

Eine Vorstudie zur Effektivität der Prosodischen Schrift¹ – Prosodierealisation bei japanischen Deutschanfängern –

Markus Rude
Katsufumi Narita

Zusammenfassung

Der Erwerb der Prosodie des Deutschen ist für japanische Deutschlernende nicht einfach, aber wichtig. Diesen Erwerb durch eine Visualisierung von wesentlichen prosodischen Merkmalen zu unterstützen ist das Ziel unserer Forschung. Prosodische Schrift ist eine solche Visualisierung; sie macht diese Merkmale durch gezielte Verformungen der Schrift transparent. Dieser Aufsatz behandelt eine Vorstudie, in der eine Experimentalgruppe A einen Fragekatalog in prosodischer Schrift, die Vergleichsgruppe B in Normalschrift zum Auswendiglernen erhielt; im Experiment fanden dann spontane Paargespräche statt; die Aussprache der Fragen wurde von zwei Ratern bewertet und statistisch analysiert. Das Ergebnis war zunächst negativ: Bei einigen Kriterien einer Frage war A sogar schlechter als B. Analysiert man jedoch A-I (Klasse für Ingenieurwissenschaftler) und A-G (Klasse für Geistes- und/oder Sozialwissenschaftler) getrennt ($A = A-I + A-G$), so ergibt sich bei manchen Fragen für manche Kriterien, dass A-G besser abschnitt als B; z. B. für die Intonation. Die Ergebnisse von A-I werden ebenfalls beschrieben. Z. B. hatte A-I länger gelernt und wurde dennoch meist schlechter bewertet als B. Warum die Ingenieurwissenschaftler relativ schlecht abschnitten, wird diskutiert, der Einfluss unterschiedlicher Hörerfahrung in Betracht gezogen. Trotzdem – das zeigen die A-G-Resultate – kann Prosodische Schrift einen positiven Effekt auf die Aussprache haben.

¹ Gefördert im Rahmen des JSPS-Kakenhi-Zuschusses, Projekt-Nr. 24520422.

1. Einleitung

Die Wichtigkeit adäquater Prosodie für erfolgreiche Kommunikation ist unbestritten. Manche Didaktiker halten sie seit langem für wichtiger als die adäquate Aussprache auf segmenteller Ebene, d. h. die Aussprache einzelner Laute (Hirschfeld 2001, S. 876). Für Deutschlerner in Japan gibt es nicht so viele Studien, die sich mit Intonation beschäftigen, und in den meisten Lehrbüchern werden nur die Grenztöne flüchtig erwähnt. Die Intonation vieler japanischer Deutschlerner ist entweder monoton, oder aber hat zu viel Tonhöhenbewegung in einem Satz (Narita 2007). Es gibt verschiedene Versuche, Intonation zu unterrichten, z. B. mit Visualisierungen.

Eine Visualisierung ist die der Prosodischen Schrift (Rude 2011). Konzeptionell handelt es sich dabei um eine 3D-Schrift, deren Schriftzüge sich in einem 3D-Raum winden. Auf- und Abbewegungen visualisieren die Tonhöhenbewegung (Intonation), Vor- und Zurückbewegungen visualisieren Lautstärke-Variationen (Satzakzente und Rhythmus), Dehnungen in der Horizontalen drücken zeitliche Dehnungen aus. Mit dieser Korrespondenz der drei räumlichen Dimensionen Horizontale-Vertikale-Tiefe zu den drei prosodischen Dimensionen Zeitdauer-Tonhöhe-Lautheit lassen sich die meisten prosodischen Phänomene sichtbar machen – und dadurch vielleicht leichter erlernen, wobei normale Orthographie beibehalten werden kann.

Die praktische Ausführung Prosodischer Schrift lässt noch etwas zu verfeinern übrig (siehe Abb. 1), die 3. Dimension kann zurzeit nur durch Perspektive (Größenvariation als Korrelat der Tiefenvariation) ausgedrückt werden, sie wird aber bereits recht gut von Studierenden aufgenommen. Ob sie jedoch auch wirksam ist, das ist erst ansatzweise erwiesen und sollte darum in dieser Vorstudie weiter untersucht werden.

An dieser Stelle möchten wir sehr herzlich Miriam Seel danken, die für diese Studie die SPSS-Analysen vorgenommen hat.

2. Beschreibung der Vorstudie

Um also die Wirksamkeit der Prosodischen Schrift zu prüfen, wurde diese

Vorstudie durchgeführt. Prosodie und ihre Merkmale wurden bei Deutschlernenden eingeführt, eingeübt und ihre Realisationen in einem Experiment getestet. In diesem Kapitel werden die beteiligten Universitäts-Klassen beschrieben, und wie in ihnen Prosodie eingeführt und der Versuch durchgeführt worden ist. Die detaillierte Forschungsfrage und die Hypothesen beschließen dieses Kapitel.

2.1 Klassenbeschreibung

Die Zielgruppe dieser Untersuchung sind Studierende, von denen die meisten im April 2013 an der Universität Nagoya (Japan) mit dem Studium angefangen haben und in diesem Studienjahr Deutsch als 2. Fremdsprache als Wahlpflichtfach lernen. Von den Klassen, die bei den Verfassern Deutsch lernen, wurden je zwei Klassen zur Experimental- und zur Kontrollgruppe gewählt:

Experimentalgruppe (Gruppe A)

Klasse Ingenieurwissenschaftler-D2 (unten abgekürzt zu A-I)

Klasse Geistes- und/oder Sozialwissenschaftler-D5 (A-G)

Kontrollgruppe (Gruppe B)

Klasse Naturwissenschaftler-D3 (B-N)

Klasse Geistes- und/oder Sozialwissenschaftler-D3 (B-G)

Gruppe A lernen Deutsch bei Rude mit Prosodischer Schrift, Gruppe B bei Narita ohne Prosodische Schrift. Die Unterrichtsfrequenz ist 2 Mal in der Woche. Eine Unterrichtsstunde dauert 90 Minuten. Ein Semester besteht aus 15 Wochen. Die Studierenden haben also $90 \text{ mal } 2 \text{ mal } 15 = 2700$ Minuten (45 Stunden) Deutschunterricht in einem Semester.

Die Klassen A-I und A-G bestehen je aus 32 Studierenden und haben einmal in der Woche bei Rude (native speaker) Unterricht, das andere Mal bei zwei anderen japanischen Lehrern. Die Klassen B-N und B-G bestehen ebenso aus 32 Studierenden, und Klasse B-N hat beides Mal in der Woche bei Narita (japanischer Deutschlehrer) Unterricht. Klasse B-G hat einmal in der Woche bei Narita, das andere Mal bei einem

anderen japanischen Lehrer Unterricht. Insgesamt wurden für dieses Experiment 56 Gespräche aufgezeichnet. Laut Fragebogen waren alle Studierenden japanische native speaker.

In Naritas Unterricht sollen die Studierenden zuerst ein deutsches Lesestück lesen, verstehen und dann Antworten auf Fragen zum Inhalt des Lesestücks vorbereiten. Verwendet wird das Lehrbuch: „Antworten Sie bitte auf Deutsch!“ (Narita 2005). Im Unterricht stellt Narita die Fragen und die Studierenden sagen die Antworten. Nach ein paar Fragen stellt Narita grammatisch und lexikalisch verwandte Fragen, auf die die Studierenden nicht vorbereitet sind.

Narita lässt in jeder neuen Lektion die Studierenden die Vokabeln nachsprechen, indem er gelegentlich die Wortbildung, Silbengrenzen und Laut-Buchstaben-Relationen erklärt. Danach liest er den Studierenden das Lesestück zweimal vor. Er lässt sie es weder vortragen noch nachsprechen. Prosodie wird im regulären Unterricht von Narita weder explizit erklärt noch geübt.

Bei Rude wird Prosodie von Anfang an thematisiert. A-G erhält ähnlichen Unterricht wie der von Narita, dasselbe Lehrbuch wird mit derselben Fragetechnik verwendet, darüber hinaus werden gelegentlich die Lesestücke auch durch die Studierenden laut vorgelesen. In A-I kommt ein anderes Buch zum Einsatz (Vögel & Hopf 2012), das alltägliche Sprechsituationen behandelt. Der Unterricht ist eine Mischung aus der Fragetechnik wie bei Narita (siehe oben) und Partnerarbeit, bei der die Studierenden die variablen Sprechsituationen einüben.

2.2 Prosodieeinführung

Um möglichst gleiche Ausgangsbedingungen für das Experiment zu erreichen, wurde Prosodie, ihre Hauptmerkmale und Bedeutungen (unter anderem angelehnt an Duden (2009, S. 105f.)), in allen Klassen systematisch über ein Handout eingeführt. In Gruppe A erfolgte die Einführung bereits ca. drei Monate vor dem Experiment, in Gruppe B erst ca. 3 Wochen vor dem Experiment: in B ist eine solche Einführung nämlich – wie erwähnt – kein regulärer Unterrichtsbestandteil. Das Handout für Gruppe A ist hier abgebildet (Abb. 1), das Handout für Gruppe B war ähnlich, es

enthielt aber die Beispielsätze in der Blattmitte nicht in Prosodischer Schrift, sondern in Normalschrift mit Symbolik, wie z. B. mit Pfeilen, zur Kennzeichnung der zehn aufgelisteten prosodischen Merkmale. In allen Klassen wurde das Handout besprochen und die Beispielsätze mehrere Male auszusprechen geübt, wobei auf die Korrespondenz prosodischer Merkmale und entsprechender Visualisierungen hingewiesen wurde.

Zwischen dieser Einführung und dem Experiment selbst wurde noch einmal Prosodie anhand eines weiteren Handouts behandelt. Dieses enthielt einen 12-zeiligen Dialogtext (Titel: „Wollen wir ins Kino gehen?“) in Prosodischer Schrift (für Gruppe A) bzw. in Normalschrift mit der Symbolik aus der Einführung (für Gruppe B). Der Text wurde intensiv geübt durch Nachsprechen, im Chor Sprechen, Partnerübung und zeilenweises, individuelles Vorlesen mit prosodischer Korrektur.

2.3 Versuchsdurchführung

(1) Vorbereitung

Zur Vorbereitung auf das spontane Partnergespräch erhielten die Studierenden ein Handout mit 30 Fragen, wobei die ersten fünf obligatorisch, die restlichen 25 frei in einem zweiminütigen Gespräch verwendet werden sollten, und zwar zwischen Zufallspartnern. Abb. 2 zeigt die erste Seite des Handouts, das Gruppe A erhielt. Die Akzentverschiebung in „Café“ von der 2. auf die 1. Silbe in Frage 5 erfolgte aus rhythmischen Gründen. Gruppe B bekam ein ähnliches Blatt, auf dem die Fragen statt in Prosodischer Schrift in Normalschrift vorhanden waren (hier aber ohne Symbolik).

(2) Tonaufnahme

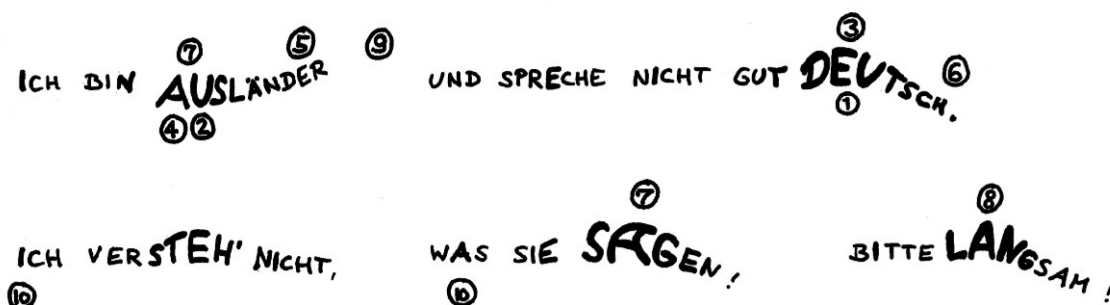
Die Aufnahme der Gespräche der Studierenden erfolgte Mitte Juli 2013 innerhalb von einer oder zwei Unterrichtsstunden im Unterrichtsraum. Zur Aufnahme verwendet wurde bei Gruppe A ein IC-Recorder Edirol R-09HR (Roland) und bei Gruppe B ein IC-Recorder ICD-UX533F (Sony) mit je dem Format MP3, 192 kbps. Die Abstände zwischen den Gesprächsführenden waren meist 3 bis 4 Sitze (ca. 2 bis 3 m), und etwa in der Mitte standen Rude bzw. Narita und nahmen die Gespräche auf.

Pronunciation: Prosodic Writing ...

... is a writing style that also encodes prosody¹ of German language.

In the style used in this class, you can see

(Prosodic feature)	(visual feature)	(possible meaning)
① main stress ²	large letter(s) ³	primary importance
② secondary stress	medium-size letter(s)	secondary importance
③ high pitch accent ⁴	slightly higher letter(s)	new information
④ low pitch accent	slightly lower letter(s)	old/inferable information
⑤ rising intonation ⁵	rising letter string	incompleteness
⑥ falling intonation	falling letter string	completeness
⑦ long stress vowel	1 stretched letter / 2 letters	distinctiveness
⑧ short stress vowel	1 standard letter	distinctiveness
⑨ pause	larger horizontal space	syntactic border; breath
⑩ rhythmic stress	bold letter(s)	establishes rhythm



Important points:

- Intonation is only shown sometimes, and not every secondary stress is visualized. Therefore: Use your ears and your feeling for speaking.
- The meaning of prosodic features is only **approximate**. There is a lot of individual variation; different stresses or intonations are possible.
- However, make sure to speak with a firm voice, emphasize the most important word, use intonation, use pauses intentionally, but not between every word.

¹ Prosody: the patterns of stress and intonation in a language.

² Stress: emphasis given to a syllable or word through a combination of relatively greater loudness, higher (or lower) pitch, and longer duration.

³ Letter(s): here: one or two vowel letters (in German language, vowels carry stress, consonants don't).

⁴ Pitch accent: a maximum or minimum (peak or valley) in the pitch contour (high or low pitch accent).

⁵ Intonation: the rise and fall of the voice in speaking; the percept of the pitch contour.

Abb. 1. Handout zur Einführung der Prosodie in Gruppe A (B erhielt ein ähnliches Blatt, aber mit Normalschrift und Symbolik anstelle der Prosodischen Schrift).

M. Rude

2-minute spontaneous conversation test

2013-07-07

When your name is called, please hold a 2-minute conversation with a random partner in German using the questions from below, by asking questions to your partner and by answering questions from your partner (please answer in whole sentences).

1. Learn the questions and replies by heart; don't use your notes or this copy.
(However, if you need, you can refer to this copy on your desk¹)
2. Use all questions from 1 to 5 (obligatory for every pair), and ask as many questions as possible from 11 to 30.
3. Be ready to reply to all questions.
(However, if you can't, you can use clarifying questions/requests²: 6 to 10).
4. Avoid long pauses.
5. Avoid unbalanced situations, in which only one student is asking questions.
(However, if your partner is not asking any question for some time, go ahead!)
6. Speak with a firm voice, use intonation and emphasize the most important word!
7. Some of the questions might be cancelled during the test, if they appear too often.
(more than about 10 times)

All other students:

- For listening: Please try to catch one keyword per interaction (the most important word) and write it down on the answer sheet.

Questions (Please repeat)

A: General Questions (1-5, obligatory)

Use each question at least once in your dialog (you or your partner)

1. WOHER KOMMEN SIE ?
2. ESSEN SIE HEUT' IN DER MENSA².
3. WAS STUDIEREN SIE DENN ?
4. UND SIE[?].
5. WOLLEN WIR INS CAFE⁻ GEHEN[?].

¹ Avoid looking at the copy in order to get a high score. If you always need to look at it, you cannot pass the test.

² Avoid these questions/requests in order to get a high score. If you always use these replies, you cannot pass the test.

Abb. 2. Handout zur Experiment-Vorbereitung für Gruppe A (B erhielt dasselbe Blatt, aber mit Normalschrift ohne Symbolik anstelle der Prosodischen Schrift).

Ab und zu wurden Geflüster und Gelächter der anderen Studierenden oder Geräusche und Widerhalle mit aufgenommen, aber nicht in dem Umfang, dass eine auditive Bewertung der Prosodie nachher unmöglich geworden wäre.

2.4 Forschungsfrage und Hypothesen

Die Forschungsfrage des Experiments lautet:

Gibt es einen Unterschied in der Performanz von Deutschlernenden beim Sprechen von auswendig gelernten Äußerungen von unterschiedlich visualisierten Materialien: Prosodische Schrift (unten abgekürzt zu PW, prosodic writing) oder Normalschrift? (Audio wurde nur im Unterricht vor dem Experiment zum Üben verwendet.)

Die Erwartung besteht darin, dass die zehn bei PW visualisierten Merkmale auch in den Realisationen der Studierenden in der Experimentalgruppe A deutlicher zum Ausdruck kommen als bei Studierenden in der Kontrollgruppe B. Allerdings sind nicht alle Merkmale in den fünf obligatorischen Fragen (hier mit Symbolik wie im Handout zur Einführung von Prosodie für Gruppe B)

1. woher ↑KOMMEN sie?
2. essen sie heute' in der ↓MENSa?
3. was stu↑DIEN sie denn?
4. und ↓SIE?
5. wollen wir ins ↑CAFÉ gehen?

vertreten, es fehlen visualisierte Nebenakzente und Pausen. Einfachheitshalber wurde auch auf die Analyse der Vokallängen verzichtet, die vier Merkmale Satzakzent, Tonhöhenakzent, finalen Grenzton und Rhythmus hielten wir für wesentlicher.

Um unsere Erwartung zu prüfen, wurden bewertbare Kriterien definiert. Drei Kriterien wurden global definiert, und zu den genannten vier Einzel-Merkmalen wurden jeweils Adäquatheit (Übereinstimmung mit dem Frage-Muster, meist kategorial, bei Rhythmus graduell) und Wahrnehmbarkeit (Deutlichkeit der Wahrnehmung, metrisch) als Kriterien festgelegt.

Die Hypothesen lauten nun, dass es statistisch nachweisbare Unterschiede bei den

11 (3 + 4 x 2) Kriterien-Bewertungen zwischen Gruppe A und B gibt, und zwar dass Gruppe A besser ist als B.

Hier sind die 11 Kriterien zusammengefasst (verkürzte Darstellung):

Kriterien zum Gesamteindruck der Aussprache im prosodischen Bereich

1. Aussprache (Nähe zum Übungsmuster), Werte: 5 – 1 (sehr gut – sehr schlecht).
2. Intonation (Sprechmelodie, d. h. Tonhöhenbewegung), Werte: 5 – 1 (dito).
3. Rhythmus (Betonungen, insbesondere der für das Deutsche typische Wechsel von betonten u. unbetonten Silben; Satzaccent), Werte: 5 – 1 (dito).

Detail-Kriterien (Werte bei Adäquatheit: wie bei 4., wenn nicht anders vermerkt):

4. Hauptakzent-Adäquatheit, Werte: 3, 2, 1 (richtig, nicht entscheidbar, falsch).
5. Hauptakzent-Wahrnehmbarkeit, Werte: 3 – 1 (deutlich – gerade noch wahrnehmbar).
6. Tonhöhenakzent-Adäquatheit (hoch für Frage 1, 3, 5 ; tief für Frage 2, 4).
7. Tonhöhenakzent-Wahrnehmbarkeit, Werte: 3 – 1.
8. Adäquatheit des finalen Grenztons (Tonbewegung am Ende der Frage. Steigend für Frage 2, 4, 5; fallend für Frage 1, 3).
9. Wahrnehmbarkeit des finalen Grenztons, Werte: 5 – 1.
10. Rhythmus-Adäquatheit (Übereinstimmung des Rhythmus mit dem Muster: 3, 3, 3, 1, 4 betonte Silben jeweils in Frage 1, 2, ... 5), Werte: 4 – 1 (hoch – keine).
11. Rhythmus-Wahrnehmbarkeit, Werte: 4 – 1 (deutlich – gerade noch wahrnehmbar).

Diese 11 Kriterien wurden für jede der 5 Fragen bei den aufgenommenen 56 Gesprächen von beiden Ratern (Verfassern) auditiv bewertet. Wahrnehmbarkeit wurde nur bewertet, wenn Adäquatheit vorlag, sonst wurde der Wert 99 vergeben (Ausschluss). Zur Analyse wurde die Software SPSS verwendet (Version 20 für Windows).

3. Ergebnisse

Zunächst kommen einige Überlegungen zur Zuverlässigkeit der Bewertungen.

Darauf folgen die Vergleiche A vs. B und A-G vs. B. Die Gruppe A-I nimmt eine Sonderrolle ein und wird in der Diskussion behandelt. Eine Zusammenfassung der Ergebnisse beschließt das Kapitel.

3.1 Korrelation der Bewertungen

Zuverlässig sind Bewertungen von Ratern dann, wenn sie ausreichend korrelieren und wenn sie nicht signifikant voneinander abweichen. Um dies zu prüfen, wurden entsprechende Kennwerte exemplarisch für Frage 2 berechnet.

Eine starke Korrelation ($\geq 0,7$) zwischen den Bewertungen von Rude und Narita lag vor für Aussprache, Tonhöhenakzent-Adäquatheit und Grenzton-Adäquatheit. Eine mittlere Korrelation (0,4 bis 0,7) ergab sich für Intonation, Rhythmus, Grenzton-Wahrnehmbarkeit und Rhythmus-Wahrnehmbarkeit. Schwach (0,2 bis 0,4) war die Korrelation bei Rhythmus-Adäquatheit.

Signifikante Abweichungen der Bewertungen von Rude und von Narita gab es 5-mal, bei der Menge $M1 = \{\text{Aussprache, Rhythmus, Grenzton-Adäquatheit, Rhythmus-Adäquatheit und Rhythmus-Wahrnehmbarkeit}\}$.

Wir nehmen an, dass sich diese Kennwerte bei den anderen Fragen ähnlich verhalten. Wir nehmen weiterhin an, dass wir uns auf die Bewertungen verlassen können, die mindestens schwach korrelieren, das ist die Menge $M2 = \{\text{Aussprache, Intonation, Rhythmus, Tonhöhenakzent-Adäquatheit, Grenzton-Adäquatheit, Grenzton-Wahrnehmbarkeit, Rhythmus-Adäquatheit, Rhythmus-Wahrnehmbarkeit}\}$, und die gleichzeitig keine signifikanten Abweichungen aufweisen.

Dann können wir uns auf die Kriterien $M2/M1$ ($M2$ ohne $M1$) verlassen, d. h. auf die Kriterien Intonation, Tonhöhenakzent-Adäquatheit und Grenzton-Wahrnehmbarkeit. Bedauerlicherweise fällt der Satzakzent durch dieses Raster, obwohl er für mündliche Kommunikation eine entscheidende Funktion besitzt.

3.2 Vergleich der Gruppen A und B

In diesem Abschnitt werden die Gruppen A und B verglichen. Außer bei Frage 2 gab es keine signifikanten Unterschiede zwischen A und B. Tabelle 1 zeigt diese

Unterschiede. Gruppe A war signifikant schlechter als B in den Kriterien Intonation und Tonhöhenakzent-Adäquatheit, zusätzlich in Aussprache und Rhythmus, wobei die letzteren beiden nicht zu den zuverlässigen Kriterien zählen.

Tabelle 1. Ergebnis des T-Tests für Mittelwerte zwischen den Gruppen A und B bei Frage 2 („Essen Sie heut' in der Mensa?“) AD steht für „Adäquatheit“, WAHR für „Wahrnehmbarkeit“. Die Mittelwerte zu den Kriterien mit * weisen signifikante Unterschiede ($p < .05$) auf. Unterstrichen sind die Kriterien, die nach Abschnitt 3.1 als zuverlässig angenommen werden.

Kriterium	Group	N	Mean	Std. Deviation
Aussprache*	A	32	3.14	.88
	B	23	3.87	.79
<u>Intonation*</u>	A	32	3.31	.98
	B	23	3.87	.91
Rhythmus*	A	32	2.86	1.03
	B	23	3.43	.79
Hauptakzent_AD	A	32	1.86	.53
	B	23	1.83	.65
Hauptakzent_WAHR	A	11	2.09	.70
	B	10	1.95	.76
<u>Tonhöhenakzent_AD</u>	A	28	2.45	.76
	B	18	2.33	.79
Tonhöhenakzent_WAHR	A	17	2.06	.70
	B	11	2.36	.64
Grenzton_AD	A	32	2.67	.71
	B	23	2.74	.54
<u>Grenzton_WAHR</u>	A	27	3.04	1.23
	B	21	3.21	1.37
Rhythmus_AD	A	32	2.86	.84
	B	23	3.13	.80
Rhythmus_WAHR*	A	29	2.31	.74
	B	23	2.80	.60

3.3 Vergleich von Klasse A-G und Gruppe B

In diesem Abschnitt werden Klasse A-G und Gruppe B (B-N + B-S) verglichen. Hier gab es signifikante Unterschiede bei Frage 1 (siehe Tabelle 2) und Frage 4 (siehe Tabelle 3). Bei beiden Fragen war Klasse A-G im Kriterium Intonation besser als die Kontrollgruppe, bei Frage 1 zusätzlich auch bei der Grenzton-Wahrnehmbarkeit.

Tabelle 2. Ergebnis des T-Tests für Mittelwerte zwischen Klasse A-G (Gruppe A, Geistes- und Sozialwissenschaftler) und Gruppe B bei Frage 1 („Woher kommen Sie?“) Die Mittelwerte zu den Kriterien mit * weisen signifikante Unterschiede ($p < .05$) auf. Unterstrichen sind die Kriterien, die als zuverlässig angenommen werden.

Kriterien	Group	N	Mean	Std. Deviation
<u>Intonation*</u>	A-G	16	4.06	.60
	B	25	3.54	.76
Hauptakzent_AD*	A-G	16	2.41	.69
	B	25	1.92	.69
<u>Tonhöhenakzent AD</u>	A-G	15	2.73	.41
	B	20	2.80	.41
Grenzton_AD*	A-G	16	2.63	.72
	B	25	1.96	.90
<u>Grenzton WAHR*</u>	A-G	14	4.04	.82
	B	16	3.16	1.36

Tabelle 3. Ergebnis des T-Tests für Mittelwerte zwischen Klasse A-G und Gruppe B bei Frage 4 („Und Sie?“) Die Mittelwerte zu den Kriterien mit * weisen signifikante Unterschiede ($p < .05$) auf. Unterstrichen sind die Kriterien, die als zuverlässig angenommen werden.

Kriterien	Group	N	Mean	Std. Deviation
<u>Intonation*</u>	A-G	24	4.17	.62
	B	40	3.66	1.13
<u>Tonhöhenakzent AD</u>	A-G	23	2.65	.46
	B	40	2.50	.53
<u>Grenzton WAHR</u>	A-G	23	3.78	1.00
	B	37	3.32	1.33

3.4 Zusammenfassung

Zunächst ergab sich durch eine Betrachtung der Rater-Korrelationen und Rater-Unterschiede, dass nur drei der elf Kriterien als zuverlässig gelten können. Da es sich hier um eine Vorstudie handelt, wurden die anderen Kriterien dennoch mitbetrachtet. Beim anschließenden Vergleich der Experimentalgruppe A mit der Kontrollgruppe B war A entgegen den Erwartungen nicht besser als B, sondern unterschied sich nur wenig. In Frage 2 war A sogar schlechter als B. Darum wurde die zu B ähnelichere Teilgruppe A-G mit B verglichen (gleiches Lehrbuch), und hier

ergaben sich bessere Werte für A-G im Vergleich zu B bei zwei Fragen: Frage 1 und Frage 4, z. B. bei der Intonation. Andere Teilgruppenvergleiche wurden auch angestellt, können aber hier aus Platzgründen nicht aufgeführt werden.

4. Diskussion

In diesem Abschnitt werden zunächst die Lernzeiten betrachtet, danach die Ergebnisse der aus dem Rahmen fallenden Klasse A-I im Vergleich zu Gruppe B berichtet und diskutiert. Im Resümee wird versucht, die Ergebnisse insgesamt zu sehen und zu interpretieren.

4.1 Lernzeiten

Im Fragebogen wurden auch die Lernzeiten eruiert: Gruppe A lernte erheblich länger als Gruppe B. Insbesondere Klasse A-I verwendete viel Zeit für das Auswendiglernen der Fragen. Warum die Lernzeiten für A länger waren als für B, kann nur vermutet werden: Es könnte daran liegen, dass das Experiment für A auch als Test zählte, für B nicht. Oder aber, A hatte eine größere Zeitspanne für die Vorbereitung zur Verfügung. Oder es hat mit der Prosodischen Schrift zu tun, eine Vermutung, die noch einmal im Resümee aufgegriffen wird.

4.2 Vergleich von Klasse A-I und Gruppe B

Das schlechte Abschneiden der Klasse A-I gegenüber der Gruppe B (B-N + B-S) in Frage 2 ist bemerkenswert. Die Unterschiede bei Aussprache, Rhythmus, Rhythmus-Adäquatheit und Rhythmus-Wahrnehmbarkeit sind hoch signifikant, die Unterschiede also größer als zwischen allen anderen Gruppen, Fragen und Merkmalen. Betrachtet man einmal statt der Unterschiede die Absolutwerte, so fallen die hohen Bewertungen für B bei Aussprache und Intonation auf (siehe Tabelle 1): A-I war vielleicht gar nicht so schlecht, sondern B war außergewöhnlich gut.

Eine mögliche Erklärung wäre eine gewisse Vertrautheit der Gruppe B mit Frage 2: B hatte eine fast identische Frage („Essen Sie heute Abend in der Mensa?“) einige Zeit

vor dem Experiment aus dem Lehrbuch (Narita 2005) zwar nicht produziert, aber einige Male gehört und beantwortet.

4.3 Resümee

Die Hypothesen konnten in dieser Vorstudie für A im Vergleich zu B nicht bestätigt werden: Bei Frage 2 war sogar die Kontrollgruppe B in einigen Merkmalen besser als die Experimentalgruppe A. Jedoch konnten signifikante Unterschiede nach der Trennung der Experimentalgruppe A in A-I (Ingenieurwissenschaftler) und A-G (Geistes- und/oder Sozialwissenschaftler) gefunden werden, aber auch dann nur für bestimmte Merkmale und bestimmte Fragen; diese Unterschiede waren für A-G im Einklang mit den Hypothesen, für A-I aber nicht.

Es ergab sich für Klasse A-G eine positive Wirkung manchmal bei Intonation, Grenztöne-Adäquatheit und -Wahrnehmbarkeit, und für Klasse A-I eine negative Wirkung manchmal bei Aussprache, Rhythmus, Rhythmus-Adäquatheit und -Wahrnehmbarkeit. PW hatte also eher eine positive Wirkung auf Intonation und auf Grenztöne, dagegen eine negative Wirkung auf Aussprache und Rhythmus. Für die Merkmale Satzakzent und Tonakzent gibt es keine solchen Tendenzen.

Klare Ergebnisse sind das allerdings nicht. Natürlich könnte PW eine ungeeignete Methode sein, um Prosodie zu erwerben: Die kognitive Belastung (cognitive load) könnte bei gleichzeitiger symbolischer und geometrischer Dekodierung zu groß sein. Es könnte aber auch sein, dass schlicht zu wenig geübt worden war: Bei der Experimentvorbereitung wurde nur 2-mal nachgesprochen, eine prosodische Korrektur fand dabei nicht statt, und Audio-Aufnahmen bekamen die Studierenden ebenfalls nicht. Es ist vielleicht kein Zufall, dass PW gerade bei einfacheren Fragen eine positive Wirkung zeigte (Frage 1: „Woher kommen Sie?“ und Frage 4: „Und Sie?“: niedrige Komplexität), eine negative Wirkung aber bei einer schwierigeren (Frage 2: „Essen Sie heut' in der Mensa?“: hohe Komplexität).

Es würde nicht wundern, wenn die kognitive Vernetzung einer prosodischen Struktur mit der entsprechenden Visualisierung bei einfachen Strukturen aus wenigen Silben schneller gelingt, bei Strukturen höherer Komplexität mindestens anfänglich

versagt. Wir möchten dies als Komplexitätshypothese bezeichnen.

Selbst die in der Analyse unauffälligen Fragen 3 und 5 (keine signifikanten Gruppen-Unterschiede) würden zu dieser Interpretation passen: Frage 3 („Was studieren Sie denn?“) besitzt mittlere Komplexität, und Frage 5 („Wollen wir ins Café gehen?“) ist zwar komplexer, wurde aber in leicht abgewandelter Form („Wollen wir ins Kino gehen?“) mit allen Klassen gleichermaßen geübt (siehe Abschnitt 2.2).

Wäre die Komplexitätshypothese korrekt, so würden längere Fragen in Prosodischer Schrift das Auswendiglernen eher erschweren. Diese Hypothese wird auch gestützt durch die Lernzeiten (A lernte länger als B), durch mindestens einen studentischen Kommentar („How do you find PW?“ „We can tell our real feelings by using PW. That is good. PW is hard to learn by heart. That is not so good.“), sowie durch die Beobachtung, dass einige Studierende die Antworten auf die Fragen in PW nicht direkt darunter schrieben (Platz war genug!), sondern Fragen und Antworten auf ein extra Blatt in Normalschrift.

Trotz der mageren statistischen Ausbeute dieser Vorstudie wird diese Forschung fortgesetzt. Dafür spricht auch die überwiegend positive Meinung der Studierenden zur Prosodischen Schrift, die aus den restlichen Kommentaren des begleitenden Fragebogens zum Ausdruck kam. 50 der 64 Studierenden der Gruppe A gaben auf obige Frage „How do you find PW?“ eine klassifizierbare qualitative Antwort (sehr schlecht, schlecht, neutral, gut, sehr gut), nämlich 6-mal sehr gut, 34-mal gut, 3-mal neutral (z. B. die obige) und 7-mal schlecht. Ästhetik spielt bei der weiteren Verbesserung der Akzeptanz auch eine Rolle.

5. Ausblick

Bei Folgestudien sollten die Ausgangsbedingungen besser kontrolliert, die Bewertungen strikt blind durchgeführt und die Rater (die Verfasser) besser trainiert werden. Auch die Testmaterialien sollten überdacht werden.

So gibt es prinzipielle Unterschiede zwischen Klassen, denn sie unterscheiden sich nicht nur nach Fachrichtungen, sondern auch nach Lehrbuch und Lehrmethode, ebenso nach der Unterrichtssprache des jeweiligen Lehrers. Weiterhin gab es zeitliche

Unterschiede in der Einführung der Prosodie, ihrer Einübung und der Zeitspanne zwischen Vorbereitung des Experiments mit Ausgabe des Handouts und dem Experiment selbst (bei A: 2 Wochen, bei B: ca. 1 Woche). Auch war die Rolle des Experiments innerhalb des Unterrichts unterschiedlich, bei A nämlich gleichzeitig Test und Experiment, bei B nur Experiment und ein Teil des Unterrichts. Diese Unterschiede sollten verkleinert werden.

Die insgesamt häufig nicht befriedigende Aussprache bei den längeren Fragen (2 und 5) könnte auf generelle Beschränkungen der prosodischen Systeme bei Anfängern hinweisen. Aus dem Rahmen fiel das sehr gute Abschneiden (Absolutwerte) der Kontrollgruppe bei der komplexen Frage 2, wobei lediglich eine ähnliche Hörerfahrung, aber keine Sprecherfahrung vorlag (Abschnitt 4.2), ein sehr interessantes Nebenergebnis, das näher untersucht werden sollte. Ungeachtet dieses Einzelphänomens sollten aber Testmaterialien bei zukünftigen Studien vielleicht vorwiegend aus Fragen niedriger oder mittlerer Komplexität bestehen.

[Literatur]

[Duden 2009]: Die Grammatik. – Duden-Verlag, 2009.

Hirschfeld, Ursula: Vermittlung der Phonetik. In: Gerhard Helbig et al. (Hrsg.): Deutsch als Fremdsprache. Ein internationales Handbuch. – Walter de Gruyter, 2001 (S. 872–879).

Narita, Katsufumi: Intonation Features in German Sentences Read by German Speakers and Japanese Speakers. – Journal of the Phonetic Society of Japan. Vol. 11 No. 2, 2007 (S. 40–54).

Narita, Katsufumi: Antworten Sie bitte auf Deutsch! – Dogakusha Verlag, 2005.

Rude, Markus: Versuch einer Bestandsaufnahme – die nukleare Kontur im Deutschen und ihre Vermittlung an japanische Deutschlernende. In: Mapping zwischen Syntax, Prosodie und Informationsstruktur. – IUDICIUM, 2011 (S. 55–69).

Vögel, Bertlinde und Anja Hopf: Und du? Sprechsituationen im Unterricht - NEU! – Osaka University Press, 2012.