

Wieviel Halbtöne braucht die Frage? Merkmale, Dimensionen, Kategorien

Anton Batliner

Angaben zur Veröffentlichung / Publication details:

Batliner, Anton. 1989. "Wieviel Halbtöne braucht die Frage? Merkmale, Dimensionen, Kategorien." In *Zur Intonation von Modus und Fokus im Deutschen*, edited by Hans Altmann, 111–62. Tübingen: Niemeyer.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



WIEVIEL HALBTÖNE BRAUCHT DIE FRAGE? MERKMALE, DIMENSIONEN, KATEGORIEN.

Anton Batliner

1. Einleitung

1.1 Überblick

In diesem Beitrag wird die Rolle unterschiedlicher Ausprägungen eines fallend-steigenden Tonverlaufs für die Satzmodusindizierung bei Verb-Erst-Fragesatz und Verb-Erst-Exklamativsatz untersucht. In Perzeptionsexperimenten mit unterschiedlichem Design werden Kategoriengrenze und Abstufungen innerhalb der Kategorien sowie der Zusammenhang zwischen akustischen und phonetischen Merkmalen, perzeptiven Dimensionen und linguistischen Kategorien genauer bestimmt.

1.2 Fragestellung

Der Tonverlauf eines Verb-Erst-Fragesatzes kann steigend oder fallend sein. Es gibt keine gesicherten Angaben darüber, wann ein Sprecher den einen und wann er den anderen Verlauf wählt. Traditionellerweise nahm man einen steigenden 'interrogativen' Verlauf als üblich an; Klein (1982) setzt dagegen den fallenden Verlauf als normal, den steigenden als mit besonderen Merkmalen behaftet an ('erstaunte Frage' o.ä.). Fest steht aber, daß der Tonverlauf (und insgesamt die Intonation) funktional unterschiedlich belastet ist: In vielen Fällen reichen andere Merkmale wie Verbsemantik und/oder Modalpartikeln aus, um den Fragemodus festzulegen, wie bei *Gehört das (etwa) Ihnen hier?* oder *Stellt Ihr euch etwa an?* In anderen Fällen (graduierbare Verben und/oder andere bzw. ambige Modalpartikeln) muß allerdings die Intonation zwischen einer Frage- und einer Nicht-Frage- (d.h. Exklamativ- oder Imperativ-) Interpretation unter-

scheiden. Beispiele sind *Stellt Ihr euch (vielleicht) an ?/!* oder *Säuft der Leo ?/!*¹

Wenn auch Verwendungsbedingungen und damit einhergehende Merkmalsausprägungen noch kontrovers sind, so dürfte wohl darüber Einigkeit bestehen, daß ein leichter initialer Abfall (FALL) der Grundfrequenz (Fo) und ein starker finaler Anstieg (RISE) der Fo auf einem Satz wie *Säuft der Leo* eindeutig eine Frage indizieren – sei das nun eine 'normale' Frage oder eine 'erstaunte' ('Versicherungs' bzw. 'Rück'-) Frage. Bei den im folgenden beschriebenen Perzeptionsexperimenten wurde daher eine entsprechende eindeutige Kontur für den einen 'Eckstimulus' der generierten Kontinua gewählt. Bei den weiteren Stimuli wurden der initiale FALL und der finale RISE systematisch variiert, wobei allerdings die Richtung der initialen und der finalen Bewegung nie gegenläufig wurde, d.h. aus dem initialen FALL wurde kein RISE und aus dem finalen RISE kein FALL. Diese Beschränkung ergab sich schlicht aus der Notwendigkeit, die Anzahl der Stimuli und damit den Umfang der Experimente im Rahmen zu halten. Es soll damit nicht impliziert werden, daß diese anderen Verläufe nicht auch mögliche oder vielleicht sogar 'bessere' Verläufe für Frage- oder Exklamativ-Sätze wären; vgl. dazu die Diskussionen der einzelnen Perzeptionsexperimente in Teil 3 sowie die Produktionsdaten in Teil 4. Ebenso wenig impliziert der Gebrauch der Begriffe FALL und RISE, daß wir der Fo-Bewegung eine größere perzeptive Relevanz beimessen als der Fo-Höhe – man könnte ja statt des finalen RISE die Höhe des Fo-Offset und statt des initialen FALL die Höhe des Fo-Onset als Parameter ansetzen. In diesem Punkt schließen wir uns Thorsen (1987:99) an:

"[...] it may be argued that our perception of pitch is continuous, i.e. we fill in – by interpolation – the empty spaces left by unvoiced sounds in the speech chain. On the other hand, it may also be argued that we anchor our perception of intonational phenomena on certain points in the time varying course of pitch and disregard what lies between such fix points. I do not know that anyone has yet settled this argument, or suggested an experiment which can resolve the issue."

1.3 Inhaltlicher Aspekt

In diesem Beitrag soll untersucht werden, wie groß der finale RISE sein muß, um eine Frage zu indizieren, und welche Veränderung des initialen FALL in wel-

¹ In manchen Fällen kann sogar Merkmalsneutralisation vorliegen, d.h. der Kontext muß über die Satzmoduszuweisung entscheiden (vgl. Oppenrieder 1988a, Batliner 1988a und Batliner et al. 1989).

chem Ausmaß Auswirkungen auf die Frageindizierung hat. Komplementär dazu ist die Frage, welche Ausprägungen der FALL-RISE-Bewegung nicht nur keine Frage indizieren, sondern einen - mehr oder weniger 'guten' - Exklamativ. Das zugrundegelegte Satzmodussystem ist in Altmann (1987) beschrieben.

1.4 Methodischer Aspekt

Wir haben in Batliner (1989) gezeigt, daß unterschiedliche Fragestellungen und damit experimentelle Designs durchaus zu prima facie widersprüchlichen Ergebnissen führen können. Es soll deshalb untersucht werden, ob unterschiedliche Testanordnungen und Instruktionen der Versuchspersonen (VPn) bei gleichem bzw. vergleichbarem Stimulusmaterial im Ergebnis konvergieren oder sich systematisch und interpretierbar unterscheiden. Ein Korrektiv zu den Perzeptionsdaten stellen Produktionsdaten dar, auf die wir in Teil 4 eingehen.

2. Testaufbau

2.1 Material

Als Ausgangsstimulus wurde der Satz *Säuft der Leo* gewählt, der auch schon in Batliner (1989) Verwendung fand. Modusindizierende 'Kontextsätze' waren für die Frage *Bist du sicher?* mit steigendem, steigend-fallendem und fallendem Fo-Verlauf sowie für den Exklamativ *Also sowas hätt' ich nicht gedacht!* Die Fo-Verläufe der Kontextsätze finden sich in Fig.9. Der Autor produzierte die Kontextsätze mit natürlicher Intonation (vgl. Batliner 1988b:245) und den Testsatz mit monotoner Intonation mehrmals im schallarmen Raum des Instituts für Phonetik in München (Aufnahmegerät: Telefunken M15, Geschwindigkeit 19 cm/sec). Die monotone Intonation sollte dabei einen möglichst neutralen Ausgangspunkt für die Manipulation darstellen. Die besten Realisierungen wurden auf einer PDP11/50 mit einer Abtastrate von 20 KHz digitalisiert und mit einem Tiefpaß von 8 KHz gefiltert. Die Kontextsätze blieben unmanipuliert, der Testsatz wurde mit einem Segmentierungsprogramm periodenweise geschnitten und diente damit als Grundlage zur pitchsynchrone Manipulation auf Halbton- (Ht-) Basis.

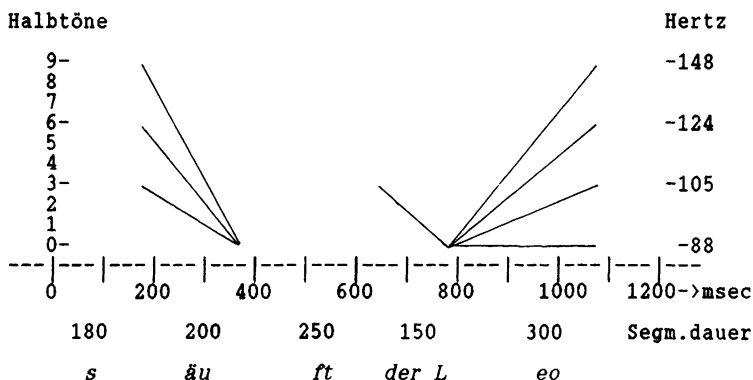
Die Stimuli in Batliner (1989) hatten eine Schrittgröße von zwei Htn. Dadurch wurden abrupte Übergänge in den Bewertungsfunktionen der Tests vermieden, bei denen die exakte Kategoriengrenze nicht genau bestimmbar ist. Da in diesem Beitrag zum einen die Kategoriengrenze nicht im Mittelpunkt des Interesses steht, zum anderen die Zahl der Paare bei dem geplanten *All-Step-Test*

(Jeder Stimulus wird mit jedem anderen gepaart) exponentiell zur Zahl der Stimuli steigt (10 Stimuli ergeben 100, 30 schon 900 Paare), legten wir die Schrittgröße auf drei Ht fest.

Neben den zwölf Stimuli mit einer FALL-RISE-Kontur (von nun an FR-Stimuli, vgl. Fig.1 und Tab.1) wurden weitere sechs mit einem Tonsprung zwischen der ersten und der zweiten Phrase (von nun an TS-Stimuli, vgl. Fig.2 und Tab.2) erzeugt, mit denen untersucht werden sollte, inwiefern ein mehr oder weniger ausgeprägter Tonsprung frage- oder exklamativtypisch ist. Die beiden Stimulusgruppen sind in Fig.1 und Tab.1 bzw. in Fig.2 und Tab.2 dargestellt. Zur Verdeutlichung sind in Fig.2 die Kurzbezeichnungen der Stimuli eingetragen.

Bei den FR-Stimuli steht vor dem senkrechten Strich der Betrag des initialen FALL, nach dem Strich der des finalen RISE. '0' steht für einen ebenen (LEVEL-) Verlauf. In der zweiten Phrase erhielt *der L* immer einen FALL von 3 Htn. Stimuli mit einem LEVEL-Verlauf auf *säuft* oder *der L* wurden nicht aufgenommen, da sie im Verhältnis zu den anderen Stimuli sehr unnatürlich klangen.

Fig.1: FALL-RISE-Stimuli, schematisch dargestellt

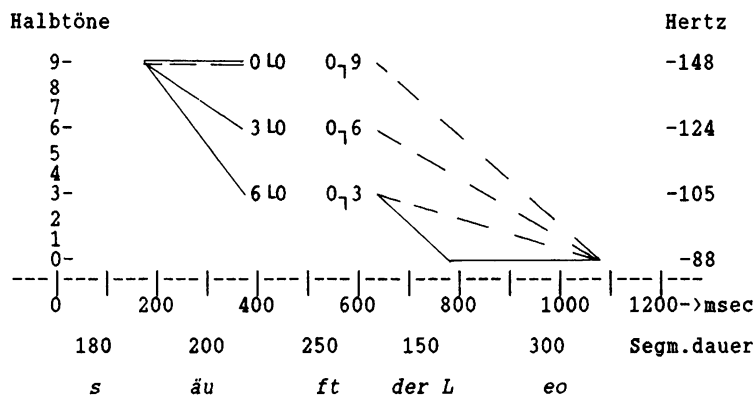


Tab.1: FALL-RISE-Stimulus-Matrix

		initialer FALL in Halbtönen		
		9	6	0
finaler RISE in Halb- tönen	0	9 0	6 0	3 0
	3	9 3	6 3	3 3
	6	9 6	6 6	3 6
	9	9 9	6 9	3 9

Bei den TS-Stimuli steht am Anfang der Betrag der Bewegung auf *säuft*, am Ende der Betrag der Bewegung auf *Leo* bzw. *der Leo*. 'L' steht für finalen LEVEL-Verlauf auf *Leo*, und '┘' für initialen LEVEL-Verlauf auf *säuft*. Bei den Stimuli in der ersten Spalte von Tab.2 wurde der Fo-Verlauf auf *der Leo* konstant gehalten (FALL von 3 Htn auf *der* und LEVEL auf *Leo*); der FALL auf *säuft* variiert derart, daß sich Tonsprünge von 0, 3 und 6 Htn ergeben. Bei den Stimuli in der zweiten Spalte von Tab.2 wurde der LEVEL-Verlauf auf *säuft* konstant gehalten; Anfang und Ausmaß des FALL auf *der Leo* variieren derart, daß sich ebenfalls Tonsprünge von 0, 3 und 6 Htn ergeben.

Fig.2: Tonsprung-Stimuli



Tab.2: Tonsprung-Stimulus-Matrix (Verlauf und Ausmaß des FALL in Halbtönen)

		LEVEL: FINAL - INITIAL	
TONSPRUNG	0	6 L0	0┘9
	3	3 L0	0┘6
	6	0 L0	0┘3

2.2 Hypothesen und statistische Analyse

Für die FR-Stimuli nehmen wir die folgenden Hypothesen an:

- Je höher der finale RISE, desto wahrscheinlicher ist eine Frage- und desto schlechter eine Exklamativinterpretation.
- Je höher der initiale FALL, desto wahrscheinlicher ist eine Exklamativinterpretation.

In all diesen Fällen können wir bei der Prüfstatistik einseitig testen. Für die Fragen läßt sich beim initialen FALL keine Vorhersage machen, da zum einen eine bessere Exklamativinterpretation eine schlechtere Frageinterpretation bedingen könnte; zum anderen haben wir aber in Batliner (1989) gesehen, daß beim gleichen Testsatz ein höherer initialer FALL auch eine bessere Frageinterpretation bedingen kann. In diesen Fällen müssen wir also zweiseitig testen.

Die FR-Stimuli könnten im Rahmen eines Varianzanalytischen Designs untersucht werden: Es gibt zwei Faktoren, den Faktor FALL mit drei Klassen (9, 6, 3 Ht) und den Faktor RISE mit vier Klassen (0, 3, 6 und 9 Ht). Da die dafür benötigten Voraussetzungen der Varianzhomogenität und der Normalverteilung aber bei unseren experimentellen Daten mehr schlecht als recht erfüllt sind, nehmen wir grundsätzlich einen 'konservativen' Standpunkt ein und testen mit nicht-parametrischen Verfahren. Dabei werden die Klassen eines Faktors gemittelt und jeweils benachbarte Klassen des anderen Faktors mit einem Wilcoxon-Test für abhängige Stichproben auf signifikante Unterschiede geprüft. Da dabei über die gleiche Stichprobe mehrfach getestet wird, muß das gewählte Signifikanzniveau α von 0.05 nach der Formel $1-(1-\alpha)^{1/c}$ abgesenkt werden, wobei c für die Zahl der durchgeführten Tests steht. Bei den einzelnen Experimenten sind die durchgeführten Vergleiche und die dabei erzielten p-Werte sowie das abgesenkte Signifikanzniveau α angegeben. Signifikante Werte sind mit einem Stern gekennzeichnet; wenn einseitig getestet wurde, sind die Werte unterstrichen.

Bei den TS-Stimuli können keine Hypothesen aufgestellt werden, da ein größerer Tonsprung gleichermaßen eine bessere Frage- oder Exklamativinterpretation bewirken könnte. Zusätzliche (zweiseitige) Tests für diese Stimuli hätten darüber hinaus den Nachteil, daß das Signifikanzniveau weiter abgesenkt werden müßte. Wir begnügen uns deshalb bei diesen Stimuli mit einer explorativen, deskriptiven Statistik.

2.3 Datenerhebung

Die Daten wurden bei allen Experimenten auf die gleiche Weise erhoben: VPn waren StudentInnen der Germanistischen Linguistik oder der Phonetik, die für ihre Teilnahme bezahlt wurden. Sie saßen im Sprachlabor des Instituts für Phonetik in München vor einem Abstimmkästchen und hörten die Stimuli über Kopfhörer. Genau nach jedem Stimulus(-paar) wurde die Abstimmung freigegeben; dies wurde durch eine Lampe an jedem Kästchen angezeigt. Die VPn hatten dann vier Sekunden Zeit, ihre Antwort zu überlegen und die entsprechende Taste zu drücken. Die Antworten wurden auf einer PDP11/03 gesammelt und zur weiteren Verarbeitung aufbereitet. Die statistische Auswertung wurde mit dem Programmpaket SPSS-PC+ vorgenommen.

3. Perzeptionsexperimente

Es wurden die folgenden Perzeptionsexperimente durchgeführt; in Klammern stehen die im weiteren gebrauchten Abkürzungen sowie ein Hinweis auf das jeweilige Unterkapitel:

Identifikationstest	(IT - 3.1)
Antwortobligationstest	(AT - 3.2)
Kontexttests	(KT1 - 3.3.1; KT2 - 3.3.2; KT3 - 3.3.3)
Diskriminationstest	(DT - 3.4)
Ähnlichkeits- (<i>All-Step</i> -) Test	(ÄT - 3.5)

3.1 Identifikationstest (IT)

3.1.1 Design

Beim IT handelt es sich um einen sog. 'forced-choice'-Test: Die VPn erhalten zwei (manchmal auch mehrere) Möglichkeiten vorgegeben, zwischen denen sie sich entscheiden **müssen**. ITs werden üblicherweise zusammen mit DTs bei Untersuchungen zur kategorialen Wahrnehmung verwendet (vgl. Repp 1984; zur Kategorialität von Intonationsphänomenen vgl. Lindsay/Ainsworth 1985, Schiefer/Batliner 1988 und Batliner 1989).

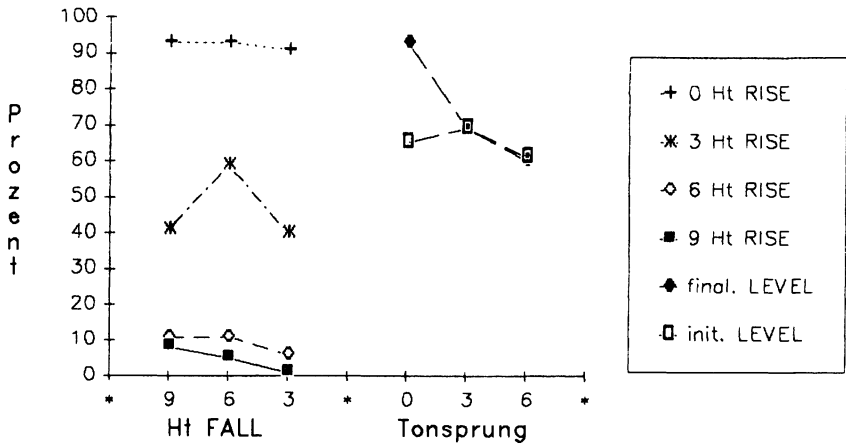
Die 18 Stimuli wurden 17 VPn randomisiert je 5mal dargeboten. Die Instruktion lautete: "Bitte entscheiden Sie, ob es sich bei den Stimuli um eine Frage oder um keine Frage handelt, und drücken Sie die linke Taste für 'Frage' oder die rechte für 'Nicht-Frage'".

3.1.2 Ergebnis

Fig.3 zeigt das Ergebnis; auf der y-Achse sind in Prozent die Nicht-Frage-Klassifizierungen aufgetragen. Hier wie in den folgenden Abbildungen sind auf der x-Achse für die FR-Stimuli die drei Klassen des initialen FALL und für die TS-Stimuli die Tonsprünge aufgetragen. Die Legende zeigt die Symbole der Klassen des finalen RISE für die FR-Stimuli bzw. der TS-Stimuli mit finalem vs. initialem LEVEL-Verlauf. Die FR-Stimuli mit LEVEL-Verlauf (0 Ht RISE) werden eindeutig als Nicht-Fragen klassifiziert, die mit 6 und 9 Ht RISE eindeutig als Fragen; die mit 3 Ht RISE liegen im Zufallsbereich, sind also ambig. Signifikante Unterschiede zeigen sich zwischen 0 und 3 Ht RISE und zwischen 3 und 6 Ht RISE. Das Ergebnis stimmt mit Batliner (1989) überein, wo sich beim gleichen Testsatz, aber einer kleineren Schrittgröße von 2 Htn eine Kategoriengrenze zwischen 2 und 4 Ht RISE ergab.

Fünf der sechs TS-Stimuli liegen leicht über dem Zufallsbereich, der sechste, Stimulus 6 L0, wird in über 90% als Nicht-Frage klassifiziert. Das läßt den Schluß zu, daß die TS-Stimuli grundsätzlich schlechte Fragekandidaten sind.

Fig.3: Identifikationstest



Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0102$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
<u>0.0001*</u>	<u>0.0002*</u>	<u>0.0807</u>	0.0465	0.0277

3.2 Antwortobligationstest (AT)

3.2.1 Design

Mit diesem Experiment soll untersucht werden, ob sich bei einer abgestuften Antwortvorgabe andere Bewertungsfunktionen ergeben als bei der binären des ITs. Implizit wird mit diesem Design natürlich ein Konzept der Kategorialität, das sich mit dem Ermitteln von Kategoriengrenzen begnügt, infrage gestellt. Der AT leitet über zu den Kontexttests (KTs), da er mit dem IT die Explizitheit der Bewertung, mit den KTs aber die Abstufungen der Bewertung gemeinsam hat.

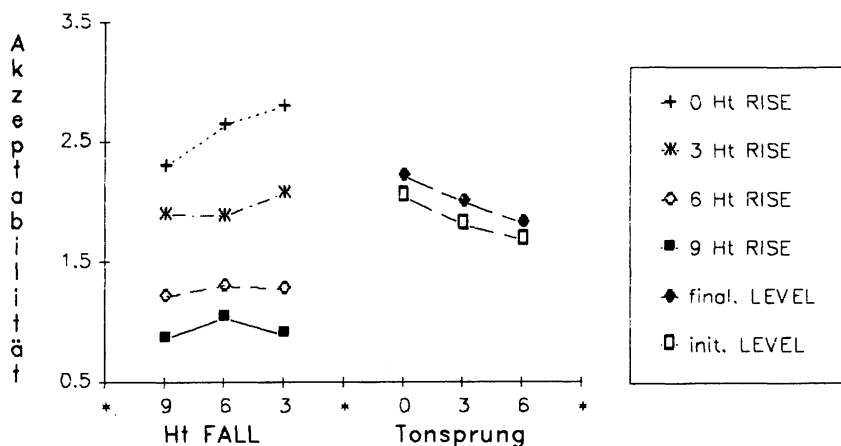
18 VPn wurden die 18 Stimuli randomisiert je 5mal dargeboten. Die Instruktion lautete: "Bitte geben Sie an, wie stark Sie sich durch die Äußerungen zu einer Antwort gedrängt fühlen und drücken Sie dann die entsprechende Taste der Tasten 1 bis 5." Die Skala lautete: 1 'sehr stark', 2 'stark', 3 'mittel', 4 'wenig', 5 'sehr wenig'. Diese Skala wurde für die Auswertung und für die Darstellung in eine Skala von 0-4 umgerechnet. Als Maßzahl pro Item und VP erhält man also einen Mittelwert aus den fünf Einzelbeurteilungen.

3.2.2 Ergebnis

Fig.4 zeigt das Ergebnis. Es war zu erwarten, daß im Gegensatz zum IT die Extrembereiche bei einer solchen Instruktion nicht besetzt sind, da manche VPn die extremen Bewertungen 1 oder 5 selten bis nie vergeben. Interessant ist aber, daß die Abstufung zwischen den Klassen von RISE kontinuierlicher ist und sich alle vier Klassen vom jeweiligen Nachbarn signifikant unterscheiden, also auch 6 Ht RISE von 9 Ht RISE. Lenkt man demnach die Aufmerksamkeit der VPn auf eine graduelle Abstufung, so zeigt sich auch innerhalb des Fragemodus - in unserem Fall bei den Stimuli mit 6 oder 9 Ht RISE - eine Differenzierung. Eine stärkere Antwortobligation beim Hörer hat ihre Ursache in einer stärkeren '**Frag Gehaltigkeit**' beim Stimulus, und diese wiederum ist offensichtlich bedingt durch eine stärkere Ausprägung des frageindizierenden finalen RISE. (Der Begriff '**Frag Gehaltigkeit**' sei hier informell eingeführt, da er für sich sprechen kann; er wird unten in Teil 5 genauer erläutert und abgeleitet.)

Die TS-Stimuli liegen alle, einschließlich des Stimulus 6 L0, im mittleren, unentschiedenen Bereich. In der leicht fallenden Tendenz beim Tonsprung von 0 über 3 nach 6 Htn deutet sich an, daß ein stärkerer Tonsprung - wenn auch in geringem Ausmaß - die Frageindizierung verbessern kann.

Fig.4: Antwortobligationstest



Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0102$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
<u>0.0025*</u>	<u>0.0025*</u>	<u>0.0054*</u>	0.0979	0.8970

3.3 Kontexttests (KTs)

Die Kombination der Teststimuli mit Kontexten, die die Kategorie (in unserem Fall also den Satzmodus) indizieren, bietet die Möglichkeit einer indirekten und auch 'natürlicheren' Ermittlung der Bewertungsfunktionen. Man kann durch eine gezielte Veränderung des Kontexts unterschiedliche Einflüsse untersuchen. So variieren wir zum einen den indizierten Modus durch zwei unterschiedliche Vorgängeräußerungen, die entweder einen Fragesatz oder einen Exklamativsatz indizieren. Zum anderen variieren wir den Intonationsverlauf der fragesatzindizierenden Vorgängeräußerung, um damit zu untersuchen, inwiefern eine intonatorische Kongruenz von benachbarten (adjazenten) Teiläußerungen gefordert ist.

Als fokusneutralen Kontextsatz, der einen Fragesatz indizieren soll, wurde der Satz *Bist du sicher?*, als exklamativindizierender Kontextsatz der Satz *Also sowas hätt' ich nicht gedacht!* verwendet. Der Satz *Bist du sicher?* erhielt drei unterschiedliche Konturverläufe, einen steigenden (RISE), einen steigend-fallenden (RISE-FALL) und einen fallenden (FALL). In KT1 (Teil 3.3.1.) hatte *Bist du sicher?* den RISE-, in KT2 (Teil 3.3.2.) den FALL-Verlauf; in KT1 und KT2 wurde der gleiche exklamativindizierende Kontextsatz *Also sowas hätt' ich nicht ge-*

dacht! verwendet, Testsätze waren die FR- und die TS-Stimuli; die Verläufe der Kontextsätze finden sich unten in Fig.9. KT3 unterscheidet sich von KT1 und KT2 darin, daß kein manipuliertes, sondern nur natürlichsprachliches Material verwendet wurde; im einzelnen s. Teil 3.3.3.

3.3.1 Kontexttest 1 (KT1)

3.3.1.1 Design

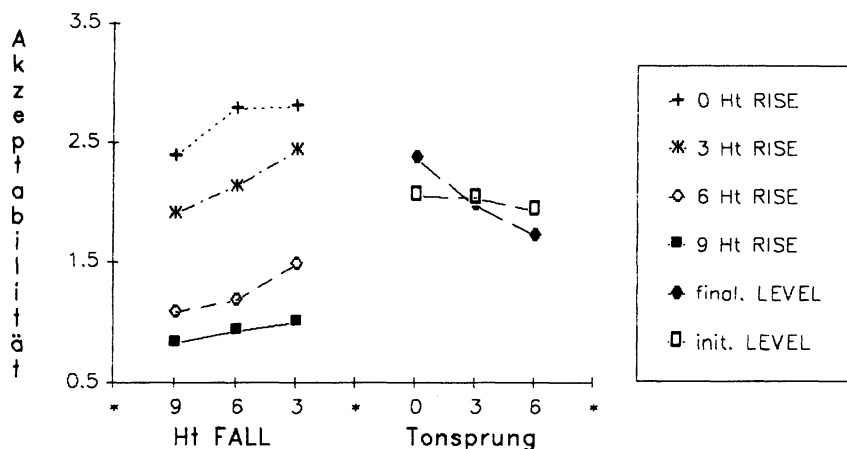
Kontextsätze waren *Also sowas hätt' ich nicht gedacht!* und *Bist du sicher?* mit RISE-Verlauf. Die 36 Paare aus Kontextsatz und Testsatz wurden 15 VPn je viermal, randomisiert und aufgeteilt auf zwei Testsitzungen, dargeboten. Das Intervall innerhalb der Stimuli betrug 0.5 sec, das zwischen den Paaren 4 sec. Die VPn gaben ihre Antwort auf einer fünfstufigen Ratingskala ab. Sie entschieden damit, ob die beiden Sätze als Äußerung ein und desselben Sprechers 'sehr gut' = 1, 'gut' = 2, 'noch gut' = 3, 'schlecht' = 4 oder 'sehr schlecht' = 5 zusammenpaßten. Diese Skala wurde für die Auswertung und für die Darstellung wieder in eine Skala von 0-4 umgerechnet. Als Maßzahl pro Item und VP erhält man also einen Mittelwert aus den vier Einzelbeurteilungen.

3.3.1.2 Ergebnis

Fig.5 zeigt die Ergebnisse und die Signifikanz der erhaltenen p-Werte beim Vergleich der Klassen für den Fragekontext. Die Verläufe in Fig.5 sind ähnlich denen beim AT in Fig.4. Wieder unterscheiden sich alle vier RISE-Klassen signifikant vom jeweiligen Nachbarn. Obwohl sich die Klassen des Faktor FALL (bedingt durch die nötige Absenkung des Signifikanzniveaus) nicht signifikant voneinander unterscheiden, kann man doch die Tendenz feststellen, daß ein ausgeprägter initialer FALL eine Frageinterpretation begünstigt. Setzt man den finalen RISE als Hauptmerkmal, den initialen FALL als Nebenmerkmal bei der Fragemarkierung an, so läßt dieses Ergebnis den Schluß zu, daß Nebenmerkmale dann relevant werden, wenn die Aufgabenstellung 'natürlicher' ist.

Die TS-Stimuli liegen wieder im mittleren, ambigen Bereich.

Fig.5: Kontexttest 1, Fragekontext



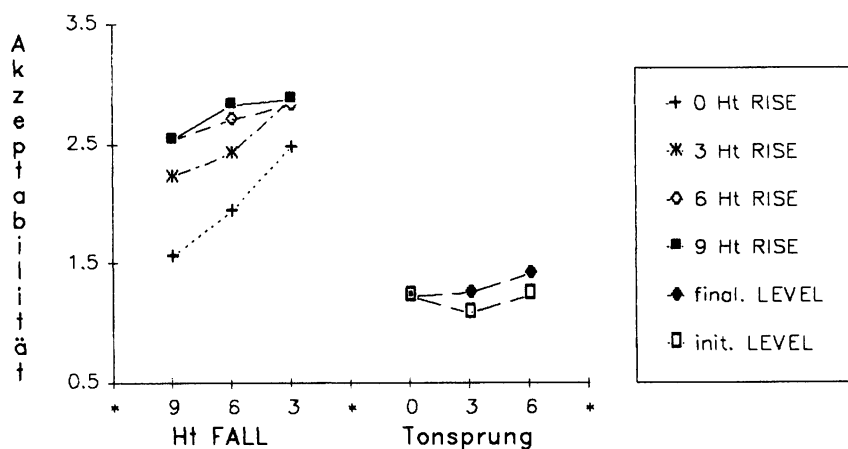
Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0051$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
<u>0.0007*</u>	<u>0.0007*</u>	<u>0.0010*</u>	0.0995	0.0144

Fig.6 zeigt die Ergebnisse für den Exklamativkontext. 0 und 3 Ht RISE unterscheiden sich signifikant, sowie 9 und 6 bzw. 6 und 3 Ht FALL. Damit bestätigt sich die Annahme, daß unter unseren Stimuli ein hoher initialer FALL und ein finaler LEVEL-Verlauf am exklamativtypischsten ist (Stimulus 9|0) und daß sowohl eine Absenkung des FALL als auch eine Anhebung des RISE (von 0 auf 3 Ht RISE) eine Verschlechterung mit sich bringen.

Beim Vergleich von Fig.5 und Fig.6 zeigt sich, daß der Stimulus 9|0 in Fig.6 schlechter bewertet ist als die fragetypischen mit 9 oder 6 Ht RISE in Fig.5. Die schematische Fo-Manipulation allein kann nicht dafür verantwortlich gemacht werden, da dann ja auch der gleiche Effekt einen Einfluß auf die fragetypischen Stimuli haben müßte. Man kann also vermuten, daß der bei unseren Teststimuli 'exklamativtypischste' Stimulus 9|0 an sich eine weniger exklamativtypische Fo-Kontur aufweist; wir werden bei der Behandlung der Produktionsdaten in Teil 4 auf diesen Punkt zurückkommen.

Fig.6: Kontexttest 1, Exklamativkontext



Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0051$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
0.0037*	0.5098	0.2585	<u>0.0038*</u>	<u>0.0038*</u>

3.3.2 Kontexttest 2 (KT2)

3.3.2.1 Design

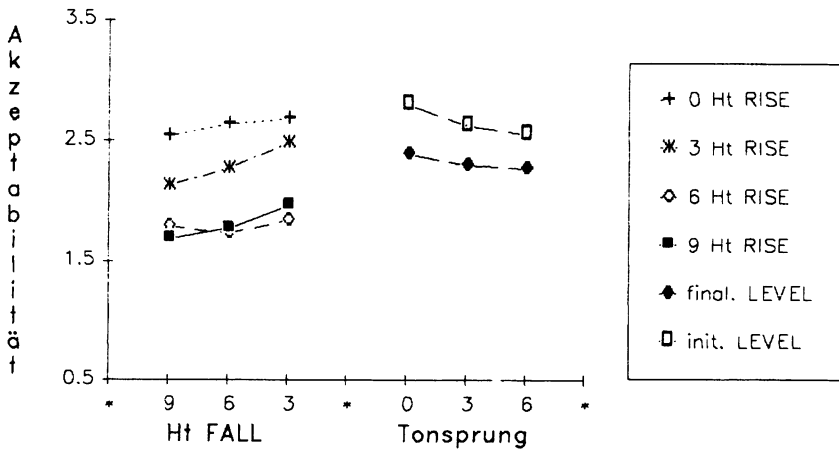
Design und Ablauf des Experiments waren identisch mit dem in 3.3.1. geschilderten, mit einer Ausnahme: Der frageindizierende Kontextsatz wies einen FALL auf (cf. Fig.9). Es nahmen 14 VPn teil.

3.3.2.2 Ergebnis

Fig.7 zeigt das Ergebnis und die Signifikanzen der erhaltenen p-Werte für den Fragekontext: Im Gegensatz zum eben beschriebenen Experiment unterscheiden sich 9 und 6 sowie 3 und 0 Ht RISE nicht signifikant voneinander; einzig zwischen 6 und 3 Ht RISE, also im Bereich der Kategoriengrenze, zeigt sich ein signifikanter Unterschied.

Vergleicht man Fig.5 mit Fig.7, so sieht man, daß in Fig.7 9 und 6 Ht RISE deutlich schlechter bewertet werden als in Fig.5; dieser Unterschied muß auf den unterschiedlichen Intonationsverlauf der Kontextsätze zurückzuführen sein. Eine eingehende Diskussion stellen wir bis zum nächsten Experiment zurück.

Fig.7: Kontexttest 2, Fragekontext



Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0051$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
<u>0.0252</u>	<u>0.0028*</u>	<u>0.4860</u>	0.5302	0.3152

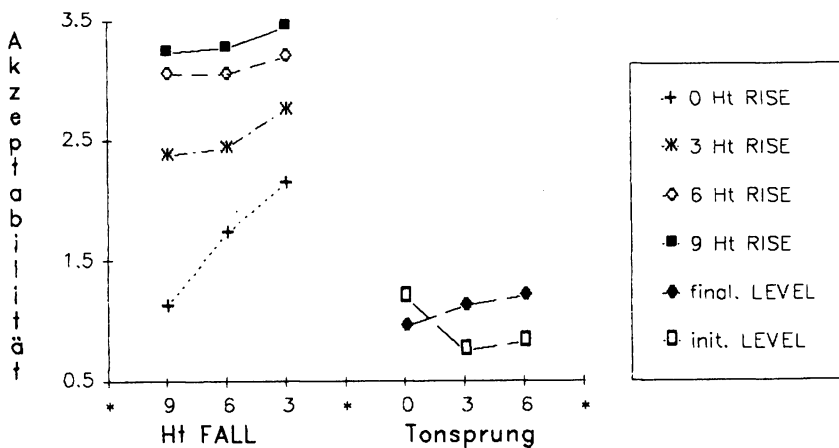
Fig.8 zeigt die Ergebnisse für den Exklamativkontext und die Signifikanzen der erhaltenen p-Werte. Die Verläufe von 0 und 3 Ht RISE entsprechen denen in Fig.6; sie unterscheiden sich signifikant. Allerdings werden die 'schlechten' Reihen 6 und 9 Ht RISE deutlich schlechter als in Fig.6 bewertet; daher auch der zusätzliche signifikante Unterschied zwischen 3 und 6 Ht RISE. Dieses Ergebnis ist überraschend; wenn kein unkontrollierter VPn-Effekt vorliegt, so muß es darauf zurückgeführt werden, daß die Tatsache, daß alle Kontextsätze einen

fallenden Fo-Verlauf aufweisen, eine Auswirkung auf die Bewertung der Testsätze mit deutlich steigenden Verläufen (6 und 9 Ht RISE) hat.

Es bleibt aber festzustellen, daß die **fragetypischen** Testsätze mit deutlichem RISE bei der Kombination mit einem FALL im frageindizierenden Kontextsatz (KT2) deutlich schlechter bewertet werden als bei der Kombination mit einem RISE im frageindizierenden Kontextsatz (KT1), die **exklamativtypischen** Testsätze – bei Kombination mit dem gleichen exklamativindizierenden Kontextsatz wie in KT1 – dagegen in KT2 gleich gut.

Die TS-Stimuli ergeben etwa das gleiche Bild wie in KT1, d.h. eine schlechte Bewertung beim Frage- und eine gute beim Exklamativkontext.

Fig.8: Kontexttest 2, Exklamativkontext



Paarweiser Vergleich der Klassen mit Wilcoxon ($\alpha = 0.0051$)

0 Ht RISE	3 Ht RISE	6 Ht RISE	9 Ht FALL	6 Ht FALL
3 Ht RISE	6 Ht RISE	9 Ht RISE	6 Ht FALL	3 Ht FALL
0.0019*	0.0019*	0.0229	<u>0.0121</u>	<u>0.0039*</u>

3.3.3 Kontexttest 3 (KT3)

Beim zuletzt besprochenen Experiment konnte nicht eindeutig entschieden werden, ob die beobachteten Effekte nur auf einen systematischen Faktor (Anforderung der intonatorischen Kongruenz, d.h. schlechtere Bewertung von Sätzen mit gegenläufigen Konturen) oder zusätzlich auf einen anderen Faktor zurückzuführen sind. Wir beschreiben deshalb nun als KT3 ein Teilexperiment, das zusammen mit den in Batliner (1988b) dargestellten Experimenten durchgeführt, dort aber nicht weiter erwähnt wurde.

3.3.3.1 Design

Das experimentelle Design ist ausführlich in Batliner (1988b:224ff) beschrieben; wir beschränken uns deshalb auf die wichtigsten Angaben:

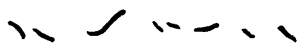
Im Tonstudio wurden vom Autor jeweils mehrere Realisationen der in Fig.9 aufgeführten Kontext- und Testsätze produziert und mit 19 cm/sec auf einer Telefunken M15 aufgenommen. Nach einer Beurteilung durch mehrere kompetente Sprecher und einer Überprüfung der relevanten phonetischen Parameter anhand von Mingogrammen mit Zeitsignal, Fo- und Intensitätskurve wurde je ein typisches Exemplar ausgewählt. Jeder Kontextsatz wurde mit jedem Testsatz gepaart und (zusammen mit anderen Kontext-Testsatz-Kombinationen, vgl. Batliner 1988b) in randomisierter Reihenfolge im Tonstudio analog auf ein Testband kopiert; das Intervall zwischen Kontext- und Testsatz betrug ca. 300-800 msec, das zwischen den Paaren ca. 4 sec. 22 VPn nahmen teil; die Aufgabenstellung war die gleiche wie in KT1 und KT2. Abgesehen davon, daß jedes Paar nur einmal vorkam, ist also der wesentliche Unterschied zwischen KT1 und KT2 auf der einen und KT3 auf der anderen Seite, daß in KT3 nicht nur die Kontextsätze, sondern auch die Testsätze natürlich produziert waren. Zur Verdeutlichung werden bei KT3 die Exklamativ-Testsätze 'E-Testsätze' und die Frage-Testsätze 'F-Testsätze' genannt; wir beschränken uns im folgenden auf die 21 Kombinationen der sieben Testsätze aus Fig.9 mit dem frageindizierenden Kontextsatz *Bist du sicher?* mit den drei unterschiedlichen Fo-Verläufen aus Fig.9, da alle anderen Kombinationen schon in Batliner (1988b) behandelt wurden.² In Fig.9 sind bei den Testsätzen die Phrasen, die den Hauptakzent tragen, unterstrichen.

² Die beiden E-Testsätze mit Satzakkzent auf *säuft* unterscheiden sich durch den Fo-Verlauf auf *säuft*, der einmal fallend, das andere Mal leicht steigend und dann fallend ist (vgl. Fig.9). Der erste Fall entspricht also dem FR-Stimulus 9|0, der zweite Fall weist den exklamativtypischen RISE-FALL-Verlauf mit Dehnung auf.

Fig.9: Schematische Fo-Verläufe der Kontext- und der Testsätze

KONTEXTSÄTZE:

exklamativindizierend:



Also sowas hätt' ich nicht gedacht!


frageindizierend:



Bist du sicher?
RISE




Bist du sicher?
RISE-FALL



Bist du sicher?
FALL


TESTSÄTZE, Satzakzent unterstrichen:

Exklamativtestsätze
'E-Testsätze':


Säuft der Leo!



Säuft der Leo!

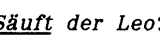


Säuft der Leo!


Fragetestsätze
'F-Testsätze':


Säuft der Leo?

FALL:

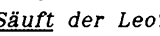


Säuft der Leo?



Säuft der Leo?

RISE:



Säuft der Leo?



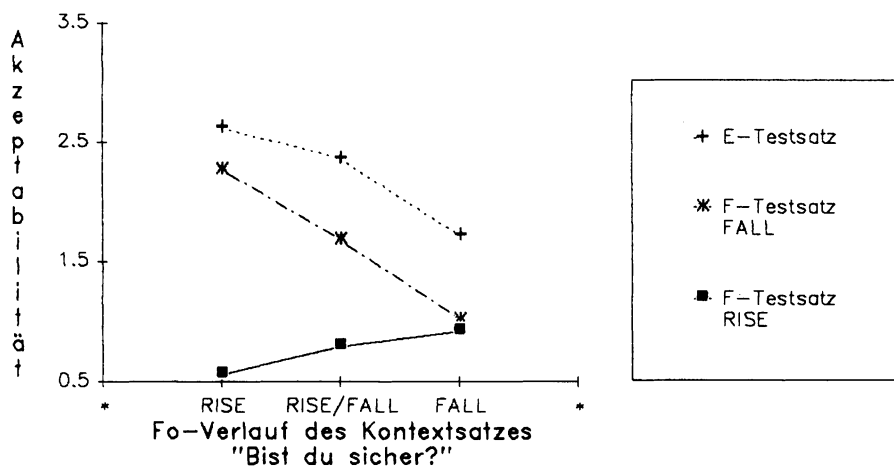
Säuft der Leo?

3.3.3.2 Ergebnis

Die Darstellung in Fig.9 und die der Ergebnisse in Tab.3 und Fig.10 sind zur Erleichterung des Vergleichs möglichst ähnlich aufgebaut, analog zu einer Matrix mit drei Spalten (Kontextsätze) und drei Zeilen (Testsätze); die Reihenfolge ist identisch. Der Kontextsatz *Bist du sicher?* indiziert keinen bestimmten Satzakkzent; dementsprechend sind die Unterschiede bei der Moduszuweisung zwischen den Einzelwerten für die zwei unterschiedlichen Positionen des Satzakkzents nicht groß, vgl. Tab.3; die Werte können also gemittelt werden. ('S' steht für 'Satzakkzent auf *säuft*', 'L' für 'Satzakkzent auf *Leo*'.) Mit dieser Mittelung kann die Verteilung der Werte normalisiert und zugleich die Zahl der prüfstatistischen Vergleiche verringert werden. Die Mittelwerte aus den in Tab.3 aufgeführten Einzelwerten pro Kästchen sind in Fig.10 aufgetragen. Wieder wurde nicht-parametrisch getestet (Wilcoxon für abhängige Stichproben, zweiseitige Fragestellung, abgesenktes Signifikanzniveau $\alpha = 0.0046$). Die erhaltenen p-Werte sind in Tab.3 aufgeführt und, sofern signifikant, mit einem Stern gekennzeichnet; die p-Werte stehen immer zwischen den verglichenen Paaren, wobei jeweils ein Faktor wechselt, der andere konstant bleibt.

Es war zu erwarten, daß die E-Testsätze bei allen drei Fragekontextsätzen am schlechtesten abschneiden. Allerdings ist der Verlauf von RISE über RISE-FALL zu FALL in Fig.10 fallend, mit einem signifikanten Unterschied zwischen den letzten beiden Klassen. Das gleiche gilt für die F-Testsätze mit FALL-Verlauf, die sich aber für RISE/FALL- und FALL-Verlauf des Kontextsatzes signifikant von den E-Testsätzen unterscheiden. Die F-Testsätze mit FALL- und die mit RISE-Verlauf sind klar gegenläufig, mit zwei signifikanten Unterschieden bei RISE- und RISE/FALL-Verlauf des Kontextsatzes. In varianzanalytischer Terminologie können wir also von einer **Interaktion** zwischen den beiden Faktoren Fo-Verlauf auf dem Kontextsatz und Fo-Verlauf auf dem Testsatz sprechen.

Fig.10: Kontexttest 3 (natürlichsprachliche Testsätze)

Tab.3: Paarweiser Vergleich der benachbarten Klassen (Wilcoxon, zweiseitig geprüft, abgesenktes Signifikanzniveau $\alpha = 0.0046$)

		<u>KONTEXTSATZ</u>					
		RISE		RISE-FALL		FALL	
<u>TESTSATZ</u>							
EXKLAMATIV	S	4.00	$p=0.0687$	3.19	$p=0.0005^*$	2.43	$p=0.0009^*$
	S	3.41		3.57		2.82	
	L	3.45		3.33		2.91	
		$p=0.0720$		$p=0.0040^*$			
FRAGE FALL	S	3.23	$p=0.0033$	2.45	$p=0.0026^*$	1.59	$p=0.6155$
	L	3.32		2.95		2.45	
				$p=0.0001^*$			
FRAGE RISE	S	1.45	$p=0.3066$	1.76	$p=0.7223$	1.76	
	L	1.68		1.82		2.09	

Ein kongruenter steigender Verlauf auf Kontextsatz und Testsatz schneidet am besten ab; eine Kombination Kontextsatz/FALL + Testsatz/RISE ist nicht viel schlechter, und beide sind bedeutend besser bewertet als die Kombination Kontextsatz/RISE + Testsatz/FALL. Es kann sich nicht um eine simple Oberflächen-

kongruenz handeln, da zum einen die Kombination der beiden gegenläufigen Konturen Kontextsatz/FALL + Testsatz/RISE gut bewertet wird, und da zum anderen die hier schlecht bewertete Abfolge Kontextsatz/RISE + Testsatz/FALL in anderem Zusammenhang, bei geschlossenen Alternativfragen, völlig normal ist. So ist die Abfolge *Möchtest Du Zucker?* mit steigendem und *Oder möchtest Du Zitrone?* mit fallendem Verlauf bei einem Fünf-Uhr-Tee dann, wenn keine Sahne oder Milch vorhanden sind, die natürliche (vgl. dazu in Batliner/Oppenrieder 1989 die geschlossenen Alternativfragen im 2. Korpus, Nr. 42 und 44).

Um den Unterschied erklären zu können, wollen wir drei Annahmen machen:

1. Ein steigender Intonationsverlauf bei einem Verb-Erst-Fragesatz (der ja regulär auch mit einem fallenden Verlauf produziert werden könnte) indiziert ein gesteigertes Interesse des Sprechers an einer Antwort (Appellfunktion).
2. Im unmarkierten Regelfall hat ein fallender Intonationsverlauf auch beim Verb-Erst-Fragesatz eine geringe(re) Appellfunktion.
3. Was am Ende eines Redebeitrags (*Turn*) signalisiert wird, ist wesentlich für seine funktionale Bedeutung.

Wenn nun sowohl Kontextsatz als auch Testsatz fallenden oder steigenden Intonationsverlauf aufweisen, dann besteht auch funktional zwischen ihnen eine Kongruenz, d.h. beide Kombinationen sind akzeptabel. Hat der Kontextsatz fallenden und der Testsatz steigenden Verlauf, so entspricht die Abfolge 'wenig Appell - Appell' einem unmarkierten Redebeitrag, an dessen Ende die Aufforderung zur Gesprächsübernahme (*Turn-Taking*) steht. Wenn die Reihenfolge aber umgekehrt ist, so folgt einem deutlichen Appell, indiziert durch den steigenden Intonationsverlauf, eine Rücknahme des Appells, indiziert durch den fallenden Verlauf und verstärkt durch die Position am Ende des Satzes. Die 'Botschaft' des Sprechers ist damit widersprüchlich und deshalb weniger akzeptabel.

3.4 Diskriminationstest (DT)

Bei den bisherigen Experimenten hat es sich gezeigt, daß ein systematischer Zusammenhang zwischen der Ht-Manipulation auf der einen und der Frage/Exklamativ-Zuweisung auf der anderen Seite existiert; m.a.W. die unterschiedliche Ausprägung der Ht auf FALL und RISE ist linguistisch/phonetisch relevant. Es soll nun untersucht werden, inwiefern sie im gleichen Ausmaß auditiv/psychoakustisch relevant ist, und ob sich in einem weiteren Schritt aus Ähnlichkeits- und Unähnlichkeitsrelationen zwischen unseren Stimuli Faktoren erschließen lassen, die die Bewertungsfunktionen der bisherigen Experimente 'erklären' können. Die Beziehung zwischen der Ht-Manipulation und den lingui-

stischen Kategorien kann linear und trivial sein: ein ausgeprägter RISE indiziert eine Frage, ein fehlender RISE eine Nicht-Frage. Die Beziehung kann aber auch nicht-trivial sein, wenn sich etwa ein komplexer und nicht-linearer Zusammenhang zwischen Merkmalen und Kategorialität ergibt. Um diesen Fragen nachzugehen, wurden ein Diskriminationstest (DT, vgl. Teil 3.4) und ein Ähnlichkeitstest (ÄT, vgl. Teil 3.5) durchgeführt.

Im DT werden benachbarte oder auch entferntere Stimuli eines physikalisch äquidistanten Kontinuums, die zu Paaren (AX-Test) zusammengestellt sind, auf Gleichheit beurteilt. Es werden die Stimulusabfolgen AB, BA und die 'gleich'-Paare AA bzw. BB dargeboten. Im Paradigma der Kategorialen Wahrnehmung (vgl. Repp 1984) wird angenommen, daß Paare, innerhalb derer eine Kategoriengrenze liegt, deutlich öfter als 'verschieden' bewertet werden als Paare, bei denen beide Stimuli zur gleichen Kategorie gehören. In unserem Fall sind die Kategorien Frage und Exklamativ (bzw. Nicht-Frage). Das Konzept des strikten Paradigmas der Kategorialen Wahrnehmung und die dort durchgeführte Berechnung der Kategorialität wurden von uns in Schiefer/Batliner (1988) kritisiert; vgl. dazu auch die Arbeiten in Harnad (1987). Das ändert aber nichts daran, daß wir es auch hier mit Kategorien zu tun haben, und daß die grundsätzliche Fragestellung sinnvoll ist, ob solche Kategoriengrenzen, wie wir sie in den bisher besprochenen Experimenten gefunden haben, auch bei der Diskrimination sichtbar werden.

3.4.1 Design

Bei den TS-Stimuli ist ein Diskriminationstest nicht sinnvoll, da sie nicht eindeutig zwei verschiedenen Kategorien angehören, und da sie auch kein 'echtes' Kontinuum bilden, sondern höchstens zwei Kontinua mit je drei Stimuli. Wir beschränken uns deshalb auf die FR-Stimuli. In einem Diskriminationstest, der auf drei Sitzungen aufgeteilt war, wurden für jede Kombination benachbarter Stimuli pro Spalte und pro Zeile (vgl. Tab.1) die Anordnung AB und BA sowie die 'gleich'-Paare AA und BB je dreimal randomisiert dargeboten; insgesamt waren es 46 Paare. Die Instruktion lautete: "Bitte entscheiden Sie, ob die Stimuli innerhalb eines Paares in allen Belangen gleich klingen, und drücken Sie dann die linke Taste für 'gleich' oder die rechte für 'verschieden'." Es nahmen 11 VPn teil.

Anhand der Matrix von Tab.1 lassen sich zwei Kontinua bilden, wenn man von Stimulus 9|0 ausgehend im Uhrzeigersinn oder im Gegenuhrzeigersinn zu Stimulus 3|9 geht. Beim ersten Kontinuum wird das exklamativindizierende Merkmal initialer FALL auf 3|0 abgesenkt und dann das frageindizierende Merkmal RISE angehoben. Beim zweiten ist es umgekehrt: Das frageindizierende Merkmal RISE wird angehoben, während der FALL auf 9 Htn bleibt; ab 9|9 wird der FALL abgesenkt. Im zweiten Kontinuum gibt es also 'widersprüchliche' Merkmalsausprägungen: hoher FALL und hoher RISE. Es interessiert in beiden Fällen, ob und wo man einen Diskriminationsgipfel beobachten kann. Da darüberhinaus auch alle anderen möglichen Paare, also auch solche, die nicht zu den beiden Kontinua gehören, getestet wurden, kann grundsätzlich untersucht werden, ob sich ein 3 Ht-Unterschied bei FALL und bei RISE unterschiedlich auswirkt.

3.4.2 Ergebnis

In Fig. 11 sind das Ergebnis für das erste Kontinuum, in Fig.12 das Ergebnis für das zweite dargestellt. Neben den Verläufen für die Anordnungen AB und BA sind die Verläufe für die 'Gleich'-Paare und das Ergebnis des IT (vgl. Fig.3) angegeben. Die BA-Kurve zeigt bei beiden Kontinua eher als die AB-Kurve einen Diskriminationsgipfel korrespondierend zum Zufallsbereich beim IT (Stimulus 9|3 bzw. 3|3).³ Wir können also - wenn wir uns mit dieser introspektiven Betrachtung der Kurvenverläufe begnügen - feststellen, daß Identifikations- und Diskriminationstest die gleiche Kategoriengrenze bei beiden Kontinua ergeben: beim RISE von drei Htn.⁴

³ Üblicherweise werden AB- und BA-Werte gemittelt und nur diese gemittelten Werte weiter betrachtet; vgl. aber die Überlegungen in Schiefer/Batliner (1988) dazu, daß die Stimulusabfolge einen ähnlichen Status wie die Schrittgröße hat. D.h. man kann sich für die Interpretation entweder auf AB oder auf BA beschränken.

⁴ Interessant ist, daß in Fig.11 zwar alle 'Gleich'-Paare sehr gut als 'gleich' erkannt wurden, daß sich aber in Fig.12 genau bei der Kategoriengrenze (Stimuli 9|3 und 9|6) ein deutlicher Gipfel zeigt; eine mögliche Erklärung für dieses Phänomen findet sich in Schiefer/Batliner (1988:279f).

Fig.11: Diskriminationstest (9/0 - 3/0 - 3/9)

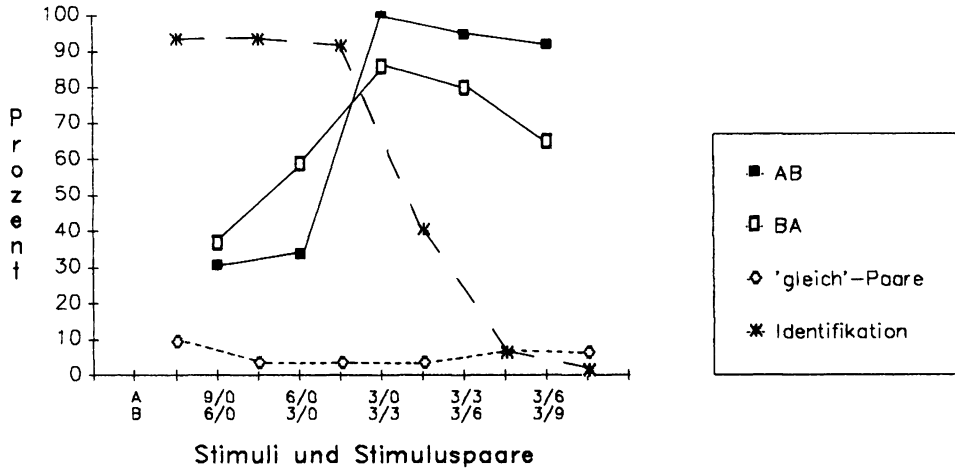
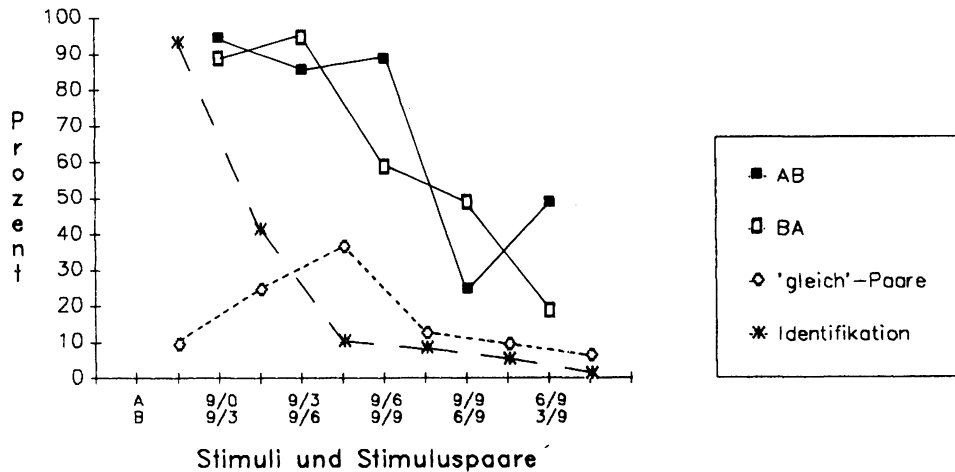


Fig.12: Diskriminationstest (9/0 - 9/9 - 3/9)



In Tab.4 sind zum einen - in der Anordnung der Matrix wie in Tab.1 - die gemittelten 'Verschieden'-Antworten in Prozent für jeden Vergleich angeführt. Tabellarisch folgen die Bewertungen aufgeteilt nach AB und BA. Zu den schon anhand von Fig.11 und 12 gemachten Feststellungen lassen sich zwei weitere hinzufügen:

1. Der finale RISE wird durchgehend besser diskriminiert als der initiale FALL. Das entspricht bekannten Ergebnissen, vgl. Rossi (1978); man muß allerdings beachten, daß bei unseren Stimuli der RISE immer länger war als der FALL, vgl. Fig. 1; der Faktor Dauer spielt sicher auch eine Rolle.
2. Der in Batliner/Schiefer (1987) und Schiefer/Batliner (1988) besprochene Ordnungseffekt tritt bei den RISE-Paaren deutlich in Erscheinung: Wenn der höhere RISE an zweiter Stelle steht, werden die Stimuli besser diskriminiert als wenn er an erster Stelle steht.

Tab.4: Diskrimination zwischen benachbarten Stimuli, gemittelt und aufgeschlüsselt nach AB und BA.

1. Mittelwerte aus AB und BA, Matrixanordnung

9 0	33%	6 0	46%	3 0
91%		90%		92%
9 3	44%	6 3	56%	3 3
90%		87%		86%
9 6	41%	6 6	38%	3 6
73%		80%		77%
9 9	36%	6 9	33%	3 9

2. AB und BA getrennt

a) Variation des initialen FALL

	9 0-6 0	6 0-3 0	9 3-6 3	6 3-3 3	9 6-6 6	6 6-3 6	9 9-6 9	6 9-3 9
AB	30%	33%	58%	76%	36%	48%	24%	48%
BA	36%	58%	30%	36%	45%	27%	48%	18%

b) Variation des finalen RISE

	9 0-9 3	9 3-9 6	9 6-9 9	6 0-6 3	6 3-6 6	6 6-6 9	3 0-3 3	3 3-3 6	3 6-3 9
AB	94%	85%	88%	94%	88%	91%	100%	94%	91%
BA	88%	94%	56%	85%	85%	70%	85%	79%	64%

3.5 Ähnlichkeitstest (AT)

3.5.1 Fragestellung und statistisches Modell

Im DT konzentriert sich das Interesse auf das Stimuluspaar, zwischen dem die erwartete Kategoriengrenze liegt. Die 'Verschieden'-Antworten können dabei als Abstände zwischen den Stimuli ein Maß für die Ähnlichkeit dieser Stimuli sein. Da allerdings immer nur benachbarte Stimuli auf ihre Ähnlichkeit/Unähnlichkeit hin überprüft werden, ist das sich ergebende Bild (vgl. Tab.4) nicht hinreichend genau. Ein genaueres Bild kann man mit einem Experiment erreichen, bei dem jeder Stimulus mit jedem anderen Stimulus gepaart ist, dem *All-Step-Test*. Dabei können auch die TS-Stimuli bewertet werden. Man erhält in unserem Fall bei 18 Stimuli eine 18x18-Matrix mit Ähnlichkeitswerten, vgl. unten Tab.5. Eine solche Matrix enthält alle nötige Information, ist aber – gerade deshalb – kaum interpretierbar, weil die Information ausgesprochen unübersichtlich dargeboten ist. Ein Verfahren zur anschaulichen, (mehr oder weniger) leicht interpretierbaren Darstellung ist die Nonmetrische Multidimensionale Skalierung (NMDS). Die Grundzüge dieses Verfahrens werden von Kruskal/Wish (1978:7) wie folgt beschrieben:

Multidimensional scaling [...] refers to a class of techniques. These techniques use *proximities* among any kind of objects as input. A proximity is a number which indicates how similar or how different two objects are, or are perceived to be, or any measure of this kind. The chief output is a spatial representation, consisting of a *geometric configuration of points*, as on a map. Each point in the configuration corresponds to one of the objects. This configuration reflects the "hidden structure" in the data, and often makes the data much easier to comprehend. By reflecting the data structure we mean that the larger the dissimilarity (or the smaller the similarity) between the two objects, as shown by their proximity value, the further apart they should be in the spatial map.

Nonmetrisch ist die Skalierung dann, wenn die Abstände zwischen den Items nicht als intervallskaliert (metrisch) aufgefaßt werden, sondern als ordinalskaliert (nonmetrisch). In die Analyse gehen in einem solchen Fall nicht die Abstände, sondern die Rangplätze ein.⁵ Das Verfahren der NMDS ist nicht konfirmatorisch in dem Sinn, daß es den strikten Anforderungen genügt, die an ein prüfstatistisches Verfahren gestellt werden; solche Anforderungen können aller-

⁵ Die NMDS ist ausführlich beschrieben in Borg (1981), eine Anwendung auf Modus/Fokus zeigt Batliner (1988b); dort werden auch einige Kritikpunkte diskutiert. Eine multidimensionale Skalierung von Tonverläufen und Tönen wurde von Gandour (1978) und Gandour/Harshman (1978) durchgeführt.

dings auch in Frage gestellt werden, cf. Henkel (1976) und Guttman (1977). Die NMDS ist ein anschauliches Verfahren, das nutzbringend explorativ eingesetzt werden kann; als solches sei es auch im folgenden verstanden.

3.5.2 Design

Jeder der 18 FR- und TS-Stimuli wurde mit jedem anderen und mit sich selbst gepaart (324 Paare); die Stimuli wurden pseudorandomisiert auf 6 Sitzungen mit je 54 Stimuluspaaren aufgeteilt. (Die Pseudorandomisierung sah so aus, daß jeder Stimulus dreimal an erster Stelle vorkam und die zweite Stelle nach einem Zufallsverfahren vergeben wurde; damit konnte eine, bei einer 'echten' Randomisierung immer mögliche, ausgeprägte ungleiche Verteilung zwischen den einzelnen Teilsitzungen vermieden werden.) Die Pause innerhalb der Paare betrug 0.5 sec, die zwischen den Paaren 4 sec. Die Anweisung lautete: "Bitte beurteilen Sie die Ähnlichkeit innerhalb eines jeden Paares und bewerten Sie diese Ähnlichkeit auf einer Skala von 1 bis 5 (1: 'kein Unterschied = identisch', 2: 'ähnlich', 3: 'noch ähnlich', 4: 'unähnlich', 5: 'sehr unähnlich'). Es kommen Paare mit identischen Stimuli vor, Paare mit sehr unähnlichen, und Paare mit mehr oder weniger ähnlichen Stimuli." Die sechs Testsitzungen wurden im Abstand von ein bis drei Wochen abgehalten; die Zahl der Teilnehmer schwankte (12, 16, 16, 18, 18, 19); sechs Teilnehmer nahmen an allen Sitzungen teil, die anderen an weniger.⁶

Die NMDS-Lösungen (getrennt für die FR-Stimuli, die TS-Stimuli und für FR- und TS-Stimuli zusammen) wurden mit der SPSS-Prozedur ALSCAL am Leibniz-Rechenzentrum in München berechnet. Eine getrennte und eine gemeinsame Berechnung der FR- und der TS-Stimuli ermöglichte es, innerhalb der beiden Gruppen eindeutiger Dimensionen und über die Gruppen hinweg beiden gemeinsamen Dimensionen zu ermitteln.

⁶ Die relativ geringe Zahl der Teilnehmer und die nicht ganz regelmäßige Besetzung der Gruppen kann natürlich den experimentellen Fehler vergrößern; mehr Teilnehmer und ein ausbalanciertes Verhältnis unter den Gruppen hätte aber den Aufwand stark erhöht. Da die NMDS in unserem Zusammenhang als exploratives Verfahren eingesetzt wurde, schien dieses Vorgehen vertretbar. Indiz dafür, daß die erzielten Lösungen relativ stabil sind, war ein Vergleich der sechs VPn, die an allen Tests teilnahmen, mit den anderen: Bei den niedrigen Dimensionen 1 und 2 waren die Lösungen praktisch identisch.

3.5.3 Bestimmung der Dimensionalität

Tab.5 zeigt die Matrix der Ähnlichkeitswerte; die Werte für die 'gleich'-Paare in der Diagonalen sind fett und kursiv. Mit ALSCAL wurden ein- bis vierdimensionale Lösungen berechnet; als Abstandsmaß wurde das euklidische genommen, die Matrix wurde nicht symmetrisiert, sondern ging als asymmetrische in die Berechnungen ein. (Der beim Diskriminationstest erwähnte Ordnungseffekt ist wahrscheinlich zum großen Teil für die Asymmetrie verantwortlich; auf ihn kann aber in diesem Zusammenhang nicht weiter eingegangen werden.) Es gibt kein 'wasserdichtes' Kriterium dafür, wieviele Dimensionen man für eine NMDS-Lösung akzeptieren kann; zumindest die folgenden groben Kriterien können aber bei der Entscheidung behilflich sein:

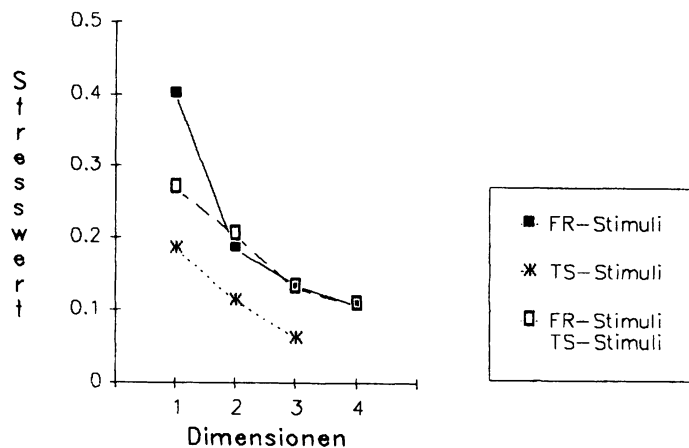
1. Das Gütemaß der Lösung, bei ALSCAL der sog. 'Stress-Wert': Je kleiner dieser Wert ist, desto 'besser' ist die Lösung. (Ein Stress-Wert unter 0.10 gilt als 'gut'; eine blinde Anwendung dieses oder eines ähnlichen Kriteriums ist aber sinnlos.)
2. Ein sog. 'Ellbogen' bei einer grafischen Darstellung des Stress-Wertes vs. Zahl der Dimensionen im 'Stress-Diagramm', vgl. Fig. 13: ein Ellbogen tritt dann auf, wenn höhere Dimensionen wenig und niedrigere viel zur Verbesserung der Lösung beitragen.
3. Das Maß der 'erklärten Varianz' (R^2). Je mehr der Wert gegen 1.0 geht, desto 'besser' wird die Varianz erklärt.
4. Die 'Eindeutigkeit' der Lösung: eine Lösung wird dann als eindeutig angenommen, wenn die Zahl der Items mindestens viermal größer ist als die Zahl der Dimensionen.
5. Die Interpretierbarkeit der Lösung.

Tab.5: Ähnlichkeitsmatrix

	9 0	6 0	3 0	3 3	3 6	3 9	6 3	6 6	6 9	9 3	9 6	9 9	0 0	3 0	6 0	0 9	0 6	0 3
9 0	0.11	0.50	1.75	2.72	2.37	3.41	1.33	1.81	2.13	1.11	2.17	2.29	2.50	1.61	0.47	2.83	2.62	2.37
6 0	0.50	0.17	0.89	1.47	1.83	2.56	1.28	1.87	2.17	1.38	2.17	2.78	2.50	2.00	0.58	3.31	3.19	3.00
3 0	1.08	0.31	0.67	1.16	1.58	2.50	0.88	2.06	2.05	1.81	2.33	2.44	3.31	2.92	1.87	3.89	2.05	2.89
3 3	1.44	1.11	0.87	0.17	1.44	1.94	3.16	1.00	2.08	1.25	2.00	1.94	3.58	3.33	0.84	3.78	3.56	1.89
3 6	2.50	1.33	1.50	0.58	0.16	0.56	0.75	0.19	1.00	1.82	1.19	1.72	3.32	3.33	2.17	3.31	3.11	3.31
3 9	2.25	1.94	1.74	1.13	0.37	0.44	1.50	1.06	0.19	2.00	1.21	1.17	3.06	3.39	2.72	3.83	3.50	3.17
6 3	1.25	0.89	1.72	0.68	1.00	1.69	0.06	1.17	1.88	0.25	1.44	1.94	2.89	2.19	1.67	3.61	3.33	3.11
6 6	1.75	1.61	1.25	1.00	0.58	1.13	0.68	0.06	1.17	1.44	0.37	1.44	3.16	2.94	2.00	3.44	3.25	3.25
6 9	2.06	1.62	2.39	1.33	0.56	0.19	1.11	0.38	0.28	2.06	1.13	0.47	3.33	2.83	2.42	3.18	2.68	2.88
9 3	1.28	1.56	1.75	2.08	2.25	2.68	1.11	1.37	2.06	0.33	1.17	1.69	2.47	2.06	1.17	3.19	2.56	2.94
9 6	1.31	1.94	1.58	2.50	1.79	1.94	0.94	0.33	0.94	0.74	0.25	0.69	3.17	1.87	1.47	3.06	2.94	2.94
9 9	1.75	1.89	2.39	1.58	1.06	1.11	1.44	0.75	0.39	1.50	0.42	0.31	3.22	3.06	1.63	3.31	3.58	3.37
0 0	2.13	2.58	3.13	3.28	3.37	3.58	3.67	3.56	3.13	2.32	3.00	3.56	0.11	0.83	1.38	1.16	1.31	0.78
3 0	1.06	1.81	2.89	3.33	3.37	3.56	2.50	2.53	2.81	1.75	2.67	2.42	0.92	0.13	0.89	2.94	1.56	1.17
6 0	0.25	1.61	2.25	2.58	3.67	3.32	1.75	2.89	2.67	1.61	1.95	2.50	2.00	1.06	0.28	2.88	2.00	2.19
0 9	3.39	3.08	3.44	3.58	3.50	3.89	3.21	3.33	3.11	2.94	2.95	3.33	1.38	1.89	2.56	0.06	0.39	1.31
0 6	2.17	2.72	2.92	3.89	3.44	3.68	2.94	3.83	3.37	2.67	3.13	3.37	0.75	1.89	1.67	0.31	0.42	0.31
0 3	2.31	3.11	3.42	3.62	3.42	3.78	2.58	3.56	3.67	3.33	3.33	3.05	0.44	0.73	1.84	1.19	0.63	0.45

Fig.13 zeigt das Stress-Diagramm, Tab. 6 die Stress-Werte, die 'erklärte Varianz' R^2 , die Eindeutigkeit ('+' für 'eindeutig', '-' für 'nicht eindeutig') sowie die Interpretierbarkeit ('+' für 'interpretierbar', '-' für 'nicht interpretierbar'; vgl. im einzelnen 3.5.3. - 3.5.5.) pro Dimension.

Fig.13: Stress-Diagramm



Tab.6: Kriterien zur Wahl der Dimensionalität

Stimuli	Dimensionen	Stress	R ²	Eindeutig- keit	Interpretier- barkeit
FALL-RISE	4	.108	.92	-	+
	3	.135	.89	+	+
	2	.185	.84	+	+
	1	.399	.59	+	+
TONSPRUNG	3	.062	.98	-	-
	2	.113	.95	-	+
	1	.185	.91	+	+
FALL-RISE + TONSPRUNG	4	.107	.93	+	-
	3	.131	.91	+	+
	2	.204	.85	+	+
	1	.269	.82	+	+

Man sieht, daß die Kriterien 1 bis 4 sich z.T. widersprechen: bei den FR-Stimuli zeigt sich z.B. ein Ellbogen bei Dimension 2, der Stress-Wert ist aber dabei noch recht hoch. Wir werden uns deshalb auf das fünfte Kriterium der

Interpretierbarkeit der Lösungen stützen, auch wenn dabei die Lösungen noch mit einem relativ hohen Stress-Wert belastet oder uneindeutig im o.a. Sinn sein können; dieses Vorgehen ist bei einer solchen explorativen Studie sinnvoller als ein mechanisches Anwenden der anderen Kriterien.

Fig.14 bis 18 zeigen in jeweils zweidimensionaler Darstellung die interpretierbaren Dimensionen (DIM): Fig.14 DIM1 und DIM2 sowie Fig.15 DIM3 und DIM4 für die FR-Stimuli, Fig.16 DIM1 und DIM2 für die TS-Stimuli und Fig.17 DIM1 und DIM2 sowie Fig.18 DIM2 und DIM3 für die FR- und TS-Stimuli. Zur besseren Vergleichbarkeit sind die FR-Stimuli an ihrer Position mit den Zeichen wiedergegeben, die ihnen auch in Fig. 3 bis Fig.8 zugewiesen waren: '■' für 9 Ht RISE, 'o' für 6 Ht RISE, '*' für 3 Ht RISE und '+' für 0 Ht RISE. Die TS-Stimuli werden einheitlich mit '#' wiedergegeben.

3.5.4 FALL-RISE-Stimuli

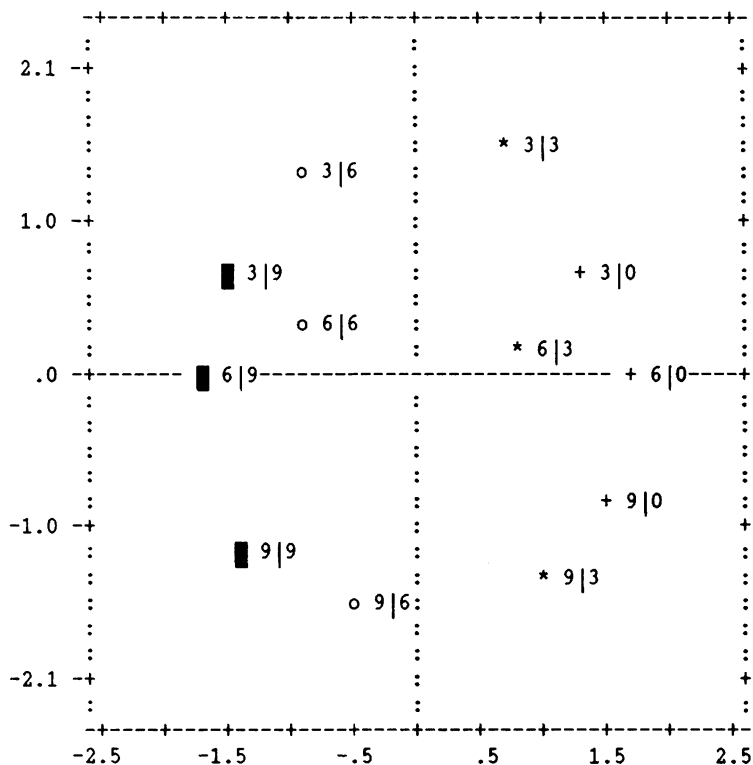
DIM1 kann als Ausprägung des finalen RISE = Höhe des Offset interpretiert werden, DIM2 als Ausprägung des initialen FALL = Höhe des Onset, cf. Fig.14. Die Dimensionen stehen ziemlich orthogonal zum Achsenkreuz. (Man beachte, daß das keine Bedingung ist; die Punkte können beliebig rotiert oder gespiegelt werden.) Die Stimuli mit 9 und 6 Ht RISE liegen auf DIM1 nahe beieinander, ebenso die mit 3 und 0 Ht RISE; die mit 6 und 3 Ht RISE liegen weiter auseinander. Diese Positionen widerspiegeln die Kategoriengrenze zwischen Frage und Nicht-Frage bzw. Exklamativ, die bei den anderen Experimenten auch zwischen 3 und 6 Ht RISE lag. Bei DIM2 liegt 6|6 näher an 3|6 als an 9|6, und 6|9 näher an 3|9 als an 9|9. Verantwortlich dafür ist wahrscheinlich die geringere perzeptive Relevanz des FALL im Verhältnis zum RISE; d.h. ein größeres Ausmaß des FALL ist nötig, um deutliche Unterschiede zu markieren.

Die zweidimensionale Darstellung von DIM1 und DIM2 entspricht der Matrix von Tab.1, die man erhält, wenn man die Punkte um 90° gegen den Uhrzeiger rotiert und dann um die Senkrechte spiegelt. Die beiden wichtigsten Dimensionen stehen also für die von uns vorgenommenen Manipulationen bei RISE und FALL; daß RISE die wichtigere ist, zeigte sich schon an der besseren Diskriminierbarkeit der RISE-Unterschiede im Diskriminationstest. Die Eckstimuli der Matrix 3|9, 3|0, 9|9 und 9|0 liegen näher zur Mitte hin, als das bei einer Matrixanordnung der Fall ist; sie 'erzeugen' damit die Kreisstruktur in Fig.14.⁷

⁷ Wir vermuten, daß man dieser Kreisstruktur keine 'kognitive Relevanz' zuschreiben kann, sondern daß sie sich schlicht aus dem experimentellen Design ergibt, bei dem die Ähnlichkeit nicht nur benachbarter, sondern aller Stimuli, also auch der unähnlichen, gewichtet ist.

Fig.14:FR-Stimuli

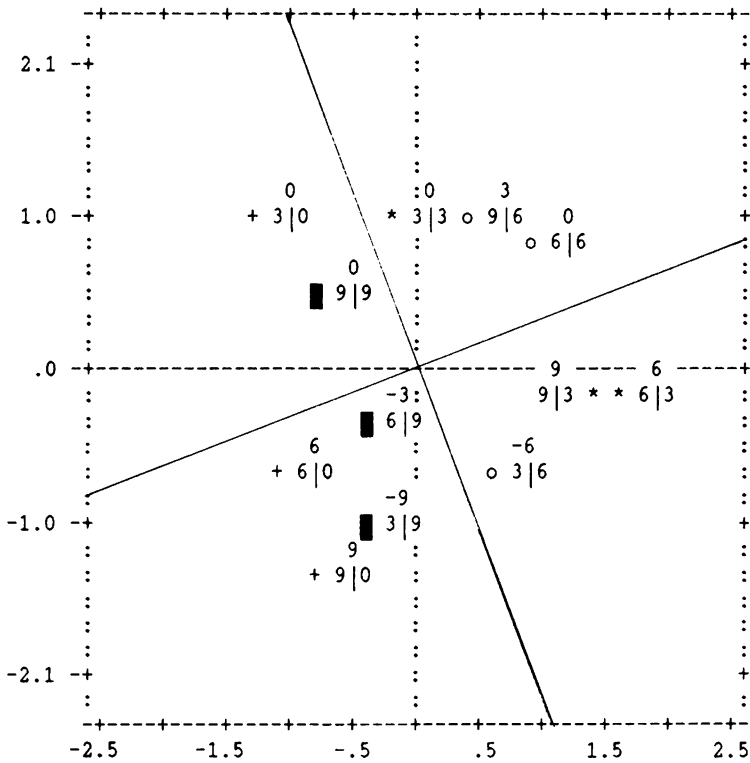
DIM1 (HORIZONTAL):finaler RISE / Höhe des Fo-Offset
 DIM2 (VERTIKAL):initialer FALL / Höhe des Fo-Onset



DIM3 und DIM4 tragen deutlich weniger zur erklärten Varianz bei als DIM1 und DIM2, vgl. Tab.6. Beide sind aber sinnvoll interpretierbar. Sie stehen zwar orthogonal zueinander, aber leicht rotiert im Verhältnis zum Achsenkreuz. In Fig.15 ist deshalb durch den Ursprung per Hand die Ausrichtung der Dimensionen eingetragen (durchgezogene Linien). Geht man bei DIM3, also in der Horizontalen, von links nach rechts, so kommen zuerst die Stimuli mit 0 Ht RISE, dann die mit 9 Ht RISE, dann, auf der rechten Seite, die restlichen mit 3 und 6 Ht RISE. Wir können also diese Dimension die des extremen Endpunkts nennen. (Eine solche Dimension, die nicht entlang eines Kontinuums angesiedelt ist, fanden auch Gandour 1978 und Gandour/Harshman 1978.)

Fig.15:FR-Stimuli

DIM3 (HORIZONTAL): Extremer Endpunkt
 DIM4 (VERTIKAL): Direktionalität der Kontur



Wenn wir im folgenden die 'Existenz' einer vierten Dimension ansetzen, dann natürlich unter der Voraussetzung, daß sie replizierbar ist: die Lösung ist in unserem Fall bei vier Dimensionen nicht mehr eindeutig, vgl. Tab.6. Betrachtet man nun die Stimuli von links oben nach rechts unten, also entlang der eingezeichneten Linie, so fällt auf, daß zuerst Stimuli kommen, bei denen sich 'wenig tut', wie 3|0 oder 3|3, zugleich aber auch Stimuli, bei denen sich 'viel tut', wie 9|9 und 6|6. Am anderen Ende sind Stimuli mit nur einem eindeutigen RISE (3|9) oder einem eindeutigen FALL (9|0) angesiedelt. Die Zuweisung zu den Positionen links und rechts von der Mitte wird eindeutig, wenn wir die Stimuli mit einem primitiven Verfahren gewichten: Ein FALL von 3 Ht ist perzeptiv unauffällig, da er im unmarkierten Fall wegen des Einflusses der Deklination 'erwartet' wird; wir können ihm also den 'perzeptiven Wert' '0' zuweisen. Ein RISE von 3 Ht ist ebenfalls perzeptiv unauffällig - nicht in dem Sinn, daß er nicht wahrgenommen wird, aber in dem Sinn, daß er zu gering ist, um eindeutig einen

phonetisch relevanten RISE zu kennzeichnen; vgl. die Kategoriengrenze bei den weiter oben besprochenen Experimenten und die Verteilung entlang DIM1 in Fig.14. Wir können also auch diesem RISE von 3 Ht einen perzeptiven Wert von '0' zuweisen. Wir setzen nun bei den anderen Ht-Werten den 'objektiven' Wert mit dem 'perzeptiven' gleich und ermitteln die Differenz zwischen dem FALL und dem RISE pro Stimulus; dieser Wert ist in Fig.15 bei jedem Stimulus über dem Strich angegeben. Wir erhalten dabei die folgende Aufteilung:

0 +3 'ausgeglichene' Kontur		
-9 -6 -3 entweder FALL oder RISE		+6 +9 entweder FALL oder RISE
unidirektional	bidirektional bzw. nondirektional	unidirektional

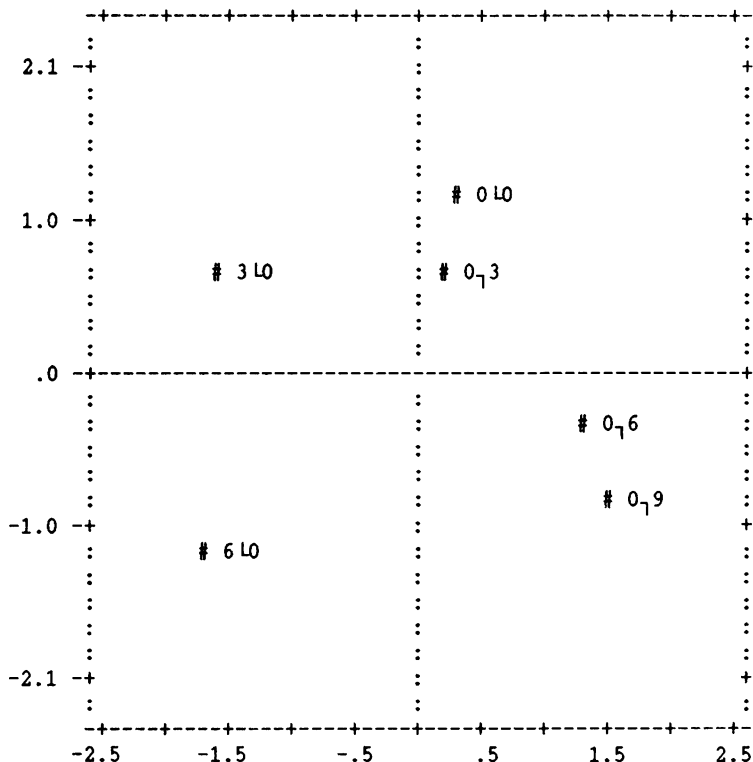
Man kann also DIM4 als die Dimension der Direktionalität bezeichnen: unidirektional vs. bi- oder nondirektional. Wir müssen es dahingestellt lassen, ob die Asymmetrie, die sich bei Stimulus 6|9 mit dem Wert '-3' zeigt, objektive Gründe hat oder nicht doch auf einen experimentellen Fehler zurückzuführen ist.

3.5.5 TONSPRUNG-Stimuli

DIM1 kann als Position des FALL: initial vs. final bzw. komplementär dazu als Position des LEVEL-Verlaufs: final - final+initial - initial bezeichnet werden, DIM2 als Tonsprung: groß - klein. Man erhält in etwa die Darstellung von Tab.2, wenn man an der Waagrechten spiegelt. Die große Ähnlichkeit von 0L0 und 0₇3, vgl. Fig.2, positioniert die beiden Stimuli nahe beieinander; daraus ergibt sich die Halbkreisform. Die Stimuli mit der gleichen Tonsprungdifferenz, 6L0 und 3L0 bzw. 0₇9 und 0₇6, sind unterschiedlich weit voneinander entfernt; das ist möglicherweise dadurch bedingt, daß ein steiler Abfall (*Slope*) perzeptiv auffälliger ist als ein flacher.

Fig.16: TONSPRUNG--Stimuli:

DIM1 (HORIZONTAL): Position des FALL (initial vs. final)
 DIM2 (VERTIKAL): Tonsprung



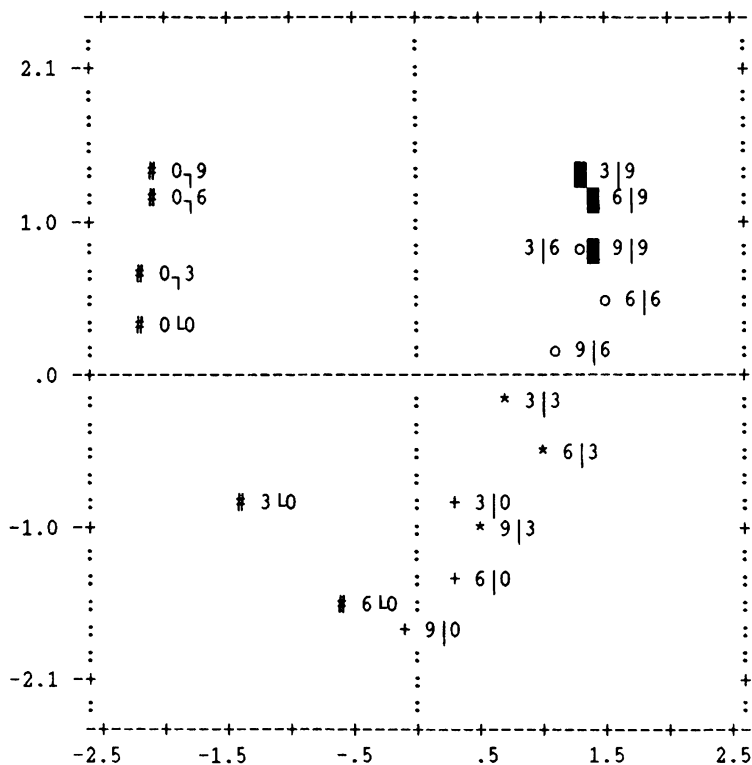
3.5.6 FALL-RISE + TONSPRUNG - Stimuli

In Fig.17 ist DIM1 eine Mischung aus einer Cluster-Bildung mit den TS-Stimuli auf der linken und den FR-Stimuli auf der rechten Seite und einer Abfolge entlang der Dimension RISE innerhalb der FR-Stimuli. Diese Dimension RISE findet sich auch in DIM2 wieder, die man als Ausmaß der finalen Bewegung / des finalen RISE bezeichnen kann: oben ist die finale Bewegung groß, unten gering bzw. ein LEVEL-Verlauf. Insgesamt sieht es so aus, als ob die bei der getrennten Analyse der FR- und der TS-Stimuli wichtigste Dimension (DIM1) eine 'Nahtstelle' hätte bei den ähnlichsten Stimuli 6 L0 und 9 |0, und ansonsten die

jeweilige geometrische Form, also der Kreis von Fig.14 und der Halbkreis von Fig.16, aufgebogen würden und zusammen ein Hufeisen ergeben.⁸

Fig.17: FALL-RISE- und TONSPRUNG-Stimuli:

DIM1 (HORIZONTAL): Cluster (TS- vs. FR-Stimuli)
 DIM2 (VERTIKAL): finaler RISE

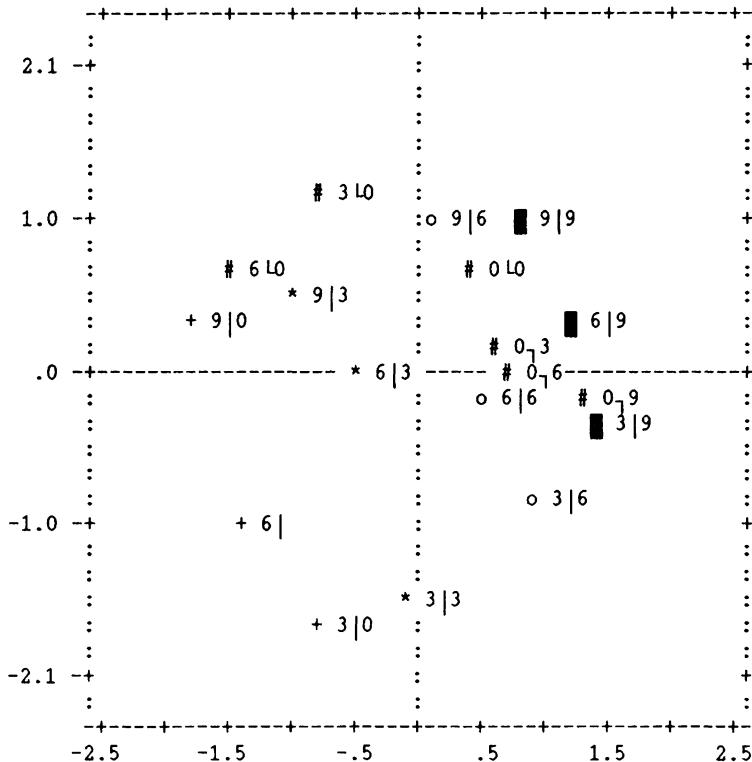


⁸ Eine solche Hufeisenform findet sich oft; sie wird manchmal als Indiz dafür genommen, daß man es eigentlich nur mit einer Dimension zu tun hat, vgl. Shepard (1974:385ff). In unserem Fall dürfte aber die Kombination der beiden unterschiedlichen Stimulusgruppen dafür verantwortlich sein.

In Fig.18 entspricht DIM2 versus DIM3 ziemlich genau DIM1 versus DIM2 bei den FR-Stimuli in Fig.14. Das weist darauf hin, daß bei einer Kombination der beiden Stimulusgruppen die erste Dimension einfach die beiden Gruppen trennt, vgl. die o.a. Cluster-Bildung, und die zweite und die dritte von der systematischen und größeren Gruppe der FR-Stimuli geprägt ist. Es bleibt zu fragen, wo in DIM2 und DIM3 die TS-Stimuli angesiedelt sind. Wenn wir die nächsten Nachbarn betrachten, so liegen 6|0 bei 9|3, 0|0 bei 9|6, 0₇3 und 0₇6 bei 6|6 und 0₇9 bei 3|9. Es geht also von einer starken Ausprägung der initialen Bewegung (FALL) zu einer starken Ausprägung der finalen Bewegung (RISE bei 3|9 oder FALL bei 0₇9), m.a.W. zu extremen Endpunkten; vgl. DIM3 in Fig.15.

Fig.18: FALL-RISE- und TONSPRUNG-Stimuli

DIM2 (HORIZONTAL): finaler RISE (FR-Stimuli)
 DIM3 (VERTIKAL): initialer FALL (FR-Stimuli)



3.5.7 Dimensionen versus linguistische Kategorien

Welcher Zusammenhang besteht nun zwischen den Dimensionen und den Bewertungsfunktionen der anderen Experimente? Sinnvoll läßt sich diese Frage nur für die FR-Stimuli stellen, da die TS-Stimuli zu einheitlich bewertet werden: Sie sind grundsätzlich bessere Exklamativ- und schlechtere Fragesätze. Ein solcher Zusammenhang kann dadurch hergestellt werden, daß man die Dimensionen als unabhängige (Prädiktor-) und die anderen Bewertungsfunktionen als abhängige (Kriteriums-) Variablen einer Regressionsanalyse ansetzt und versucht, die 'Erklärung' der abhängigen durch die unabhängigen Variablen zu optimieren. Maßzahl dafür ist die 'erklärte Varianz' R^2 . Es ist natürlich abzusehen, daß als wichtigste Dimension (DIM1) der finale RISE für die Frage/Nicht-Frage-Distinktion entscheidend ist. Interessant ist also der Beitrag der weniger wichtigen Dimensionen DIM2, DIM3 und DIM4. Da wir allerdings nur 12 Datenpunkte, nämlich die jeweiligen Werte der 12 Stimuli, vorhersagen können, besteht bei einer Erhöhung der Zahl der unabhängigen Variablen bald die Gefahr, daß keine systematischen Zusammenhänge entdeckt werden, sondern daß sozusagen der experimentelle Fehler 'gelernt' wird. (Wenn die Zahl der unabhängigen gegen die Zahl der abhängigen Variablen geht, wird die Vorhersage perfekt, aber sinnlos.) Die folgende Auswertung muß also immer unter dem Caveat betrachtet werden, daß es sich um eine explorative Auswertung handelt.

In Tab.7 sind die Ergebnisse der Regressionsanalysen für alle Experimente aufgeführt. In der ersten Spalte stehen das betreffende Experiment und ein Verweis auf die Abbildung, in der zweiten Spalte die Dimensionen, in der dritten der R^2 -Wert als Maß der erklärten Varianz bei einer multivariaten Regressionsanalyse mit schrittweiser Hinzufügung der unabhängigen Variablen (zuerst die wichtigste, dann die zweitwichtigste, usw.) Ein Strich zeigt an, daß diese Variable das Toleranzkriterium von 0.01 nicht passiert hat. In der vierten steht der Schritt, bei dem die Dimension in die Analyse eingefügt wurde; in der fünften ist angegeben, ob die Korrelation positiv oder negativ ist. (Man beachte, daß 'positiv' oder 'negativ' hier arbiträre Begriffe sind: 'positiv' heißt, daß ein positiver Wert auf der entsprechenden Dimension der NMDS mit einem hohen Wert, also einer schlechten Bewertung bei der Kontextverträglichkeit, korreliert!) In der letzten Spalte steht R^2 für eine univariate Regressionsanalyse, bei der jeweils nur die eine Dimension einging. (Die Wurzel aus diesem Wert ergibt den

Korrelationskoeffizienten r). '1-4' in Spalte 2 und die entsprechenden R^2 -Werte in Spalte 3 stehen für Analysen, bei denen alle vier Dimensionen in die Analyse eingingen.

Wie erwartet, ist DIM1 (finaler RISE) die wichtigste Dimension, die allein schon bei den Fragen einen sehr guten Wert ergibt. DIM2 (initialer FALL) und DIM3 (extremer Endpunkt) tragen bei den Fragen auch noch bei, DIM4 (Direktionalität) ist bei den Fragen immer irrelevant. Das Bild bei den Exklamativen ist etwas widersprüchlich, es fällt aber auf, daß DIM3 nichts beiträgt, und daß beim steigenden Verlauf des Fragekontextes DIM2 und DIM4 relativ viel beitragen.

Dieses Ergebnis stimmt gut mit anderen Ergebnissen aus der Literatur, z.B. mit denen von Studdert-Kennedy/Hadding (1973), überein: Diese Autoren manipulierten bei ihren Stimuli sowohl die finale Bewegung als auch den Fo-Verlauf vor der finalen Bewegung. Sie stellten fest, daß ein steigender finaler Fo-Verlauf in über 90% der Fälle eine Frage (*question*) indizierte und vice versa ein fallender Fo-Verlauf in über 90% der Fälle eine Nicht-Frage (*statement*). Sie stellten aber ebenfalls fest, daß z.B. ein präfinaler hoher Fo-Gipfel eine Frageinterpretation begünstigen kann, m.a.W. daß nicht nur der finale Fo-Verlauf für die Frage/Nicht-Frage-Distinktion relevant ist.

Wir wollen nun im nächsten Teil die bisherigen Ergebnisse an Produktionsdaten überprüfen und danach auf die Relevanz der Dimensionen für die beiden Satzmodi zurückkommen.

Tab.7: Regressionsanalysen mit NMDS-Dimensionen als Prädiktoren und Bewertungsfunktionen als Kriterium

Experiment	DIM	R ² multiv.	Schritt	Korr.	R ² univ.
IT:	1	.86	1	+	.86
Frage-Nicht-	2	-	-	+	.00
frage	3	.94	2	-	.13
(Fig.3)	4	-	-	-	.03
	1-4	.99			
AT	1	.91	1	+	.91
	2	.97	3	+	.03
	3	.94	2	-	.08
(Fig.4)	4	-	-	+	.03
	1-4	.97			
KT1	1	.90	1	+	.90
Fragekontext	2	.98	2	+	.08
(RISE)	3	.99	3	-	.06
(Fig.5)	4	-	-	-	.00
	1-4	.99			
KT1	1	.51	1	-	.51
Exklamativkont.	2	.75	2	+	.25
(Fragekont. RISE)	3	-	-	-	.05
(Fig.6)	4	.89	3	+	.18
	1-4	.93			
KT2	1	.82	1	+	.82
Fragekontext	2	.96	3	+	.07
(FALL)	3	.91	2	-	.15
(Fig.7)	4	-	-	-	.01
	1-4	.97			
KT2	1	.81	1	-	.81
Exklamativkont.	2	-	-	+	.03
(Fragekont. FALL)	3	-	-	-	.08
(Fig.8)	4	-	-	+	.09
	1-4	.93			

4. Produktionsdaten

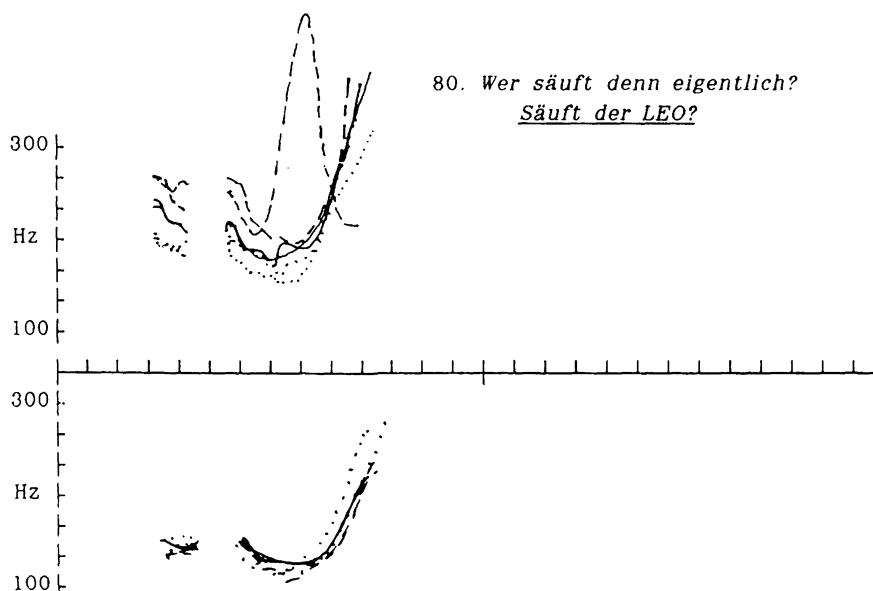
4.1 Material

Es sollen nun den bisherigen Ergebnissen aus den Perzeptionsexperimenten Produktionsdaten mit den gleichen Satzmodi und gleichen oder vergleichbaren Kontext- und Testsätzen gegenübergestellt werden. Es handelt sich dabei um die Sätze Nr.80 bis 84 aus dem vierten Korpus, das ausführlich in Batliner/Oppenrieder (1989) dokumentiert ist; es wird deshalb an dieser Stelle nur kurz charakterisiert: Sechs Sprecher (drei weibliche, drei männliche) produzierten die Testsätze mit dem beigefügten Kontext zwei- bis viermal im schallarmen Raum des Instituts für Phonetik. Die Kontexte indizierten die intendierte Modus- und Fokus-Konstellation. In Fig. 19 sind die Verläufe für die zwei Fragesätze, in Fig. 20 die Verläufe für die drei Exklamativsätze mit dem uns in diesem Zusammenhang interessierenden Testsatz *Säuft der Leo* aufgeführt. Jede Äußerung wurde einem 'Natürlichkeitstest' und einem 'Kategorisierungstest' unterzogen. Im ersten Test wurde die Äußerung auf einer fünfstufigen Ratingskala (1-5) danach beurteilt, wie gut sie in den Kontext paßt. Bei einem Wert von < 2.5 wird die Äußerung als 'natürlich' bezeichnet. Im zweiten Test wurde die Äußerung ohne Kontext dargeboten und die VPn ordneten sie einer Satztypkategorie zu. Bei einer mindestens 80%igen Übereinstimmung und richtiger Zuordnung wird die Äußerung als 'richtig kategorisiert' bewertet. 'Natürliche' und zugleich 'richtig kategorisierte' Äußerungen wurden als Prototypen bezeichnet, die anderen als Nicht-Prototypen; zur genaueren Darstellung und Motivation dieses Vorgehens vgl. im einzelnen Oppenrieder (1988a,b) und Batliner (1988a). Die Anordnung in Fig.19 und 20 ist wie folgt: Die y-Achse zeigt den Fo-Verlauf in Hz, die x-Achse die Zeit. Über der x-Achse stehen die Konturen der Sprecherinnen, unter der x-Achse die der Sprecher. Links stehen die Prototypen, rechts die Nicht-Prototypen. In den Testsätzen ist das Element, auf dem die Hauptakzent-silbe erwartet wurde, groß geschrieben.

Da die Zahl der Fälle pro Konstellation vergleichsweise gering ist, begnügen wir uns mit einer interpretativen Analyse und sehen davon ab, die Ergebnisse einer (Prüf-)Statistik zu unterziehen.

Fig.19: Fragesätze

80. Wer säuft denn eigentlich?
Säuft der LEO?



81. Stimmt das mit dem Leo?
SAUFT der Leo?

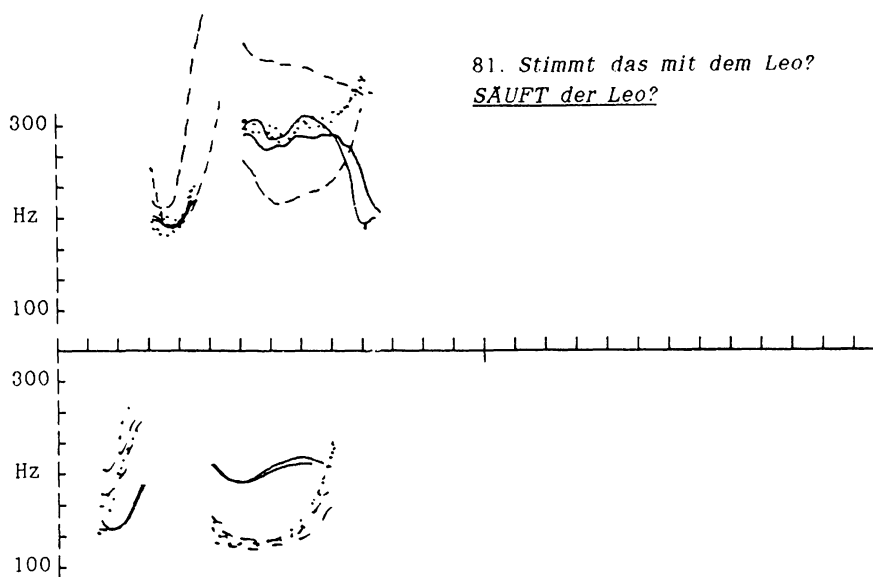
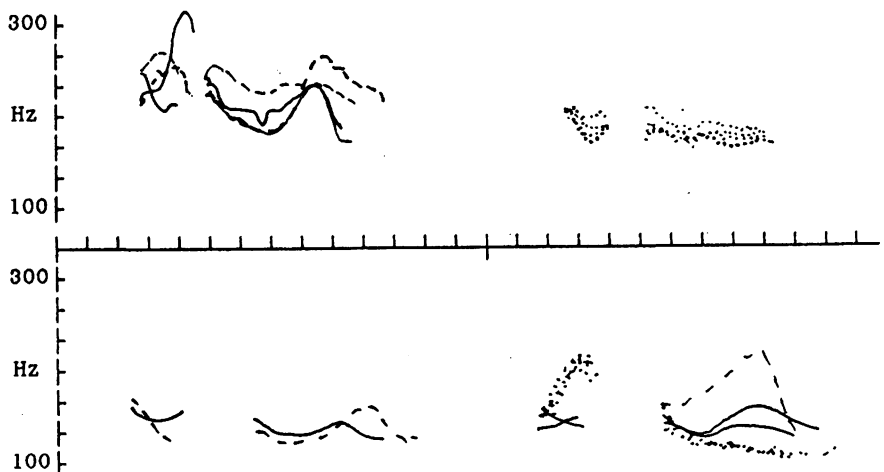


Fig.20: Exklamativsätze

82. Gestern waren wir im Wirtshaus. Mann - Säuft der LEO!



83. Gestern war ich mit dem Leo im Wirtshaus. Mann - SAUFT der Leo!

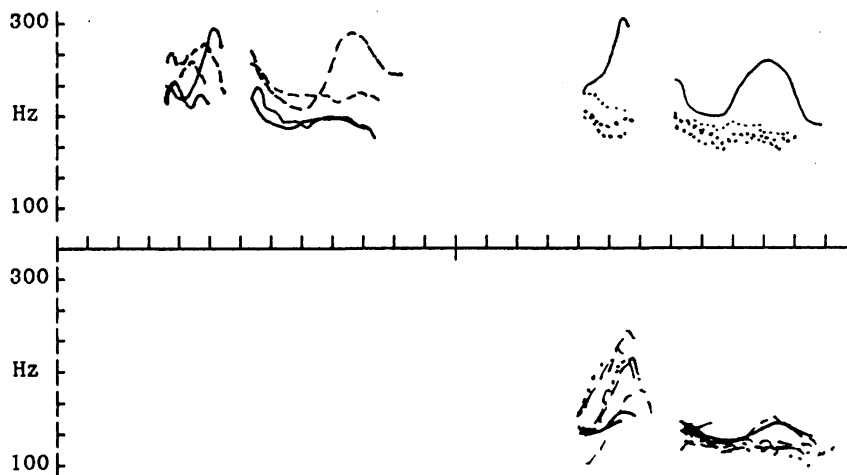
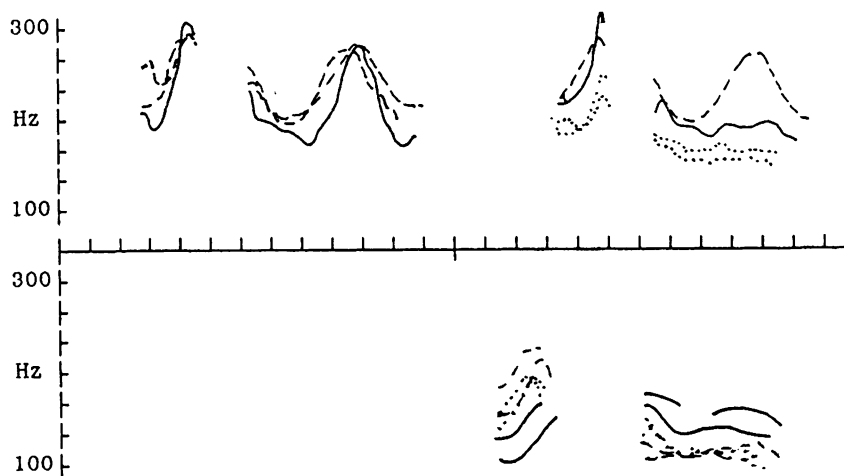


Fig.20:Exklamativsätze (Fortsetzung)

84. *Also sowas hätt' ich nicht gedacht!* SAUFT der Leo!



4.2 Fragesätze

Bei Satz 80 mit fokussiertem *Leo* ist der Verlauf mit einer Ausnahme immer ein schwacher FALL bzw. ein LEVEL auf *säuft* und ein starker RISE auf *Leo*. Die Äußerungen sind alle natürlich und eindeutig kategorisiert. Der Verlauf entspricht unserem Stimulus 3|9, der auch in den Perzeptionstests die besten Fragezuweisungen erhielt.

Einen Verlauf, der den Verläufen von Satz 81 mit Fokus auf *säuft* entspricht, fand sich nicht in unserem Stimulusmaterial: Man kann aber sehen, daß bei einem nicht-finalen fokussierten Element der Offset auch tief sein kann, ohne daß sich an der Natürlichkeit oder Kategorisierbarkeit etwas ändert.

4.3 Exklamativsätze

Bei allen drei Konstellationen fällt auf, daß sich kein Verlauf findet, der dem Stimulus 9|0 vergleichbar ist. Es hat sich aber in Batliner (1988b:226ff) gezeigt, daß ein solcher Verlauf beim Exklamativsatz durchaus akzeptabel ist. Möglicherweise handelt es sich dabei um einen weniger typischen bzw. weniger oft auftretenden Verlauf.

Bei den Exklamativsätzen zeigt sich oft eine Doppelakzentuierung (vgl. dazu Batliner 1988b,c). Bei den Prototypen von Nr.84 ist neben einem RISE auf *säuft* immer ein RISE-FALL auf *Leo* realisiert. Wenn er fehlt, so werden die Äußerungen als Fragen fehlklassifiziert (die genauen Kennwerte dafür finden sich in Batliner/Oppenrieder 1989). Ähnliche Fehlklassifikationen bei entsprechenden Verläufen sind auch bei Nr.82 und Nr.83 zu beobachten. Die 'nicht-prototypischen', fast ebenen Verläufe sind zwar als natürlich bewertet, konnten aber ohne Kontext nicht ausreichend gut klassifiziert werden. Die Verläufe entsprechen dem Stimulus 3|0. Sie sind offensichtlich beim Exklamativsatz dann akzeptabel, wenn ihnen ein exklamativindizierendes *Mann-* vorausgeht, zu dem sie quasi einen intonatorischen Nachlauf bilden können. (Zur Frage, ob es sich hier überhaupt noch um Exklamativsätze handelt, vgl. Batliner 1988c:254f). Da eine Entsprechung zu *Mann-* in unseren Perzeptionstests fehlt, schneidet der Stimulus 3|0 schlechter ab als die Stimuli 6|0 und 9|0.

4.4 Die Relevanz der NMDS-Dimensionen für die Satzmodi 'Verb-Erst-Frage'- und 'Verb-Erst-Exklamativsatz'.

Fragesätze haben typischerweise einen hohen oder einen tiefen Offset; bei einem tiefen Offset ist ein globaler FALL-Verlauf möglich (vgl. Oppenrieder 1988b, Batliner et al. 1989). Darin spiegelt sich die Relevanz von DIM1 (RISE), DIM2 (FALL) und DIM3 (extremer Endpunkt). DIM4 (Direktionalität) bestimmt Fragen vermutlich nicht, da Non- oder Bidirektionalität bei Fragen nicht auftritt oder dann, wenn sie auftritt, die wichtigeren Dimensionen DIM1 bis DIM3 'zuständig' sind.

Die Relevanz von DIM1 für die Exklamativsätze ist bei unseren Perzeptionsexperimenten sicher durch die Art der Stimulusmanipulation bedingt: Bei unseren Stimuli dominieren typische Frageverläufe bzw. -merkmale.⁹ DIM1 definiert also die Exklamativsätze 'negativ': wenn sie stark ausgeprägt ist, ist eine Exklamativinterpretation nicht möglich, wenn sie schwach ausgeprägt ist, dann ist eine solche Interpretation möglich, wenn andere Merkmale vorhanden sind. Diese anderen Merkmale sind mit bestimmt durch DIM2 (FALL) und DIM4 (Direktionalität); wir haben bei den Produktionsdaten gesehen, daß dort bei den Exklamativsatzrealisationen sowohl non-direktionale als auch bidirektionale (allerdings RISE-FALL, nicht FALL-RISE) Verläufe gehäuft auftreten. Exklamative

⁹ Wie sich in anderen Untersuchungen gezeigt hat, vgl. Batliner (1988b,c), ist für den Exklamativ ein hoher Gipfel bei einer RISE-FALL-Bewegung, eine Dehnung der Hauptakzentsilbe, eventuell eine späte Fo-Gipfelposition sowie eine ausgeprägte Doppelakzentuierung typisch - alles Merkmale, die bei unseren Stimuli nicht vorkamen.

zeichnen sich also durch 'sehr viel Bewegung' aus, oder beim Vorliegen gewisser Randbedingungen ('intonatorischer Nachlauf') durch sehr geringe Bewegung. DIM3 (extremer Endpunkt) ist irrelevant: Ein ausgeprägter finaler RISE - besonders als steiler linearer Anstieg wie bei unseren Stimuli - kommt bei Exklamativsätzen praktisch nicht vor. Ein tiefer Offset findet sich, aber der ist durch die Dimension (globaler) FALL abgedeckt. Typisch für den Exklamativ sind aber auch 'mittelhohe' Offsetwerte.¹⁰

5. Allgemeine Diskussion

5.1 Abstufung der Fragehaltigkeit

Es ist eine Eigenheit der Satzmodi, daß intonatorische Indikatoren beim gleichen Satzmodustyp strikt entgegengesetzte Ausprägungen annehmen können. So kann der Verb-Erst-Fragesatz sowohl einen hohen als auch einen tiefen Offset aufweisen. Normalerweise spricht man hier von freien Varianten oder bestenfalls von paralinguistisch relevanten Nuancen - daß etwa ein steigender Intonationsverlauf auf Subordination oder Höflichkeit des Sprechers hinweist. In dieser Deutung hat also der unterschiedliche Intonationsverlauf keinen sprachlichen Zeichencharakter, sondern ist ein **Anzeichen** außersprachlicher Gegebenheiten. Im Gegensatz dazu nehmen wir an, daß die Variation der Intonationsverläufe innerhalb des Satzmodustyp 'Frage' mit einer Abstufung der 'Fragehaltigkeit' (und damit der Kategorialität) korreliert, also sehr wohl Zeichencharakter hat.¹¹ Für diese Annahme lassen sich phonetische und linguistische Gründe ins Feld führen, wobei eine Abstufung auf der phonetischen Ebene die Voraussetzung bildet für eine Abstufung auf der Zeichenebene.

¹⁰ Vgl. dazu die Liste der intonatorischen Prototypen in Oppenrieder (1988b:203ff): die Fo-Kontur wird dort global mit drei relationalen Stufen gekennzeichnet. Bei den Exklamativsätzen überwiegt der Verlauf '132', bei dem der Fo-Offset **höher** liegt als der Fo-Onset, bei den damit kontrastierenden Aussage- bzw. Imperativsätzen überwiegt der Verlauf '231', bei dem der Fo-Offset **tief**er liegt als der Fo-Onset.

¹¹ Die Annahme einer Abstufung der Kategorialität o.ä. ist nicht unbedingt neu; während sie aber in der traditionellen Intonationsforschung auf Intuitionen beruht, versuchen wir, eine empirisch-experimentelle Basis zu schaffen.

5.1.1 Die phonetische Ebene

Wenn man annimmt, daß Phoneme kategorial wahrgenommen werden, so läßt sich daraus nicht einfach folgern, daß die Wahrnehmung der Intonation ebenso kategorial sein muß. Die Wahrnehmung der Intonation kann kontinuierlicher sein oder zumindest mehr Zwischenstufen aufweisen als die von Segmenten: Bei den Plosivkontinua, die im Rahmen des Paradigmas der Kategorialen Wahrnehmung getestet wurden, gibt es zwischen den (Eck-)Stimuli mit einem klaren artikulatorischen Korrelat Zwischenstufen, die von natürlichen Sprechern so nicht produzierbar sind. Bei Vokalkontinua sind dagegen die Zwischenstufen produzierbar; das ist mit ein Grund dafür, daß die Wahrnehmung von Vokalen weniger kategorial ist als die von Plosiven. Solche Zwischenstufen ergeben aber distinktiv keinen Sinn, da sie immer nur einer der bedeutungsunterscheidenden Kategorien zugeschlagen werden können. Bei der Intonation liegt der Fall weniger eindeutig: Natürlich kann, wenn überhaupt, nur eine endliche und kleine Menge von Distinktionen indiziert werden, es ist aber noch nicht trivial gegeben, wieviele es wirklich sind.

5.1.2 Die linguistische Ebene

Der kommunikative Sinn, der sich aus der Äußerung eines Fragesatzes in einer bestimmten Situation ergibt, hängt teilweise davon ab, wo die Äußerung auf bestimmten **fragetypischen 'pragmatischen Bedeutungsdimensionen'** lokalisiert ist, z.B. auf einer Skala zunehmender Antwortdetermination oder auf einer Skala des angezeigten Interesses an einer bestimmten Antwort. Je nach der Position einer Fragesatzäußerung auf einer dieser Skalen kann die Frage als mehr oder weniger typisch, als mehr oder weniger 'fragehaltig' aufgefaßt werden. Unsere generelle These ist, daß die intonatorische Markierung des betreffenden Fragesatzes eines der Merkmale sein kann, das diese 'Fragehaltigkeit' bestimmt - ein anderes Merkmal ist die Verwendung bestimmter Modalpartikeln, die für sich auch schon zeigen, daß die genannten Funktionen grammatikalisiert sein können. Voraussetzung für die intonatorische Markierung derartiger 'pragmatischer' Bedeutungen ist natürlich, daß der betreffende intonatorische Parameter frei variieren kann und nicht bereits als festes Merkmal für die grammatische Satzmoduskennzeichnung benützt wird. Muß z.B. ein Fragesatz durch einen hohen Fo-Offset gekennzeichnet werden, so kann die Offset-Höhe (in diesem Fall als binäres Merkmal - hoch versus tief - definiert) keine zusätzlichen pragmatischen Bedeutungen mehr vermitteln.

Fragetypisch sind mindestens die zwei schon kurz angesprochenen und nun etwas näher zu charakterisierenden pragmatischen Bedeutungsdimensionen:

- (1) Die erste dieser Dimensionen betrifft die **Antworterwartung** des Sprechers. Am einen Ende stehen hier völlig 'offene' Fragen ohne jede vom Sprecher angezeigte Antworterwartung, am anderen Ende rhetorische Fragen mit feststehender Antwort, dazwischen Fragen mit angezeigter Antworterwartung, bei denen der Sprecher aber im Prinzip auch noch mit einer nicht erwartungskonformen Antwort rechnet. So wird durch die Verwendung der unbetonten Modalpartikel *nicht* in Entscheidungsfrasesätzen angezeigt, daß eine bejahende Antwort erwartet wird: *Sind Sie nicht Herr Chomsky?* Bekanntlich beeinflußt auch die Verwendung der Antwortpartikel *doch* die Art der Antworterwartung: Mit *doch* und nicht mit *ja* reagiert man 'negativ' auf eine Frage, auf die ein *nein* als Antwort erwartet wird, angezeigt durch einen betonten Negationsausdruck: A: *Haben Sie KEINE Angst?* B: **Ja./Doch.*
Intonatorisch kann die Offsethöhe einen Hinweis auf die 'Offenheit' der Frage geben; je 'offener' die Frage, desto häufiger wird – ceteris paribus – ein hoher Offset realisiert. Ist die Offsethöhe bereits grammatisch festgelegt, so werden andere intonatorische Merkmale, z.B. ein erhöhter Fo-Range, für die Kennzeichnung des Grades der Antworterwartung genutzt. Als Effekt ergibt sich dann, daß die entsprechende Fragesatzäußerung einen umso 'assertiveren' und weniger fragetypischen Sinn erhält, je mehr die möglichen Antworten eingeschränkt werden.
- (2) Teilweise fällt diese Dimension zusammen mit einer zweiten, die man '**angezeigtes Interesse an einer Antwort**' nennen kann. Bei rhetorischen Fragen ist das Interesse des Sprechers an einer Antwort gleich Null. Am anderen Ende der Skala stehen Äußerungen, mit denen der Sprecher ein großes Interesse an einer Antwort zu verstehen gibt. Das kann z.B. durch Verwendung der Modalpartikel *eigentlich* angezeigt werden: *Kennst du eigentlich Herrn Chomsky?* (Wenn schon durch die Situation klar ist, daß der Angesprochene zu einer Antwort verpflichtet ist, kann die daraus folgende Antwortobligation als besonderes Antwortinteresse kaschiert werden.) Unsere These ist, daß auch hier vor allem der hohe Offset ein besonderes Interesse an einer Antwort anzeigen kann. Bei festgelegter Offsethöhe wird vermutlich ein größerer Range realisiert. Eine Fragesatzäußerung ist umso typischer, je mehr der Sprecher an einer Antwort interessiert ist. Damit wird impliziert, daß (Proto-) Typizität funktional eingesetzt werden kann.

5.2 Merkmale, Dimensionen, Kategorien

Das im wesentlichen innerhalb der (strukuralistischen) Phonologie entwickelte Konzept des **Merkmals** beruht auf der Unterscheidung von distinktiven (d.h. bedeutungsunterscheidenden) und redundanten Merkmalen. Normalerweise wird dabei das Merkmal als binär aufgefaßt, d.h. es gibt nur zwei Ausprägungen, wie etwa [+ stimmhaft] oder [- stimmhaft] bei der Unterscheidung von [b] und [p]. Die VOT (*voice onset time*, das Einsetzen des Stimmtons relativ zur Verschlußlösung bei Plosiven) ist dabei als akustisch/phonetischer Parameter kontinuierlich.¹² An diesem Parameter wird ein Grenzwert festgestellt, der die Kategorien – in unserem Fall stimmhafte und stimmlose Plosive – trennt. Im Paradigma der Kategorialen Wahrnehmung werden dafür ITs und DTs eingesetzt (vgl. oben Teil 3.1 und 3.4). Daß es klar trennbare Kategorien gibt, wurde nicht nur in der Phonetik und Phonologie sehr lange implizit als Faktum angenommen, vgl. Lakoff (1987:6):

From the time of Aristotle to the later work of Wittgenstein, categories were thought be [sic] well understood and unproblematic. They were assumed to be abstract containers, with things either inside or outside the category. Things were assumed to be in the same category if and only if they had certain properties in common. And the properties they had in common were taken as defining the category.

Die 'certain properties' sind bei den Plosiven also die binären Zustände des distinktiven Merkmals der Stimmhaftigkeit.

Wir können an dieser Stelle nicht allgemein auf die Kritik am traditionellen Konzept der Kategorialität eingehen; dafür sei auf Harnad (1987) und Lakoff (1987) sowie die dort zitierte Literatur verwiesen. Wenn wir uns gleich der Kategorie 'Frage' (vs. 'Nicht-Frage') zuwenden, so macht allein schon die Tatsache, daß Fragen hohen oder tiefen Fo-Offset aufweisen können, daß aber dieses Merkmal (das man natürlich auch 'Ausmaß des finalen RISE' nennen kann, vgl. oben 1.2) der einzige Kandidat für ein **distinktives** Merkmal wäre, die traditionelle Bestimmung der Kategorien, wie sie in dem Lakoff-Zitat skizziert ist, hinfällig.¹³ Es ist also adäquater, von der Höhe des Fo-Offsets bzw. dem Ausmaß

¹² In diesem Zusammenhang sehen wir von der Frage ab, inwiefern **akustisch** kontinuierlich veränderbare Parameter einhergehen mit kontinuierlich veränderbaren **artikulatorischen** Merkmalen; so sind z.B. Vokalkontinua um einiges leichter zu produzieren als Konsonantenkontinua. – Grundsätzlich beschränken wir uns in der Diskussion auf die akustischen Merkmale und ihre perzeptiven Korrelate; auf die Frage, inwiefern eine direkte Abbildung der artikulatorischen Prozesse bzw. der Produktionsprozesse i.a. auf die Perzeption adäquater ist, können wir hier nicht weiter eingehen.

¹³ Man könnte nun einwenden, daß die **grammatische** Kategorie 'Frage' – wobei in diesem Zusammenhang irrelevant ist, ob es sich um eine Form- oder Funktionskategorie handelt – etwas prinzipiell anderes sei als die **phoneti-**

des finalen RISE als (sehr relevantem) Hauptmerkmal zu sprechen, wobei andere (Neben-) Merkmale wie die Höhe des initialen FALL bei den FR-Stimuli auch zur Bestimmung der Kategorie beitragen können; die Merkmale stehen zueinander in einer *trading relation*, d.h. daß die geringere Ausprägung des einen Merkmals durch eine größere Ausprägung eines anderen Merkmals kompensiert werden kann; vgl. Repp (1981).

Bisher haben wir so getan, als ob eine direkte Abbildung der akustischen Merkmale auf die **Kategorien** möglich wäre. Dazwischen liegen aber perzeptive und kognitive Prozesse. Die **Dimensionen**, die sich bei den NMDS-Lösungen ergaben, sind fürs erste theoretische Konstrukte der Urteilsfindung anhand solcher perzeptiver und kognitiver Prozesse. Wenn wir ihnen nun ein 'Eigenleben' zugestehen, so sind für den relativ einfachen Fall unserer Teststimuli mit ihren klar definierbaren äquidistanten akustischen Unterschieden die Dimensionen das perzeptive und kognitive Korrelat der akustischen Struktur. Aus dem Repertoire der Dimensionen kommt das Inventar der modusindizierenden intonatorischen Merkmale. Die Abbildung der akustischen Merkmale auf die Dimensionen muß weder stetig noch linear in dem Sinn sein, daß eine Veränderung entlang des akustischen Kontinuums von einem Eckwert zum anderen immer mit einer Veränderung entlang einer Dimension korrespondiert; vgl. die Unstetigkeit der Bewertungsfunktionen an der Kategoriengrenze im AT und in den KT's und DIM1 in Fig.14 sowie die nicht-lineare Abbildung der Merkmale auf DIM3 und DIM4 in Fig.15. Eine lineare Abbildung wie beim finalen RISE (DIM1) mag zwar der Regelfall sein; dafür spricht, daß DIM3 bzw. DIM4 als Prädiktoren verhältnismäßig wenig zur 'erklärten Varianz' bei IT und den KT's beitragen, vgl. Tab.7. Es mag aber auch daran liegen, daß man immer nur nach solchen linearen Abbildungen gesucht hat: Wenn man wie bei den meisten Perzeptionsexperimenten jeweils nur ein Merkmal variiert, wird man fast immer eine lineare Abbildung erhalten. Erst bei gleichzeitiger Variation mehrerer Merkmale können nicht-lineare Abbildungen 'entdeckt' werden.

5.3 Kategorialität der Wahrnehmung

Die Experimente waren nicht im Hinblick darauf konzipiert, entscheidende Argumente (für oder) gegen eine kategoriale Wahrnehmung zu finden. Ein solcher Nachweis (im Sinne eines *experimentum crucis*) ist u.E. nur sehr schwer - wenn

sche Kategorie 'Stimmhaftigkeit'. Es mehren sich aber die Anzeichen, daß dem eben nicht so ist, vgl. z.B. die einzelnen Beiträge in Harnad (1987) sowie Batliner (1989).

überhaupt - zu führen. Um es so zu formulieren: Wenn man nach Kategorialität sucht und das experimentelle Design dementsprechend anlegt, wird man sie auch oft finden. Wenn nicht, so wird man genauso oft eine nicht-kategoriale Wahrnehmung entdecken.

Unsere Überzeugung läßt sich wie folgt skizzieren: Natürlich muß der Mensch die erhaltene Information möglichst schnell auf ihren kategorialen Gehalt reduzieren, um nicht im Übermaß an Information zu ertrinken. (vgl. Hörmann 1976:460ff oder Bühlers (1934:42ff) Prinzip der 'abstraktiven Relevanz'). Diese Reduzierung kann sehr schnell vor sich gehen, wie im segmentalen Bereich (ist es ein [b] oder ein [p]?), oder langsamer, wie im suprasegmentalen Bereich (indiziert die Intonation eine Frage oder nicht?). Ebenso notwendig ist aber auch ein gewisses Nachhallen der kontinuierlichen (akustischen) Information - schon um bei eventuellen Interpretationsschwierigkeiten darauf zurückgreifen zu können.¹⁴ Kategorialität ist eher dann gegeben, wenn die phonetische Information 'einfach' und möglichst auch psychophysisch kodierbar ist (im Sinne eines 'entweder-oder', wie z.B. bei der VOT und bei der Transition des zweiten Formanten, die entweder steigend oder fallend sein kann). Sie ist dann eher nicht gegeben, wenn die phonetische Information komplex und psychophysisch nicht kodierbar ist, und wenn es sinnvoll ist, die inhaltliche Information nicht rein kategorial zu gestalten. So ist es wohl möglich, aber nur selten sinnvoll, einen Plosiv zu produzieren, der zu zwei Drittel [b] und zu einem Drittel [p] ist; es kann aber sehr sinnvoll sein, dem Hörer eine Abstufung der Fragehaltigkeit zu signalisieren.

Vor diesem Hintergrund lassen sich die Ergebnisse in Bezug auf die Kategorialität interpretieren: Bei einer direkten Befragung und *forced-choice*-Aufgabe wie beim IT zeigte sich eine **Kategoriengrenze** im Bereich einer psychophysischen Grenze (RISE versus FALL). Bei einer direkten (AT) oder indirekten (KTs) Befragung **ohne forced-choice**-Aufgabe zeigten sich **Abstufungen** innerhalb der Kategorien. Diese Abstufungen waren zweifach bedingt: Zum einen durch unterschiedliche Ausprägungen der intonatorischen Parameter (RISE und in geringerem Ausmaß FALL) **innerhalb des Testsatzes**, zum anderen durch die unterschiedlichen Ausprägungen der intonatorischen Parameter von **Kontextsatz und Testsatz**. Wir nehmen an, daß die Abstufungen zeichenhaft eingesetzt werden. Damit sind sie ein genuines Beschreibungsobjekt sowohl der Linguistik als auch der Phonetik.

¹⁴ Die Begriffe 'Reduzierung' bzw. 'Nachhallen der Information' sind hier informell verwendet; wir möchten uns jeder konkreten Aussage über die genauen Prozesse im Kurz- bzw. Langzeitgedächtnis enthalten. Es kann sich auch um verschiedene Arten des Zugriffs oder der Aufmerksamkeitsteuerung handeln.

LITERATUR

- Altmann, H. (1987): Zur Problematik der Konstitution von Satzmodi als Formtypen. In: Meibauer, J. (Hg.) (1987): Satzmodus zwischen Grammatik und Pragmatik. Tübingen. S.22-56.
- Altmann, H. (Hg.) (1988): Intonationsforschungen. Tübingen.
- Batliner, A. (1988a): Produktion und Prädiktion. Die Rolle intonatorischer und anderer Merkmale bei der Bestimmung des Satzmodus. In: Altmann, H. (Hg.) (1988). S.207-221.
- Batliner, A. (1988b): Modus und Fokus als Dimensionen einer Nonmetrischen Multidimensionalen Skalierung. In: Altmann, H. (Hg.) (1988). S.223-241.
- Batliner, A. (1988c): Der Exklamativ: Mehr als Aussage oder doch nur mehr oder weniger Aussage? Experimente zur Rolle von Höhe und Position des F0-Gipfels. In: Altmann, H. (Hg.) (1988). S.243-271.
- Batliner, A. (1989): Eine Frage ist eine Frage ist keine Frage. Perzeptionsexperimente zum Fragemodus im Deutschen. (In diesem Band)
- Batliner, A./Oppenrieder, W. (1988): Rising intonation: Not passed away but still alive. A reply to R. Geluykens. In: Journal of Pragmatics 12. S.227-233.
- Batliner, A./Oppenrieder, W. (1989): Korpora und Auswertung. (In diesem Band)
- Batliner, A./Nöth, E./Lang, R./Stallwitz, G. (1989): Zur Klassifikation von Fragen und Nicht-Fragen anhand intonatorischer Merkmale. (erscheint in: Tagungsband der 15. Gemeinschaftstagung der Deutschen Arbeitsgemeinschaft für Akustik (DAGA). 13.-16. März 1989)
- Batliner, A./Schiefer, L. (1987): Stimulus category, reaction time, and order effect - An experiment on pitch discrimination. In: Proceedings XIth ICPHS, Vol. 5. S.46-49.
- Borg, I (1981): Anwendungsorientierte Multidimensionale Skalierung. Berlin etc.
- Bühler, K. (1934): Sprachtheorie. Stuttgart.
- Gandour, J.T. (1978): Perceived Dimensions of 13 Tones: A Multidimensional Scaling Investigation. In: Phonetica 35. S.169-179.
- Gandour, J.T./Harshman, R.A. (1978): Crosslanguage Differences in Tone Perception: A Multidimensional Scaling Investigation. In: Language and Speech 21. S.1-33.
- Guttman, L. (1977): What is not What in Statistics. In: The Statistician 26. S.81-107.
- Harnad, S. (Hg.) (1987): Categorical perception. The groundwork of cognition. Cambridge etc.
- Henkel, R.E. (1976): Tests of Significance. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-004. Beverly Hills and London.

- Hörmann, H. (1976): *Meinen und Verstehen*. Frankfurt a.M.
- Klein, W. (1982): Einige Bemerkungen zur Frageintonation. In: *Deutsche Sprache*. S.289-310.
- Kruskal, J.B./Wish, M. (1978): *Multidimensional Scaling*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-011. Beverly Hills and London.
- Lakoff, G. (1987): *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago und London.
- Lindsay, D./Ainsworth, W.A. (1985): Two models of nuclear intonation. In: *Journal of Phonetics* 13. S.163-173.
- Medin, D.L./Barsalou, L.W. (1987): Categorization processes and categorical perception. In: Harnad, S. (Hg.) (1987). S.455-490.
- Oppenrieder, W. (1988a): Intonation und Identifikation. Kategorisierungstests zur kontextfreien Identifikation von Satzmodi. In: Altmann, H. (Hg.) (1988). S.153-167.
- Oppenrieder, W. (1988b): Intonatorische Kennzeichnung von Satzmodi. In: Altmann, H. (Hg.) (1988). S.169-205.
- Repp, B.H. (1981): Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. In: *Haskins Laboratories: Status Report on Speech Research SR-67/68*. S.1-40.
- Repp, B.H. (1984): Categorical perception: Issues, methods, findings. In: Lass, N.J. (Hg.): *Speech and Language*. Orlando etc. S.243-335.
- Repp, B.H. (1986): The role of psychophysics in understanding speech perception. In: Schouten, M.E.H. (Hg.): *The Psychophysics of Speech Perception*. Dordrecht et al. S.3-27.
- Rosch, E./Mervis, C.B. (1975): Family Resemblances: Studies in the Internal Structure of Categories. In: *Cognitive Psychology* 7. S.573-605.
- Rossi, M. (1978): Interactions of Intensity Glides and Frequency Glissandos. In: *Language and Speech* 21. S.384-396.
- Schiefer, L./Batliner, A. (1988): Intonation, Ordnungseffekt und das Paradigma der Kategorialen Wahrnehmung. In: Altmann, H. (Hg.) (1988). S.273-291.
- Shepard, R.N. (1974): Representation of Structure in Similarity Data: Problems and Prospects. In: *Psychometrika* 39. S.373-421.
- Studdert-Kennedy, M./Hadding, K. (1973): Auditory and linguistic processes in the perception of intonation contours. In: *Language and Speech* 16. S.293-313.
- Thorsen, N.G. (1987): Suprasegmental Transcription. In: Almeida, A./Braun, A. (Hgg.): *Probleme der phonetischen Transkription*. Stuttgart. S.79-109.