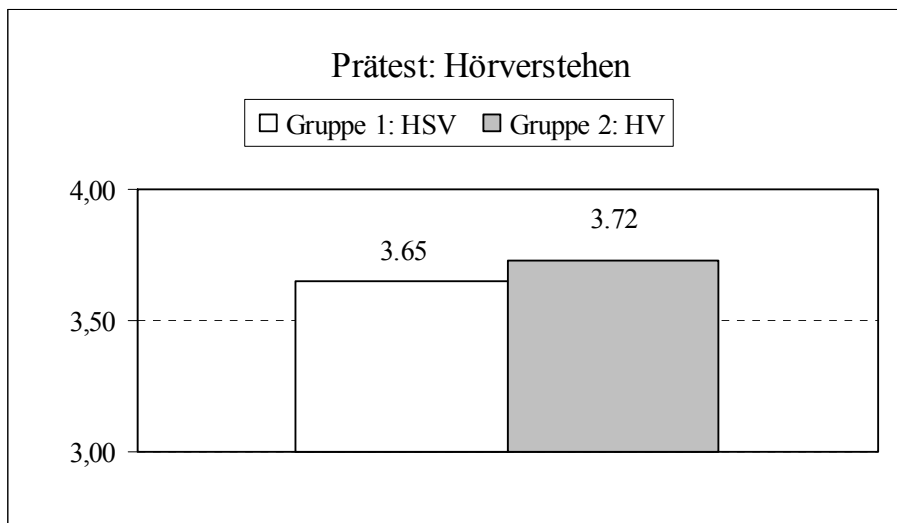
erste Item zu dem Text fragt nach einer spezifischen Information am Ende der ersten Hälfte der Gesamtsequenz (vertraglich vereinbarte Freiminuten: *quarante cinq minutes de communication par mois*). Es wird dabei selektives Zuhören auf lokaler Ebene verlangt, wobei allerdings zuvor das Globalverständnis hergestellt werden muss, was wiederum zwei Hörvorgänge rechtfertigt. Die weiteren zwei Items folgen der Chronologie der Sequenz, das vierte Item fragt nach einer Information aus dem ersten Reportage-Teil, so dass neben der hohen Itemzahl eine zusätzliche Erschwernis durch die Präsentation der Informationen und der Items besteht. Zudem ist die Sequenz *Portables* mit einer Gesamtaufnahmelänge von zwei Minuten deutlich länger, als es für die anderen beiden Videosequenzen der Fall ist. Auch das zweite Item verlangt das Erfassen einer spezifischen Information (*Pourquoi est-ce que ses parents ont changé son abonnement?*), die allerdings – im Gegensatz zum ersten Item – nicht explizit im Stimulus erwähnt, sondern stark inferierend erschlossen werden muss. Der Abgleich des Lexikmaterials in Stimulus und Item hilft in diesem Falle lösungsstrategisch nicht weiter, sondern vielmehr die Konzentration auf Einzelwörter sowie gegebenenfalls eine Hypothesenbildung auf Grund des Verstehens des Kontexts. Ähnlich wie beim zweiten Item wird auch beim dritten Item wieder das selektive Zuhören auf lokaler Ebene getestet. Es geht wieder um eine spezifische Information, die inferierend unmittelbar verstanden oder inferierend erschlossen werden kann (*En France, le contrôle parental est un service qui limite ... l'accès aux sites concernés*), wobei die Nähe von *sites pornographiques* in der Textvorlage zu *sites choquants* im Item einen deutlichen Lösungshinweis gibt, der strategisch genutzt werden kann. Der Abgleich von Text- und Itemmaterial hilft hier also weiter, so die Vermutung. Zu diesem Item beschreiben die Schüler anschließend ihre Lösungswege. Die Bildinformation gibt ebenfalls eindeutige Lösehinweise für das ausgewählte Item. Es sind Displayoberflächen mit flüchtigen, aber deutlich erkennbaren Hinweisen auf pornographische Angebote zu erkennen. Das vierte Item fragt wieder nach einer spezifischen Information (eine konkrete Zahl), die durch Konzentration auf das Einzelwort auf ganz lokaler Ebene erfasst werden kann.

3.4 Kategorisierung und Auswertung der Daten

Die Antworten der Schüler zu den halboffenen Items wurden durch eine eigens geschulte Lehramtsstudentin im Fach Französisch kodiert. Die Antworten der Schüler zu den Schwierigkeiten und Lösungsstrategien bei den drei ausgewählten Items wurden durch die Verfasser des vorliegenden Beitrags kategorisiert (induktives Vorgehen). Alle statistischen Analysen wurden mit SPSS (Version 15.0) durchgeführt.

3.5 Ergebnisse: Forschungsfrage 1

Bevor die Ergebnisse zur Forschungsfrage 1 "Zeigen Fremdsprachenlerner in einem standardisierten Leistungstest im Fach Französisch höhere Verstehensleistungen, wenn sie begleitend zum Hörtext unterstützendes Bildmaterial präsentiert bekommen?" präsentiert werden, wird die Annahme untersucht, dass die beiden Gruppen sich im Hinblick auf ihre Hörverstehensleistungen tatsächlich nicht signifikant unterscheiden bzw. dass die Gruppenzuweisung in Bezug auf das Merkmal "Hörverstehenskompetenz" tatsächlich als zufällig bezeichnet werden darf. Dazu werden die mittleren Testleistungen aus dem Prätest zum Hörverstehen verglichen. Wie Abbildung 2 zeigt, liegt die Hörverstehensleistung in der Gruppe 1 (HSV) geringfügig unter der Leistung in der Gruppe 2 (HV). Allerdings ist der Unterschied nicht signifikant, so dass bei den weiteren Analysen davon ausgegangen werden kann, dass beide Gruppen über eine gleich hohe Hörverstehenskompetenz verfügen.

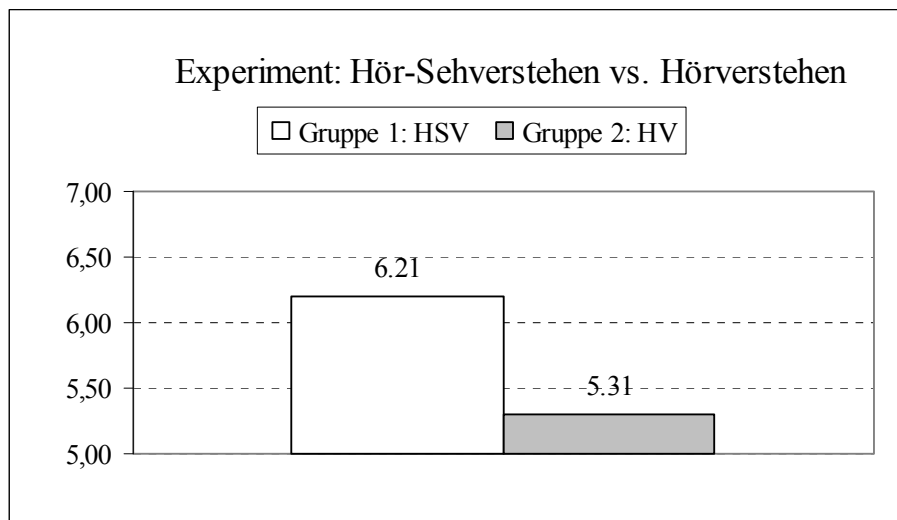


Hinweise: t -Test: $t(154) = -.325^*$, $p > .05$; *Levene-Test: Varianzgleichheit; $N_1 = 68$; $N_2 = 88$; Maximalscore = 6; Reliabilität $\alpha = .59$.

Abbildung 2: Ergebnisse (Mittelwert) aus dem Prätest nach Gruppenzugehörigkeit

Nachfolgend werden die Ergebnisse im zweiten Testteil im Vergleich betrachtet. Dazu wurden im ersten Schritt alle Items der Aufgaben zu einem Testwert zusammengefasst (Summenscore auf der Basis von 10 dichotomen Items). Aus

der Abbildung 3 wird deutlich, dass die Schüler im Mittel signifikant besser im Hör-Sehverstehenstest als im Hörverstehenstest abschneiden, obwohl alle Teilnehmer die gleichen Items zu den gleichen Hörtexten erhielten. Die Stärke des Effektes liegt bei einer Größenordnung von fast einer halben Standardabweichung ($d = .40$).



Hinweise: t -Test: $t(153.29) = 2.481^*$, $p < .05$; Levene-Test: Varianzungleichheit; $N_1 = 68$; $N_2 = 88$; Maximalscore = 10; Reliabilität Gesamttest: $\alpha_{\text{Gruppe 1}} = .57$, $\alpha_{\text{Gruppe 2}} = .70$.

Abbildung 3: Mittlere Verstehensleistungen beim Hör-Sehverstehen (HSV) und Hörverstehen (HV) nach Gruppenzugehörigkeit

Wie in Abschnitt 2.4 beschrieben, wird der Verstehensprozess der Rezipienten durch eine Interaktion von Item- und Dokumentmerkmalen bestimmt. Daher wurden die Videos unter Berücksichtigung der Anforderungen der Items dahingehend analysiert, inwieweit die Bildinformationen eine Hilfe für das Lösen der Items darstellen. Danach wurden die in Abschnitt 3.1 genannten Hypothesen formuliert, die sich auf jeweils ein Item jeder Aufgabe beziehen. Die mittlere Lösungshäufigkeit dieser Items in den beiden Gruppen wurde mittels einfaktorieller Varianzanalyse¹⁴ auf signifikante Unterschiede geprüft. Die abhängigen Variablen sind

¹⁴ Zuvor wurde der Einfluss einer Wechselwirkung zwischen den Leistungen im Prätest (HV) und der Gruppenzuweisung auf die Leistungen im Hör-Sehverstehen für die drei Items geprüft, wobei dazu die Lerner je nach Abschneiden im Prätest einer oberen oder unteren Leistungsgruppe (*Mediansplit*) zugeordnet wurden. Da keine signifikanten Effekte nachweisbar sind, werden nachfolgend lediglich die Haupteffekte für die beiden Gruppen (HSV bzw. HV) getrennt für die drei Items berichtet.

die Itemwerte, die unabhängige Variable die Gruppenzugehörigkeit. Auf eine Korrektur des Signifikanzniveaus (*Bonferoni-Korrektur*) wurde angesichts des explorativen Charakters unserer Studie verzichtet. Die Ergebnisse sind in Abbildung 4 dargestellt: Entsprechend den Hypothesen zeigt sich zwischen den Gruppen kein signifikanter Unterschied in den Leistungen für das Item 1 zu *Alimentation*. Hier liegt der Mittelwert in der Gruppe 1 (HSV) deskriptiv sogar unter dem der Gruppe 2 (HV), was auf einen möglichen störenden Effekt der zusätzlichen visuellen Verarbeitung und/oder auf eine höhere Hörverstehenskompetenz der Gruppe 2 (vgl. Abbildung 2) hindeuten könnte. Allerdings ist der beobachtete Unterschied nicht signifikant ($F[1, 156] = 3.092, p = .081$). Die Gruppe 1 schneidet jedoch erwartungsgemäß signifikant besser bei der Beantwortung der Items zu *Eau de mer* ($F[1, 156] = 8.908, p = .003$) und *Portables* ($F[1, 156] = 12.426, p = .001$) ab, wobei der Effekt in beiden Fällen hochsignifikant und mit einer Effektstärke von jeweils etwa einer halben Standardabweichung ($d = .49$ bzw. $.57$) auch deutlich ausgeprägt ist.

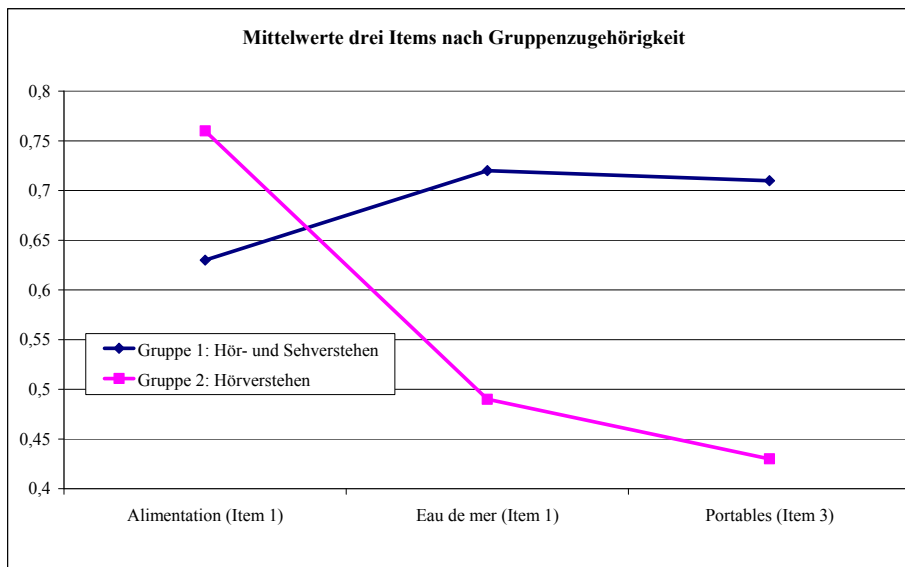


Abbildung 4: Einfaktorielle Varianzanalyse nach Gruppen

Die Daten bestätigen damit für die drei untersuchten Items die auf der Grundlage der Item- und Videoanalysen aufgestellten Hypothesen zum Grad der jeweiligen Bildunterstützung. Nachfolgend werden die weiteren eingesetzten Items hinsichtlich der Signifikanz ihrer Lösungshäufigkeit zwischen den Gruppen mittels einer einfaktoriellen Varianzanalyse geprüft. Annahme ist, dass die Bilder

keine Unterstützung darstellen, es also keinen Vorteil zugunsten der Gruppe 1 gibt.

Aufgabe	Item	Mittelwerte (Standardabweichung)		Signifikanztest
		Gruppe 1	Gruppe 2	
				F(1, 156)
Alimentation	Item 1	Vgl. Abbildung 4		
	Item 2	.74 (.44)	.67 (.47)	0.760
Eau de mer	Item 1	Vgl. Abbildung 4		
	Item 2	.75 (.44)	.53 (.50)	7.948*
	Item 3	.40 (.49)	.34 (.48)	0.517
	Item 4	.19 (.40)	.28 (.45)	1.795
Portables	Item 1	.88 (.33)	.90 (.31)	0.095
	Item 2	.40 (.49)	.32 (.47)	1.039
	Item 3	Vgl. Abbildung 4		
	Item 4	.79 (.41)	.58 (.50)	8.353*

Hinweise: $N_1 = 68$; $N_2 = 88$; * $p < .05$.

Tabelle 8: Mittelwerte, Standardabweichungen und Ergebnisse der einfaktoriellen Varianzanalyse nach Gruppen der weiteren Items

Wie der Tabelle 8 zu entnehmen ist, zeigen die Ergebnisse zu den weiteren Items, dass keine Bildunterstützung vorliegt, allerdings mit zwei Ausnahmen: Es besteht ein signifikanter Unterschied zugunsten der Gruppe 1 (HSV) für das Item 4 der Aufgabe *Portables* und für das Item 2 der Aufgabe *Eau de mer*. Der Gruppenunterschied beim Item 2 von *Eau de mer* lässt u.E. folgenden Schluss zu: Der mit anspruchsvoller Lexik (*propre à la consommation, pis encore, sous forme de glace, ...*) angereicherte Text stellt vor allem zu Beginn relativ hohe Anforderungen an das Hörverstehen, da ohne Bildunterstützung die Situierung und das notwendige Kontextwissen erst nach und nach aufgebaut bzw. abgerufen werden können. Mit Bildunterstützung sind die Betrachter dagegen sofort "im Bilde", d.h., sie können die dargebotene Information rascher situieren. Es ist zu vermuten, dass die getesteten Neuntklässler ohne Bildunterstützung den ganzen Textanfang auf Grund der hohen lexikalischen Hürden "verpassen" und erst ab Item 3 hinreichend "situieren" sind.

In Bezug auf das Item 4 von *Portables* bietet sich folgende Erklärung an: Zwar gibt es für die Beantwortung dieses Items, welches das Verstehen eines Details erfordert (Frage nach der Anzahl an maximal erlaubten Kurznachrichten, die Igor laut dem Videoausschnitt verschicken darf), keine explizite Unterstützung durch das Bild, jedoch scheint es, dass das Bild eine Strukturierungshilfe darstellt. Denn entgegen dem sonstigen Vorgehen (die Items entsprechen in der Reihenfolge der präsentierten Information), bezieht sich das Item auf eine Infor-

mation am Anfang des Videos, ist aber im Testheft als letztes Item angegeben. Da im Video der Name des Jugendlichen als Text eingeblendet wird, während dieser spricht, fällt das "Wiederfinden" der Szene den Schülern der Gruppe 1 (HSV) u.U. leichter, falls sie die Information noch nicht beim ersten Vorspielen des Videos verstanden haben.

Diese Ergebnisse zeigen einerseits, dass im Fall authentischer Videos Annahmen über Unterschiede in den Testleistungen einer gründlichen Analyse des Testmaterials bedürfen. Allerdings lassen sich nicht alle auftretenden Schwierigkeiten oder genutzte Hilfen zu komplexem Testmaterial, wie es z.B. bei Hör-Sehverstehensaufgaben der Fall ist, prognostizieren.

3.6 Ergebnisse: Forschungsfragen 2 und 3

Zur Beantwortung der Forschungsfrage 2 "Welche Schwierigkeiten haben Fremdsprachenlerner beim Lösen von Hör- und Hör-Sehverstehens(test)aufgaben und welche Strategien wenden sie an?" und der Forschungsfrage 3 "Unterscheiden sich die Untersuchungsgruppen und die einzelnen Aufgaben in Bezug auf die Zahl der genannten Schwierigkeiten und Lösungsstrategien?" wurden allen Schülern nach dem Einsammeln der Testhefte noch einmal die drei Hörtexte bzw. Videos zusammen mit den drei in Bezug auf die zu überprüfenden Hypothesen relevanten Items präsentiert und folgende schriftlich zu beantwortende Fragen vorgelegt: (a) Welche Schwierigkeiten hattest du beim Lösen? (b) Wie bist du beim Bearbeiten vorgegangen? Die Hördokumente bzw. Videos und die Items dienten dabei als Abrufreiz (*stimulated recall*; vgl. zur weiteren methodologischen Begründung z.B. Gass & Mackey 2000 sowie auch Ockey 2007: 534).

Durch Zuordnung der halboffenen Antworten bzw. der Selbsteinschätzung der Schüler konnten insgesamt zehn Kategorien von Schwierigkeitsmerkmalen der bearbeiteten Aufgaben ($N = 156$) identifiziert werden. Die Kategorien und das Kodierschema wurden dabei induktiv auf der Basis einer mehrfachen Lektüre der Schülerkommentare von den Autoren gemeinsam erarbeitet (vgl. zur Begründung z.B. Creswell 2008). Die Kategorien sind zusammen mit der Häufigkeit der entsprechenden Indikatoren in Tabelle 9 aufgeführt. Zu beachten ist, dass Schüler auch mehrere Schwierigkeiten (oder keine) genannt haben.

Bezeichnung	Typen von Aufgabenschwierigkeiten	Anzahl absolut
P1	Geschwindigkeit des Textes/der Sprecher zu hoch	50
P2	einzelnes Vokabular im gehörten Text unbekannt	61
P3	Überlastung des Arbeitsgedächtnisses	15
P4	zu viele Fragen/Items	3
P5	Bilder erschweren Verständnis	2
P6	(einzelne) Sprecher schwer verständlich	19
P7	(einzelnes) Vokabular im Item unbekannt	9
P8	Hintergrundgeräusche/-musik störend	6
P9	Sprecherwechsel (zu wenig distinkt)	3
P10	keine klare Struktur des Hörtextes erkennbar	6
P11	Anderes	15

Tabelle 9: Aufgabenschwierigkeiten laut Schüleraussagen

Häufig wurde von den Schülern vor allem die Redegeschwindigkeit als zu hoch empfunden, und einzelne Wörter im gehörten Text waren ihnen laut eigenen Aussagen unbekannt. Daneben äußerten einzelne Lerner, dass sie zu viele Informationen präsentiert bekamen und deshalb dem Hörtext nicht mehr folgen konnten. Zudem wurden einzelne Sprecher als schwer(er) verständlich bewertet. Insgesamt gesehen haben die Schüler deutlich mehr Lösungsstrategien als Schwierigkeiten angegeben, wobei einzelne Strategien auch häufig in Kombination genannt wurden (vgl. Beispiel 1 unten). Die genannten Lösungsstrategien konnten insgesamt zehn Kategorien zugeordnet werden. Diese sind zusammen mit der Häufigkeit der zugehörigen Nennungen in Tabelle 10 aufgeführt.

Bezeichnung	Typen von Lösungsstrategien	Anzahl absolut
S1	Rückgriff auf Weltwissen bzw. Fachwissen	15
S2	Nutzung Globalverständnis zur Erschließung von Einzelinformationen	27
S3	Nutzung Einzelinformationen zur Erschließung des Globalverständnisses	2
S4	Anwendung von Teststrategien	23
S5	Darbietung: Sehen/Hören für Globalverstehen, Darbietung: Beantwortung der Items oder Kontrolle der Antworten	37
S6	Raten	79
S7	Nutzung der Bildinformation	23
S8	Abgleich Gehörtes mit Antwortmaterial	94
S9	bekannte oder auffallende Einzelwörter/ Schlüsselwörter identifizieren	78
S10	häufig genannte Wörter als Indikator für richtige Lösung	2
S11	Anderes	14

Tabelle 10: Lösungsstrategien laut Schüleraussagen

Die nachfolgenden Beispiele zeigen jeweils eine Schülerantwort für die Lösungsstrategien "Raten" und "Abgleich Gehörtes mit Vokabular des Items":

Beispiel 1: Raten

Voll Feureifer stürzte ich mich wieder einmal ins Gefecht – und versagte vollkommen. Mein ursprünglicher Plan, die Antwort anzukreuzen, deren Worte ich im Text hörte, war nicht mehr in die Tat umzusetzen. Der Sprecher berichtete von allen vier Antworten. Bindeworte mussten her. Doch da war es auch zu spät. Was sollte ich jetzt tun? Unter Auferbietung all meiner geistigen Kräfte riet ich – hoffentlich richtig.

Beispiel 2: Abgleich Gehörtes mit Vokabular des Items

Ich habe mir den Text einmal angehört und Wörter in den Antworten gesucht, die auch im Text zu hören waren. (...)

Auffallend sind im Vergleich zu der Anzahl an Äußerungen möglicher Schwierigkeiten die vielen Antworten und die ausführliche Beschreibung des Vorgehens, das die Schüler nach ihrer eigenen Einschätzung angewendet haben. Die angegebenen Strategien lassen sich in insgesamt zehn Kategorien klassifizieren, wobei auch Vorgehensweisen wie (blindes) Raten auftauchen, das zwar als Lösungsstrategie bezeichnet werden kann, aber in der Regel weder als erfolgversprechend noch als kognitive Strategie bzw. als Informationsverarbeitungsstrategie betrachtet wird. Deutlich anspruchsvoller bzw. elaborierter kann dazu im Vergleich das unterschiedliche Vorgehen bei der zweimaligen Rezeption der Texte bezeichnet werden, je nachdem, ob es sich um das erste oder zweite Hören/Sehen handelt.

In Hinblick auf die genannten Schwierigkeiten und Strategien stellt sich die Frage, ob sich deren Anzahl für die Gruppen und einzelnen Aufgaben unterscheiden. Zur Beantwortung der Forschungsfrage 3 anhand der erhobenen Daten wurde auf nichtparametrische Tests zurückgegriffen (Binomialtest, Chi-Quadrat-Test, Poisson-Modell)¹⁵, die zur Anwendung kommen können, wenn die Form der Verteilung unbekannt ist und wenn kein Intervallskalenniveau, sondern lediglich ein Nominal- oder Ordinalskalenniveau der Daten vorliegt. Geprüft wird im vorliegenden Fall die Frage, ob die beobachteten Häufigkeiten einer Variablen mit den erwarteten Häufigkeiten übereinstimmen. Die Ergebnisse aus den Analysen werden wie folgt zusammengefasst:

- Die Gruppenzugehörigkeit bedingt nicht die Anzahl der insgesamt genannten Schwierigkeiten, jedoch die Gesamtzahl der genannten Strategien (jeweils unter Berücksichtigung der Gruppengröße). Die Gruppe 1 (HSV) hat deutlich mehr Strategien genannt, als auf Grund ihrer Gruppengröße zu erwarten gewe-

15 Einen Überblick über nichtparametrische Verfahren geben z.B. Bortz, Lienert & Boehnke (2008). Vgl. zum Binomialtest S. 88-92 und zum Poisson-Modell S. 20f.

sen wäre, nämlich bei einem Umfang von lediglich $N_1 = 68$ insgesamt 196 Strategien, während in der Gruppe 2 (HV; $N_2 = 88$) 198 Strategien angegeben wurden. Angesichts der Gruppengröße hätte man erwarten können, dass die Gruppe 2 eine größere Zahl von Strategien nennt als die Gruppe 1, wobei zu bedenken ist, dass es sich um Selbstaussagen der Schüler handelt. Plausibel ist, dass die Gruppe 1 mehr Strategien anwendet, weil ihr durch den visuellen Kanal mehr Möglichkeiten für das Lösen der Aufgaben zur Verfügung stehen.

- Die Gruppenzugehörigkeit bedingt auch die Anzahl der Strategienennungen bei einzelnen Aufgaben, aber nicht die Anzahl der genannten Schwierigkeiten bei den drei Aufgaben. Vor allem für die Aufgaben *Eau de mer* und *Portables* gibt es Unterschiede zwischen beobachteter und erwarteter Anzahl von Strategien. Zu *Eau de mer* wurden von der Gruppe 1 mehr Strategien angegeben als von der Gruppe 2 (46 zu 28), für die Aufgabe *Portables* ist es umgekehrt (55 zu 74).
- Die Schüler der Gruppe 2 haben deutlich häufiger als erwartet angegeben, dass sie die Lösungsstrategien "Raten" (S6) und "Abgleich Gehörtes mit Antwortmaterial" (S5) verwendet hätten.

Insgesamt zeigen die genannten Ergebnisse, dass die Schüler erwartungsgemäß bei der Rezeption von kombinierten Ton-Bild-Dokumenten mehr Strategien angewendet haben. Warum gezielte Lösungsstrategien gerade von dieser Gruppe angewendet wurden und seltener geraten wurde, ist auf Grund der auf Selbstaussagen beruhenden Daten letztendlich nicht zu beurteilen. Unsere Annahme ist, dass die Schüler, die die Stimuli lediglich gehört haben und im Vergleich geringere Testleistungen bei Vorlage der gleichen Items aufwiesen, mehr Schwierigkeiten hatten und diese entsprechend kompensiert haben. Jedoch konnten die Analysen keine Unterschiede zwischen der Anzahl der angegebenen Schwierigkeiten zwischen den beiden Gruppen aufdecken, wobei in weiteren Studien andere Instrumente wie z.B. Lautdenkprotokolle (vgl. für eine entsprechende Studie z.B. Wagner 2008) möglicherweise zu anderen Ergebnissen führen würden.

4. Diskussion und Ausblick

Mit der Einführung der Bildungsstandards für die erste Fremdsprache (KMK 2003; 2004) und des Gemeinsamen europäischen Referenzrahmens für Sprachen (Europarat 2001) wurde ein Kompetenzmodell vorgestellt, welches neben dem Sprechen, Schreiben und Leseverstehen das Hörverstehen sowie auch das Hör-Sehverstehen als zu vermittelnde fremdsprachliche Kompetenzen vorsieht. Der Fremdsprachenunterricht sollte dazu für die Schüler thematisch geeignete und möglichst authentische Texte anbieten, die auch zur Überprüfung der genannten

Kompetenzen herangezogen werden können. Im Bereich des Hörverstehens handelt es sich dabei um ausschließlich auditiv vermittelte Texte wie Mitteilungen auf Anrufbeantwortern, Bahnhofsdurchsagen oder Gesprächsausschnitte, im Bereich des Hör-Sehverstehens um audio-visuelle Texte wie Videos bzw. Film- und Fernsehaufnahmen. Letztgenannte gelten für den Fremdsprachenlerner als besonders motivierend und bieten Anlass für verschiedene Sprech- und Schreibhandlungen (vgl. Thaler 2007a; b). Außerdem kann die Authentizität der eingesetzten Materialien die Augenscheingültigkeit von Testaufgaben und damit mittelbar auch die Konstruktvalidität positiv beeinflussen (vgl. auch Feak & Salehzadeh 2001: 491; Grotjahn 2010). Allerdings stellt die Auswahl geeigneter fremdsprachlicher Videos – insbesondere wenn es sich um authentische Aufnahmen handelt – große Herausforderungen an Lehrkräfte. Es ist davon auszugehen, dass Unterschiede zwischen den Konstrukten Hörverstehen und Hör-Sehverstehen Fremdsprachenlehrern nicht hinlänglich bekannt sind und in der Ausbildung bisher zu wenig thematisiert werden, auch wenn der Stellenwert des Hörverstehens im Fremdsprachenunterricht in den letzten Jahren deutlich höher geworden ist. Auch die Bildungsstandards für die erste Fremdsprache (vgl. Abschnitt 2.6) bieten neben der Nennung unterschiedlicher Textsorten bisher kaum explizite Hinweise darauf, was die beiden Kompetenzen voneinander unterscheidet. Vor allem fehlen konkrete Beschreibungen, inwiefern kognitive Verstehensleistungen beim Hör-Sehverstehen durch den zusätzlichen visuellen Input verändert werden.

In einer ersten Pilotphase wurden am IQB Hör-Sehverstehensaufgaben zur Testung von Französischlernern in der Sekundarstufe I entwickelt, die vornehmlich authentische Sequenzen aus dem französischen Fernsehen umfassen. In der vorgestellten Studie wurden zwei Gruppen von Französischlernern der Klassenstufe 9 drei der entwickelten Testaufgaben vorgelegt, wobei die eine Gruppe die Items anhand des zugehörigen Videos und die andere Gruppe lediglich anhand der Tonspur des Videos beantworten sollte. Der zusätzliche visuelle Input hat zur Folge, dass neben Merkmalen der Texte, der Items oder der Interaktion der Texte mit den Items weitere Variablen für das Verständnis der Lerner eine Rolle spielen können, z.B. dass Rezipienten sehr unterschiedlich nonverbale Mittel verstehen und für die Beantwortung nutzen können (vgl. Wagner 2008). Trotzdem besteht kein Anlass, den visuellen Input lediglich als Quelle möglicher konstruktirrelevanter Varianz zu betrachten und die fremd-/zweitsprachliche Aufgabenentwicklung lediglich auf das reine Hörverstehen zu fokussieren. Diese Haltung wird z.B. von Buck (2001: 172) nahe gelegt, der zwar durchaus Vorteile des Einsatzes von Videos zur Messung von Hörverstehenskompetenz sieht, zugleich jedoch feststellt:

...we are usually interested in the test-takers' language ability, rather than the ability to understand subtle visual information. ... research indicates that people differ quite considerably in their ability to utilise visual information ... There are good reasons to keep the emphasis on the audio information.

Ähnlich argumentieren auch Choi, Kim & Boo (2003: 299):

The difficulty level of each test item may depend on the varying degree of relevance of visual cues to the input. Idiosyncratic differences in individual schemata may, however, lead to a misunderstanding of visual cues on the display. Furthermore, misleading pictures may serve as distractors rather than fulfilling the expected function of contextualization (Miller, 1999). A listening comprehension test should, therefore, be designed to minimize the influence of visual cues on input processing...

Eine eindeutig positive Position nimmt dagegen z.B. Wagner (2007; 2008) ein, der die Möglichkeiten betont, über die Einbeziehung der visuellen Komponente das Hörverstehenskonstrukt zu erweitern und auf diesem Wege die Validität der Schlussfolgerungen von den Testergebnissen auf die zielsprachliche Verwendungssituation (im Sinne von Bachman & Palmer 1996) zu erhöhen:

...the inclusion of the visual channel in presenting the spoken input might lead to construct relevant variance in the assessments, allowing for more valid inferences to be made from the results of those assessments (Wagner 2007: 67).

Weitere Argumente für den Einsatz von visuellem Input hatten wir bereits in Abschnitt 2 genannt. So entspricht das Verstehen von auditivem Text in Verbindung mit Bildern deutlich mehr der lebensweltlichen Realität der Lerner, und mit Hilfe von Videos können zusätzliche Lernanlässe für interkulturelles Lernen im Fremdsprachenunterricht geschaffen werden. Zudem wird auch argumentiert, dass der Einsatz von Videos den affektiven Filter herabsetzt und so zu einer höheren Verstehensleistung führt (vgl. Benson & Benson 1994). Ob dies allerdings auch in einer *high-stakes*-Testsituation gilt, ist bisher nicht geklärt. Die empirischen Ergebnisse unter Einschluss der vorliegenden Studie weisen darauf hin, dass das Hör-Sehverstehen auch empirisch vom Hörverstehen mit Hilfe von Testaufgaben in konstruktrelevanter Weise unterscheidbar ist. Insgesamt zeigen in der vorliegenden Studie die Lerner, denen audio-visuelles Material präsentiert wird, höhere Testleistungen. Zudem konnte nachgewiesen werden, dass visueller Input nur dann die Hörverstehensleistung verbessert, wenn er den auditiven Input in sinnvoller Weise ergänzt. Dieser Befund wird auch durch bereits vorliegende Forschungsergebnisse der pädagogischen Psychologie und der Fremdsprachenforschung bestätigt. Insgesamt spricht all dies für eine begründete Erweiterung des traditionellen Hörverstehenskonstrukts um visuelle Verstehenskomponenten und

zugleich auch dafür, Hör-Sehverstehen als (in Teilen) eigenständige kommunikative Kompetenz zu betrachten.¹⁶

Einschränkend muss jedoch hinzugefügt werden, dass es sich in der vorliegenden Studie lediglich um eine kleine Auswahl von Items handelt, die nicht ausschließlich für die Überprüfung der Hypothesen der Studie entwickelt worden sind. Bei den Analysen der Items und Testergebnisse wurde zudem deutlich, dass im Bereich des Hör-Sehverstehens bei einer zukünftigen Aufgabenentwicklung gezielter als bisher schwierigkeitsgenerierende Merkmale berücksichtigt werden müssen. Als weitere Einschränkung sind zudem die Testbedingungen zu nennen, die möglichst identisch mit Bedingungen in den Schulleistungsstudien vom IQB sein sollten und die insbesondere keine individuelle Zufallszuweisung zu den beiden *Treatment*-Bedingungen erlaubten. Zu den Bedingungen der Schulleistungsstudien des IQB gehört u.a. auch die zweimalige Präsentation jedes Textes, was wiederum die Anzahl der in einer bestimmten Zeitspanne testbaren Items reduziert. Wünschenswert sind daher weitere Studien zur Validitätsprüfung. Zudem handelt sich um eine *low-stakes*-Testung: Zwar können die Lehrkräfte – Kooperationspartner des IQB – als besonders motiviert gelten; in weiteren Studien sollten aber auch Merkmale wie die Motivation der Schüler oder das (spezifische) Selbstkonzept als Fremdsprachenlerner in die Auswertung einbezogen werden. Nicht berücksichtigt wurde zudem das spezifische Hintergrundwissen der Schüler zu den Themen. Schließlich bietet sich in experimentellen Versuchen die Möglichkeit an, dass eine dritte Gruppe von Schülern lediglich das Bild der Videosequenzen erhält. Im bestehenden Aufgabenpool des IQB gibt es allerdings keine Aufgaben, die lediglich das Ablesen vom Bildschirm verlangen. Das führt wiederum zur offenen Problematik einer genauen Konstruktbeschreibung des Hör-Sehverstehens. Kritisch ist in diesem Zusammenhang die Frage, was dann an spezifisch "Fremdsprachlichem" bliebe: Handelt es sich bei der Berücksichtigung von Items, die auch ohne den Fremdsprachentext lösbar sind, nicht vielmehr um die Testung einer "Film-/Medienkompetenz"? Für eine standardbasierte Testaufgabenentwicklung ist somit eine weit genauere und umfassendere Konstrukt- und Kompetenzbeschreibung des Hör-Sehverstehens notwendig, als sie bisher in den Bildungsstandards vorgelegt wurde. Fachdidaktiker und politisch Beteiligte im Bildungswesen müssen dazu u.a. die (schwierige) Entscheidung treffen, ob und inwieweit Hör-Sehverstehenskompetenz als eine Film/Medienkompetenz oder auch *Visual Literacy* (vgl. Gruba 1997; 2006) zu definieren ist und entsprechend im Fremdsprachenunterricht vermittelt und getestet werden muss. Zweifellos führt diese Zusammenführung zu einer sehr komplexen Konstruktbeschreibung,

16 Vgl. Thaler (2007: 13): "Hör-Seh-Verstehen ist keine Untervariante des Hörverstehens, sondern eine eigenständige kommunikative Kompetenz". Die Schreibweise "Hör-Seh-Verstehen" soll nach Thaler "die gleichrangige Bedeutung beider Kanäle und die Wechselwirkungen" zum Ausdruck bringen.

zu der auch das Verstehen von nonverbalen Kommunikationsmitteln zählen würde (vgl. zur Konstruktdefinition auch Ockey 2007; Wagner 2008)¹⁷.

Auch der Gemeinsame europäische Referenzrahmen für Sprachen (Europarat 2001), welcher Grundlage für die formulierten Kompetenzstufen und -beschreibungen der Bildungsstandards ist, liefert nur begrenzt eine solche Konstruktdefinition (vgl. Abschnitt 2.6). Im Bereich der audio-visuellen Rezeption wird lediglich eine Skala speziell zum Verstehen von Fernsehsendungen und Filmen angeboten, wobei zudem ausschließlich sprachliche Aspekte fokussiert werden (vgl. Europarat 2001: 77). In Bezug auf das Hör-Sehverstehen bieten daher der GeR genauso wie die aktuelle Fassung der Bildungsstandards für die erste Fremdsprache (KMK 2003; 2004) – in Anbetracht der Erkenntnisse der Medienpsychologie und empirischen Untersuchungen aus der Fremdsprachenforschung – lediglich eingeschränkt eine Orientierung für eine adäquate Konstruktdefinition. Hinzu kommt die Herausforderung, für das zu formulierende Konstrukt standardbasierte Test- und Lernaufgaben zu entwickeln und diese Kompetenzstufen zuzuordnen, so wie es bereits für weitere kommunikative Teilkompetenzen umgesetzt worden ist.

Wünschenswert ist ferner für die Entwicklung von Lern- und Testaufgaben, dass stärker medienspsychologische Aspekte und schwierigkeitsgenerierende auditive und visuelle Merkmale berücksichtigt werden. Diese Forderung betrifft nicht allein Aufgabenentwickler für den Einsatz in Schulleistungsstudien. Auch Fremdsprachenlehrern sollte in ihrer Ausbildung konsequent vermittelt werden, worauf sie bei der Auswahl und Konzeption von Aufgaben zum Hör-Sehverstehen zu achten haben. Man kann davon ausgehen, dass einige der eingangs genannten Prinzipien zur Gestaltung von Lernaufgaben bei der Nutzung von (authentischen) Videos gelegentlich zwar missachtet werden (z.B. der Einsatz von Stimuli mit "störender" Hintergrundmusik), aber trotzdem Verwendung finden können, wenn die Lerner im Hinblick auf das Medium sensibilisiert werden bzw. ihnen die Besonderheit von authentischen Videos bewusst gemacht wird. Für die Entwicklung von validen und fairen Testaufgaben sollten jedoch multimediale Gestaltungsprinzipien wie die Vermeidung von störender Redundanz besonders berücksichtigt werden. Die vorgestellte Studie kann zudem Anstoß für eine stärkere Bewusstmachung von Strategien sein, die beim Lernen und beim Testen von Hör- und Hör-Sehsequenzen eingesetzt werden und für einen Kompetenzzuwachs verstärkt im Unterricht trainiert werden sollten.

17 Wagner (2008: 239f.) argumentiert unter Bezugnahme auf die Kategorie "pragmatisches Wissen" bei Bachman & Palmer (1996) und Purpura (2004): "...the ability to utilize and comprehend the nonverbal information that a speaker conveys seems to fit well within this category of pragmatic knowledge. Being able to see the speaker allows the listener to make informed guesses about the speaker's age, status, and power and about his or her use of sarcasm, irony, and humor".

Die Ausführungen zeigen insgesamt deutlich, dass weitere empirische Studien zur medialen Nutzung im Kontext des Fremdsprachenunterrichts ein dringendes Desiderat darstellen. Zum erfolgreichen Einsatz von Strategien beim Lösen von Testaufgaben durch Fremdsprachenlerner hat die vorgestellte Studie lediglich erste Belege finden können, so dass insbesondere in diesem Bereich weitere Forschungsarbeit geleistet werden muss. Des Weiteren besteht Bedarf an einer fachdidaktischen und bildungspolitischen Diskussion des Konstrukts, das man vermitteln und messen möchte. Die empirische Arbeit weist zudem auf die besondere Herausforderung bei der Verwendung von authentischem Material im Bereich des Hör-Sehverstehens für die Forschung, die Aufgabenentwicklung aber auch den Einsatz im Unterricht hin. Nachteilig ist für den Einsatz in empirischen Erhebungen, dass nur einschränkend gezielt Itemmerkmale verändert und in ihrer Wirkung gemessen werden können (vgl. Saarbrücker TV-Projekt: Batz & Ohler 1984). Dennoch ist die Verwendung von authentischen Hör-Sehmaterialien auch für schulische Testungen äußerst wünschenswert, wenn möglichst umfassend den Prinzipien eines kommunikativen Fremdsprachenunterrichts entsprochen werden soll.

Eingang des revidierten Manuskripts 01.11.2010

Literaturverzeichnis

- Ankerstein, Hilmar S. (Hrsg.) (1972), *Das visuelle Element im Fremdsprachenunterricht*. Stuttgart: Klett.
- Bachman, Lyle F. & Palmer, Adrian S. (1996), *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.
- Baddeley, Alan D. (1992), Working memory. *Science* 255, 556-559.
- Baddeley, Alan D. (2006), Working memory: An overview. In: Pickering, Susan J. (Hrsg.) (2006), *Working Memory and Education*. New York: Academic Press, 1-31.
- Baddeley, Alan D. (2007), *Working Memory, Thought, and Action*. Oxford: Oxford University Press.
- Baltova, Iva (1999), Multisensory language teaching in a multidimensional curriculum: the use of authentic bimodal video in core French. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes* 56: 1, 31-48.
- Batz, Richard & Ohler, Peter (1984), Hörsehverstehen. Reflexionen im Umfeld eines Forschungsprojekts. In: Buße, Wolfgang; Deichsel, Ingo & Dethloff, Uwe (Hrsg.) (1984), *Fernsehen und Fremdsprachenlernen. Untersuchungen zur audio-visuellen Informationsverarbeitung. Theorien und didaktische Anwendungen*. Tübingen: Narr, 79-131.
- Bejar, Isaac; Douglas, Dan; Jamieson, Joan; Nissan, Susan & Turner, Jean (2000), *TOEFL 2000 Listening Framework: A Working Paper* (TOEFL Research Report MS-19). Prince-

- ton, NJ: Educational Testing Service. Online: www.ets.org/Media/Research/pdf/RM-00-07.pdf [28.10.2010].
- Benson, Malcolm J. & Benson, Valerie A. (1994), A new challenge: Testing the video course. *IALL Journal of Language Learning Technologies* 27: 2, 11-26.
- Betrancourt, Mireille (2005), The animation and interactivity principles in multimedia learning. In: Mayer, Richard E. (Hrsg.) (2005), 287-296.
- Bortz, Jürgen; Lienert, Gustav A. & Boehnke, Klaus (2008), *Verteilungsfreie Verfahren in der Biostatistik*. Heidelberg: Springer.
- Brett, Paul (1997), A comparative study of the effects of the use of multimedia on listening comprehension. *System* 25, 398-412.
- Brett, Paul (2001), Too many media in my multimedia? A study of the effects of combinations of media on a recall task. In: Brett, Paul (Hrsg.) (2001), *CALL in the 21st Century: Proceedings of the ESADE/IATEFL Conference, June/July 2000, Barcelona*. Whitstable, Kent: IATEFL, section 3.1, 1-20.
- Brown, Gillian (1995), Dimensions of difficulty in listening comprehension. In: Mendelsohn, David J. & Rubin, Joan (Hrsg.) (1995), *A Guide for the Teaching of Second Language Listening*. San Diego, CA: Dominic Press, 59-73.
- Brünken, Roland & Leutner, Detlef (2008), Lernen mit Medien. In: Schneider, Wolfgang & Hasselhorn, Marcus (Hrsg.) (2008), *Handbuch der Psychologie, Band Pädagogische Psychologie*. Göttingen: Hogrefe, 551-562.
- Brünken, Roland; Plass, Jan L. & Leutner, Detlef (2003), Direct measurement of cognitive load in multimedia learning. *Educational Psychologist* 38: 1, 53-61.
- Buck, Gary (2001), *Assessing Listening*. Cambridge, UK: Cambridge University Press.
- Choi, Inn-Chull; Kim, Kyoung S. & Boo, Jaeyool (2003), Comparability of a paper-based language test and a computer-based language test. *Language Testing* 20: 3, 295-320.
- Chung, Ulric K. (1994), *The Effect of Audio, a Single Picture, Multiple Pictures, or Video on Second-Language Listening Comprehension*. Ph.D. Dissertation, University of Illinois at Urbana-Champaign, USA.
- Clark, Ruth C. & Mayer, Richard E. (2002), Does practice make perfect? In: Clark, Ruth C. & Mayer, Richard E. (Hrsg.) (2002), *E-Learning & the Science of Instruction*. San Francisco: Jossey-Bass Wiley, 148-171.
- Coniam, David (2001), The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System* 29: 1, 1-14.
- Creswell, John W. (2008), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Los Angeles: Sage.
- Europarat (Hrsg.) (2001), *Gemeinsamer europäischer Referenzrahmen für Sprachen: lehren, lernen, bewerten*. München: Langenscheidt.
- Feak, Christine B. & Salehzadeh, Julia (2001), Challenges and issues in developing an EAP video listening placement assessment: A view from one program. *English for Specific Purposes* 20: Supplement 1, 477-493.
- Field, John (2004), An insight into listeners' problems: too much bottom-up or too much top-down? *System* 32: 3, 363-377.
- Field, John (2008), *Listening in the Language Classroom*. Cambridge: Cambridge University Press.
- Freedle, Roy & Kostin, Irene (1999), Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks. *Language Testing* 16: 1, 2-32.
- Gass, Susan M. & Mackey, Alison (2000), *Stimulated Recall Methodology in Second Language Research*. Mahwah, NJ: Erlbaum.

- Geranpayeh, Ardeshir & Taylor, Lynda (2008), Examining listening: Developments and issues in assessing second language listening. *Research Notes* 32, 2-5. Online: http://www.cambridgeol.org/rs_notes/rs_nts32.pdf [28.10.2010].
- Ginther, April (2001), *Exploring item characteristics that are related to the difficulty of TOEFL dialogue items* (TOEFL Research Report RR-01-16). Princeton, NJ: Educational Testing Service. Online: <http://ets.org/Media/Research/pdf/RR-01-16.pdf>. [28.10.2010].
- Ginther, April (2002), Context and content visuals and performance on listening comprehension stimuli. *Language Testing* 19: 2, 133-167.
- Goh, Christine C. M. (2000), A cognitive perspective on language learners' listening comprehension problems. *System* 28: 1, 55-75.
- Grotjahn, Rüdiger (2000), Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TESTDAF. In: Bolton, Sibylle (Hrsg.) (2000), *TESTDAF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar*. Köln: VUB Gilde, 7-55.
- Grotjahn, Rüdiger (2003), *Leistungsmessung und Leistungsbewertung*. Hagen: FernUniversität (Fernstudienbrief für den Weiterbildungs-Masterstudiengang "Deutschlandstudien. Schwerpunkt: Deutsche Sprache und ihre Vermittlung").
- Grotjahn, Rüdiger (2005), Testen und Bewerten des Hörverstehens. In: Ó Dúill, Micheál; Zahn, Rosemary & Höppner, Kristina D. (Hrsg.) (2005), *Zusammenarbeiten: Eine Festschrift für Bernd Voss*. Bochum: AKS-Verlag, 115-144.
- Grotjahn, Rüdiger (erscheint), *Testen und Evaluieren fremdsprachlicher Kompetenzen: Ein Arbeitsbuch*. Tübingen: Narr.
- Grotjahn, Rüdiger & Tesch, Bernd (2010a), Messung der Hörverstehenskompetenz im Fach Französisch. In: Porsch, Raphaella; Tesch, Bernd & Köller, Olaf (Hrsg.) (2010), 125-150.
- Grotjahn, Rüdiger & Tesch, Bernd (2010b), Messung der Leseverstehenskompetenz im Fach Französisch. In: Raphaella Porsch; Bernd Tesch & Olaf Köller (Hrsg.) (2010), 91-124.
- Gruba, Paul A. (1993), A comparison of audio and video in language testing. *JALT Journal* 15, 85-88.
- Gruba, Paul A. (1997), The role of video media in listening assessment. *System* 25: 3, 335-345.
- Gruba, Paul A. (1999), *The Role of Digital Video Media in Second Language Listening Comprehension*. Ph.D. thesis, University of Melbourne. Online: <http://eprints.unimelb.edu.au/archive/00000244/> [28.10.2010].
- Gruba, Paul A. (2006), Playing the videotext: A media literacy perspective on video mediated L2 listening. *Language Learning and Technology* 10: 2, 77-92 Online: <http://llt.msu.edu/vol10num2/pdf/gruba.pdf> [31.10.2010].
- Guichon, Nicolas & McLornan, Sinead (2008), The effects of multimodality on L2 learners: Implications for CALL resource design. *System* 36: 1, 85-93.
- Hardison, Debra M. (2010), Visual and auditory input in second-language speech processing. *Language Teaching* 43: 1, 84-95.
- Hoven, Debra (1999), A model for listening and viewing comprehension in multimedia environments. *Language Learning and Technology* 3: 1, 88-103. Online: <http://llt.msu.edu/vol3num1/hoven/index.html> [28.10.2010].
- In'nami, Yo & Koizumi, Rie (2009), A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing* 26: 2, 219-244.
- Kellerman, Susan (1992), 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied Linguistics* 13: 3, 239-258.
- Kittelberger, Rainer & Freisleben, Immo (1994), *Lernen mit Video und Film*. Weinheim/Basel: Beltz.

- Kleppin, Karin & Reich, Astrid (2010), Test-Taking Strategien. In: Berndt, Annette & Kleppin, Karin (Hrsg.) (2010), *Sprachlehrforschung: Theorie und Empirie. Festschrift für Rüdiger Grotjahn*. Frankfurt am Main: Lang, 95-112.
- KMK (2003), *Beschlüsse der Kultusministerkonferenz. Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. Beschluss vom 04.12.2003*. Online: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-erste-Fremdsprache.pdf [28.10.2010].
- KMK (2004), *Beschlüsse der Kultusministerkonferenz. Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Hauptschulabschluss. Beschluss vom 15.10.2004*. Online: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-ersteFS-Haupt.pdf [28.10.2010].
- Kostin, Irene (2004), *Exploring Item Characteristics that are Related to the Difficulty of TOEFL Dialogue Items*. TOEFL Research Report RR-04-11. Princeton, NJ: Educational Testing Service. Online: <http://ets.org/Media/Research/pdf/RR-04-11.pdf> [28.10.2010].
- Leeser, Michael J. (2007), Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning* 57: 2, 229-270.
- Mayer, Richard E. (2001), *Multimedia Learning*. Cambridge, UK: Cambridge University Press.
- Mayer, Richard E. (Hrsg.) (2005), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Mayer, Richard E. (2005a), Cognitive theory of multimedia learning. In: Mayer, Richard E. (Hrsg.) (2005), 31-48.
- Mayer, Richard E. (2005b), Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In: Richard E. Mayer (Hrsg.) (2005), 169-182.
- McGurk, Harry & MacDonald, John (1976), Hearing lips and seeing voices. *Nature* 264, 746-748.
- Meißner, Franz-Joseph (2006), Linguistische und didaktische Überlegungen zur Entwicklung von Kompetenzaufgaben im Lernbereich Mündlichkeit. *Französisch heute* 37: 3, 240-282.
- Niegemann, Helmut M.; Domagk, Steffi; Hessel, Silvia; Hein, Alexandra; Hupfer, Matthias & Zobel, Annett (2008), *Kompodium multimediales Lernen*. Berlin, Heidelberg: Springer.
- Nold, Günther & Rossa, Henning (2007a), Hörverstehen. In: Beck, Bärbel & Klieme, Eckhard (Hrsg.) (2007), *Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*. Weinheim: Beltz, 178-196.
- Nold, Günther & Rossa, Henning (2007b), Leseverstehen. In: Beck, Bärbel & Klieme, Eckhard (Hrsg.) (2007), *Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*. Weinheim: Beltz, 197-211.
- Ockey, Gary J. (2007), Construct implications of including still image or video in computer-based listening tests. *Language Testing* 24: 4, 517-537.
- Paivio, Allan (1986), *Mental Representations: A Dual Coding Approach*. Oxford: UK Oxford University Press.
- Parry, Thomas S. & Meredith, R. Alan (1984), Videotape vs. audiotape for listening comprehension tests: An experiment. *OMLTA Journal*, 47-53 [ERIC Document Reproduction Service No. ED 254107].
- Paschke, Peter (2001), Zum Problem der Authentizität in L2-Hörverstehenstests. *Fremdsprachen Lehren und Lernen* 30, 150-166.
- Pisoni, David B. & Remez, Robert E. (Hrsg.) (2005), *The Handbook of Speech Perception*. Malden, MA: Blackwell.
- Poelmans, Petra (2003), *Developing Second-Language Listening Comprehension: Effects of Training Lower-Order Skills versus Higher-Order Strategy*. Utrecht, Netherlands: LOT.

- Porsch, Raphaela; Tesch, Bernd & Köller, Olaf (Hrsg.) (2010), *Standardbasierte Testentwicklung und Leistungsmessung. Französisch in der Sekundarstufe I*. Münster: Waxmann.
- Porsch, Raphaela; Tesch, Bernd & Köller, Olaf (2010), Die Entwicklung von Kompetenzstufenmodellen zum Lese- und Hörverstehen im Fach Französisch. In: Porsch, Raphaela; Tesch, Bernd & Köller, Olaf (Hrsg.) (2010), 244-266.
- Purpura, James E. (2004), *Assessing Grammar*. Cambridge: Cambridge University Press.
- Quetz, Jürgen & Vogt, Karin (2009), Bildungsstandards für die erste Fremdsprache: Sprachenpolitik auf unsicherer Basis. *Zeitschrift für Fremdsprachenforschung* 20: 1, 63-89.
- Raabe, Horst (1997), "Das Auge hört mit". Sehstrategien im Fremdsprachenunterricht? In: Rampillon, Ute & Zimmermann, Günther (Hrsg.) (1997), *Strategien und Techniken beim Erwerb fremder Sprachen*. Ismaning: Hueber, 150-172.
- Raabe, Horst (2003), Audiovisuelle Medien. In: Bausch, Karl-Richard; Christ, Herbert & Krumm, Hans-Jürgen (Hrsg.) (2003), *Handbuch Fremdsprachenunterricht* (4., vollständig neu bearb. Aufl.). Tübingen: Francke, 423-426.
- Reid, Joy M. (Hrsg.) (1995), *Learning Styles in the ESL/EFL Classroom*. Boston: Heinle & Heinle.
- Riley, Philip (1979), Viewing comprehension: l'œil écoute. *Mélanges Pédagogiques 1979*. Centre de Recherches et d'Applications Pédagogiques en Langues, C.R.A.P.E.L, Université de Nancy II, 81-95.
- Roche, Jörg (2008), *Handbuch Mediendidaktik Fremdsprachen*. Ismaning: Hueber.
- Rost, Michael (2004), *Teaching and Researching Listening*. Harlow: Longman (Nachdruck der Auflage 2001).
- Rupp, Andre A.; Garcia, Paula & Jamieson, Joan (2001), Combining multiple regression and CART to understand difficulty in second language reading and listening comprehension test items. *International Journal of Testing* 1: 3/4, 185-216.
- Sakai, Hideki (2009), Effect of repetition of exposure and proficiency level in L2 listening tests. *TESOL Quarterly* 43: 2, 360-372.
- Schnotz, Wolfgang (2005), An integrated model of text and picture comprehension. In: Mayer, Richard E. (Hrsg.) (2005), 49-69.
- Schwerdtfeger, Inge C. (1989), *Sehen und Verstehen: Arbeit mit Filmen im Unterricht Deutsch als Fremdsprache*. Berlin: Langenscheidt.
- Sherman, Jane (2003), *Using Authentic Video in the Language Classroom*. Cambridge: Cambridge University Press.
- Shin, Dongil (1998), Using videotaped lectures for testing academic listening proficiency. *International Journal of Listening* 12, 56-79. Online: <http://www.cios.org/www/tocs/lis/012.htm> [28.10.2010].
- Skipper, Jeremy I.; van Wassenhove, Virginie; Nusbaum, Howard C. & Small, Steven L. (2007), Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex* 17: 10, 2387-2399.
- Solmecke, Gert (2000), Faktoren der Schwierigkeit von Hörtests. In: Bolton, Sibylle (Hrsg.) (2000), *TESTDAF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar*. Köln: VUB Gilde, 57-76.
- Sueyoshi, Ayano & Hardison, Debra M. (2005), The role of gestures and facial cues in second language listening comprehension. *Language Learning* 55: 4, 661-699.
- Swellers, John & Chandler, Paul (1991), Evidence for cognitive load theory. *Cognition and Instruction* 8: 4, 351-362.
- Tardieu, Hubert & Gyselinck, Valérie (2003), Working memory constraints in the integration and comprehension of information in a multimedia context. In: van Oostendorp, Herre (Hrsg.) (2003), *Cognition in a Digital World*. Mahwah, NJ: Erlbaum, 3-24.
- Thaler, Engelbert (2007a), Film-based language learning. *Praxis Fremdsprachenunterricht* 1, 9-14.

- Thaler, Engelbert (2007b), Schulung des Hör-Seh-Verstehens. *Praxis Fremdsprachenunterricht* 4, 12-17.
- Vanderplank, Robert (2010), *Déjà vu? A decade of research on language laboratories, television and video in language learning.* *Language Teaching* 43: 1, 1-37.
- Wagner, Elvis (2007), Are they watching? Test-taker viewing behavior during an L2 video listening test. *Language Learning and Technology* 11: 1, 67-86. Online: <http://llt.msu.edu/vol11num1/pdf/wagner.pdf> [31.10.2010].
- Wagner, Elvis (2008), Video listening tests: What are they measuring? *Language Assessment Quarterly* 5: 3, 218-243.
- Weidenmann, Bernd (2006), Lernen mit Medien. In: Krapp, Andreas & Weidenmann, Bernd (Hrsg.) (2006), *Pädagogische Psychologie*. Weinheim: Beltz, 423-476.
- Zydatiŕ, Wolfgang (2005), *Bildungsstandards und Kompetenzniveaus im Englischunterricht: Konzepte, Empirie, Kritik und Konsequenzen*. Frankfurt am Main: Lang.