

**DER EXKLAMATIV: MEHR ALS AUSSAGE ODER DOCH NUR MEHR ODER WENIGER AUSSAGE? EXPERIMENTE ZUR ROLLE VON HÖHE UND POSITION DES F<sub>0</sub>-GIPFELS.**

**Anton Batliner (München)**

**1. EINLEITUNG**

**LINGUISTISCHE KLASSIFIKATION:** Ausgangspunkt der im folgenden beschriebenen Experimente ist das von H. Altmann entwickelte Satzmodussystem, vgl. Altmann (1987); **Thema** ist die intonatorisch unterschiedliche Markierung der Modi **Aussage** und **Exklamativ**, und zwar bei derjenigen Minimalpaarkonstellation, bei der die Intonation allein die beiden Modi unterscheidet, also bei Verb-Zweit-Stellung ohne zusätzliche Merkmale (wie Modalpartikel o.ä.). Im Gegensatz zu den Hauptkategorien wie Aussage, Frage und Imperativ ist dabei der Status des Exklamativs weniger eindeutig bestimmbar; z.T. wird er von anderen Autoren als ein Spezialfall der (emphatischen) Aussage, also nicht als eigenständiger Modus betrachtet; vgl. dazu Näf (1987). Daß der Exklamativ der Aussage näher steht als der Frage und auch von der Aussage nicht so eindeutig zu trennen ist, zeigt sich auch in der in Batliner (1988) beschriebenen Analyse im Rahmen einer Nonmetrischen Multidimensionalen Skalierung. Nimmt man realistischerweise an, daß die Kategorialität der Modi grundsätzlich nur in ihren Kernbereichen deutlich ausgeprägt ist, daß es aber an ihren Rändern Mischtypen gibt (vgl. das Konzept der 'Fuzziness', s. u.a. Oden 1979), so kann man weiter vom Exklamativ annehmen, daß wir es hier mit einer Kategorie zu tun haben werden, die besonders 'fuzzy' ist.

**LINGUISTISCHER VERSUS PHONETISCHER GESICHTSPUNKT:** Von Haus aus interessieren sich Linguisten mehr für die formalen, strukturellen und semantischen Eigenschaften des Exklamativs und behandeln seine phonetischen Eigenschaften eher global; Näf (1987:146) etwa begnügt sich mit der Charakterisierung "portato-ritando-Intonation". Auf der anderen Seite spricht die Phonetik normalerweise nicht von Exklamativ, sondern unspezifiziert von Emphase, und begnügt sich dabei mit einem allgemeinen vortheoretischen Verständnis, ohne den

Begriff abzuleiten. Man kann auch in den allermeisten Fällen Einverständnis darüber erzielen, welche Äußerung in welcher Situation emphatisch gemeint ist. (Ein anderer Grund dürfte sein, daß Phonetiker solche Begriffsbestimmungen lieber den Linguisten überlassen.) Versucht man nun, wie in diesem Beitrag, einer phonetischen Untersuchung ein streng linguistisch definiertes Material zugrundezulegen, so erkauft man sich mit einem grundsätzlichen Vorteil eine praktische Schwierigkeit: vorteilhaft ist, daß die phonetischen Ergebnisse ohne weiteres in das linguistische Modell inkorporierbar sind; die Schnittstelle zwischen Linguistik und Phonetik ist also gegeben. Nachteilig kann sein, daß die linguistischen Vorgaben die Wahl des Stimulusmaterials stark einschränken.

**INTONATORISCHE MARKIERUNG:** Grob gesprochen, zeichnet sich der Exklamativ durch ein 'Mehr' aus: auf der funktionalen, inhaltlichen Seite durch größeres emotionales Engagement, Überraschung, Erstaunen; auf der formalen Seite bei der Intonation durch stärkere Ausprägungen der relevanten Parameter. Kennzeichen sind also im Fo-Bereich größere **Maxima**, im Zeitbereich größere **Dauer**, und eventuell größere **Intensität**. Um es phonologisch auszudrücken, wobei man sich dabei weniger binäre Oppositionen als vielmehr n-dimensionale Merkmale vorstellen sollte: die Aussage ist der Normalfall, also **unmarkiert**; der Exklamativ ist eine **markierte** Kategorie und wird deshalb auch durch markierte Parameterausprägungen indiziert. Markiert können dabei nicht nur die drei erwähnten Standardparameter sein, sondern auch andere, z.B. die Fo-**Gipfelposition**.

**EXPERIMENTELLE OPERATIONALISIERUNG:** Es wurde ein Testsatz gewählt, der bestmöglich den linguistischen und phonetischen Anforderungen entspricht. D.h. es sollte ein Satz sein, mit dem - durch unterschiedliche intonatorische Markierung - möglichst viele verschiedene Modus- und Fokuskonstellationen indiziert werden können. Wie sich zeigen wird, lassen sich die Modi auch durch eine unterschiedliche Behandlung der Fokusstruktur kennzeichnen; Modus und Fokus können also nur bedingt isoliert voneinander betrachtet werden. Phonetisch-experimentelle Kriterien sprechen für einen kurzen Satz mit möglichst vielen stimmhaften Segmenten. Die linguistischen Kriterien für einen Exklamativ erfordern hier einen definiten Ausdruck sowie ein graduierbares Prädikat. Wir entschieden uns für den Satz *Der Leo säuft*, mit dem u.a. die in Tab.1 aufgeführten Konstellationen erzeugbar sind. Wenn der Kontextsatz einen Satzakzent auf *Leo* indiziert, so sprechen wir vom 'frühen', wenn ein Hauptakzent auf *säuft* indiziert wird, vom 'späten' Fokus. In der Tabelle sind auch die Kontextsätze aufgeführt, mit denen im Experiment genau die jeweiligen Konstellationen indiziert werden sollten. Eine Bemerkung zur Terminologie: Im folgenden verwen-

den wir statt einiger expliziter Begriffe oft Kurzformen: KONTRAST etwa steht für 'Aussagesatz mit Kontrastakzent', AUSSAGE für 'normaler Aussagesatz ohne Kontrastakzent' usw. Diese Kurzformen werden auch später beim Versuchsplan für die Faktoren und ihre Klassen verwendet. 'RISE', 'LEVEL' und 'FALL' stehen für 'Anstieg', 'ebener Verlauf' bzw. 'Abfall' der Fo-Kontur, 'Gipfel' für 'Maximum und Wendepunkt der Fo-Kontur'.

**Tab.1: Modus/Fokus-Konstellationen und Kontextsätze**

---

AUSSAGE, FRÜHER FOKUS:	<i>Ihr fragt mich, wer säuft?</i>
AUSSAGE, SPÄTER FOKUS:	<i>Was soll schon mit dem Leo sein?</i>
KONTRAST, FRÜHER FOKUS:	<i>Der Theo säuft nicht,</i>
KONTRAST, SPÄTER FOKUS:	<i>Der Leo trinkt nicht,</i>
EXLAMATIV, FRÜHER FOKUS:	<i>Gestern waren wir im Wirtshaus. Mann -</i>
EXKLAMATIV, SPÄTER FOKUS:	<i>Gestern war ich mit dem Leo im Wirtshaus. Mann -</i>

---

**GRUNDLEGENDE FRAGESTELLUNG:** Wir interessieren uns in diesem Beitrag für die schon im Titel angedeutete Frage, ob sich der Exklamativ immer durch eine 'stärkere' intonatorische Markierung gegenüber der Aussage auszeichnet. Es wird sich zeigen, daß das nicht notwendig der Fall sein muß, daß es aber trotzdem auch in bezug auf die intonatorische Markierung sinnvoll ist, vom Exklamativ als einer eigenen Kategorie zu sprechen.

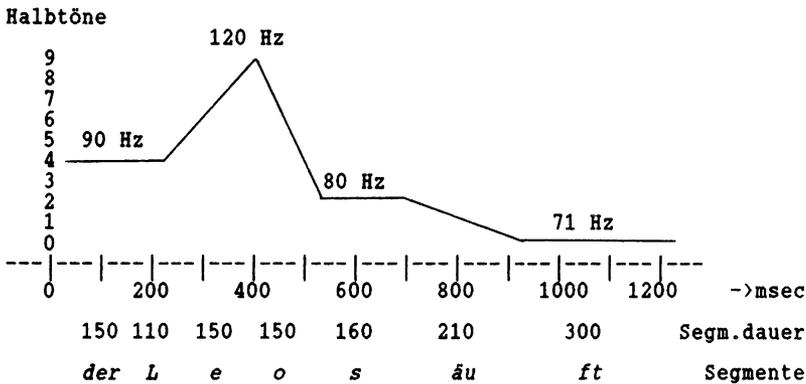
## 2. TESTAUFBAU

Um Redundanzen in der Darstellung zu vermeiden, werden die folgenden Punkte, soweit sie auf alle Experimente zutreffen, gesammelt abgehandelt. Bei den einzelnen Experimenten wird dann nur noch kurz darauf eingegangen.

**MATERIAL:** Der Autor produzierte die Kontextsätze mit natürlicher Intonation so wie den Testsatz *Der Leo säuft* mit monotoner Intonation mehrmals im schallarmen Raum des Instituts für Phonetik in München (Telefunken M15, Geschwindigkeit 19/sec). Die monotone Intonation sollte dabei einen möglichst neutralen Ausgangspunkt für die Manipulation darstellen. Die besten Realisationen wurden auf einer PDP11/50 mit einer Sample-Rate 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 Halbtonbasis. Einer der Stimuli taucht in jedem der Experimente auf. Dieser 'Bezugsstimulus'

wird nun anhand von Fig.1 beschrieben; die anderen Stimuli können leicht von ihm abgeleitet werden. Auf der y-Achse sind in Fig.1 die Halbtonwerte aufgetragen, auf der x-Achse der Zeitverlauf. Der LEVEL-Verlauf des Anfangsteils wurde nicht manipuliert, sondern von der natürlichen Äußerung übernommen. Ebenfalls nicht manipuliert waren die stimmlosen Konsonanten in *säuft*. Im vokalischen Teil von *Leo* erfolgt ein RISE um fünf Halbtöne und, nach einem Gipfel bei der Transition zwischen [e] und [o], ein FALL von sieben Halbtönen. Der Diphthong in *säuft* fällt um zwei Halbtöne. Der Intensitätsverlauf blieb unmanipuliert, d.h. monoton, aber mit den geringfügigen intrinsischen Schwankungen der natürlichen Produktion.

Fig.1: Bezugsstimulus



In Tab.2 sind die Stimulusserien I-V aufgeführt, die in den Experimenten getestet wurden; sie seien nun kurz erläutert. Wir interessierten uns hauptsächlich für den Gipfel als Indikator des Exklamativs und dabei besonders für die **Gipfelerhöhung** und **Gipfelverschiebung**. Der Intensitätsverlauf wurde deshalb nicht verändert. Die Dauer wurde für die Untersuchung des verschobenen Gipfels (Kontinua II-V) kovariierend manipuliert, da sonst zu steile Fo-Verläufe entstanden wären.

Tab.2: Dauer- und Halbtonwerte der Kontinua

		Dauerrelation (RISE/FALL) pro Kontinuum				
		I	II	III	IV	V
Halbtöne						
1 / 3	150/150	-	-	-	-	-
3 / 5	150/150	0/270	90/180	135/135	180/90	
5 / 7	150/150	0/300	100/200	150/150	200/100	
7 / 9	150/150	0/330	110/220	165/165	220/110	
9 / 11	150/150	0/360	120/240	180/180	240/120	
11 / 13	-	0/390	130/260	195/195	260/130	

Im ersten Kontinuum (I) bleibt die Dauer des manipulierten Bereichs in *Leo* konstant auf 300 msec; der Gipfel befindet sich in der Mitte bei der Transition zwischen [e] und [o]. Vor dem Schrägstrich ist immer die Dauer des RISE, danach die des FALL angegeben. Der Gipfel wird in Zwei-Halbton-Schritten von einem Halbton bis neun Halbtöne über dem Ausgangswert angehoben. Man beachte, daß wir die Höhe des Gipfels immer mit dem Wert des RISE angeben; der FALL ist jeweils um 2 Halbtöne größer. Da dieses Kontinuum noch zu wenig im Exklamativbereich lag, wurde in den Kontinua II-V der Gipfel in Zwei-Halbton-Schritten von 3 Halbtönen auf 11 Halbtöne angehoben. Die Dauer kovariert dabei in 30-msec-Schritten zwischen 270 msec und 390 msec. Die durch die pitchsynchrone Manipulation bedingten Schwankungen um +/-5 msec um die in Tab.2 angegebenen Zielwerte sind perceptiv unauffällig. Im Kontinuum IV war der Gipfel analog zum Kontinuum I in der **Mitte**, bei der Transition zwischen [e] und [o]. Im Kontinuum III kam der Gipfel **früh**, nach dem ersten Drittel im [e]; die Transition blieb in der Mitte. Im Kontinuum V kam der Gipfel **spät**, nach dem zweiten Drittel zusammen mit der Transition zwischen [e] und [o]. Die Ergebnisse der ersten Testphase führten zur Konstruktion eines **sehr frühen** Kontinuums II, bei dem der Gipfel noch im [l] lag, d.h. der RISE im [e] hat eine Dauer von Null. Bei diesem letzten Kontinuum blieb der Ausgangswert nicht konstant auf 90 Hz, sondern erhielt jeweils einen LEVEL-Verlauf, der zwei Halbtöne unter dem Wert des Gipfels lag.

**HYPOTHESEN:** Es wurden keine exakten Alternativhypothesen aufgestellt, sondern eine globale Hypothese: ein 'Mehr' an Ausprägung, an Markiertheit, entspricht einem 'Mehr' an Exklamativität. Das 'Mehr' wird dabei kodiert als Erhöhung des Gipfels bzw. als Verschiebung dieser Gipfelposition aus einer unmarkierten Lage heraus. Eigene, frühere Ergebnisse (vgl. Batliner 1988) lassen weiter vermuten, daß der Fokus beim Exklamativ weniger indiziert wird als bei der Aussage. Weitere Hypothesen, die sich aus eigenen Vorarbeiten ergaben, werden wir jeweils bei den einzelnen Experimenten gesondert erläutern.

**VERSUCHSPERSONEN** waren Studenten der Germanistik oder Phonetik, die für ihre freiwillige Teilnahme bezahlt wurden.

**DESIGN:** Es wurden Identifikationstests, Diskriminationstests und Kontexttests durchgeführt. Die beiden ersten Arten werden üblicherweise bei Untersuchungen zur Kategorialität phonetischer Parameter verwendet. Der Kontexttest wurde wegen seiner größeren Nähe zur natürlichen Kommunikationssituation hinzugenommen. Die Versuchspersonen saßen dabei immer im Sprachlabor des Instituts für Phonetik 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 Versuchspersonen hatten dann vier Sekunden Zeit, ihre Antwort zu überlegen und die entsprechende Taste zu drücken. Die Antworten wurden auf einer PDP11/3 gesammelt und zur weiteren Verarbeitung aufbereitet. Beim **Identifikationstest** lautete die Instruktion: "Bitte entscheiden Sie, ob es sich bei diesem Stimulus um eine Aussage oder einen Ausruf handelt, und drücken Sie die linke Taste für 'Aussage' oder die rechte für 'Ausruf'." Das geläufigere 'Ausruf' wurde für die Instruktion dem Fachterminus 'Exklamativ' vorgezogen. Es wurde darauf hingewiesen, daß man sich auch eine nachdrückliche (kontrastierende) Aussage vorstellen kann und daß beim Ausruf ein Merkmal des Erstaunens, der Überraschung hinzukommen muß. Beim **Diskriminationstest** lautete die Instruktion: "Bitte entscheiden Sie, ob die beiden Stimuli innerhalb eines Paares in allen Belangen gleich klingen oder nicht, und drücken Sie dann die linke Taste für 'gleich' oder die rechte für 'verschieden'." Bei den **Kontexttests** mußten die Versuchspersonen ihre Antworten auf einer fünf-stufigen Ratingskala abgeben. 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 die Darstellung in eine Skala von 0 bis 4 umgerechnet.) Da die Items öfter als einmal vorkamen (vgl. die jeweiligen Angaben zu den einzelnen Experimenten), erhielten wir für die Identifikations- und Diskriminationstests pro Versuchsperson und Item einen Häufigkeitswert für eine der Antwortmöglichkeiten und für die Kontexttests einen Mittelwert aus den Einzelbeurteilungen.

**STATISTISCHE AUSWERTUNG:** Die statistische Auswertung wurde mit dem Programmpaket SPSSPC+ vorgenommen, d.h. bei erfüllten Voraussetzungen, gegebenenfalls nach Transformation der Daten, mit der dort verfügbaren Prozedur MANOVA. Dabei wurde das multivariate Modell zugrundegelegt. Wenn die Voraussetzungen nicht gegeben waren, so wurde zuerst für die Teilgruppen der Stimuli, die die gleiche Gipfelhöhe aufwiesen, ein nicht-parametrischer Test für  $k$  abhängige Stichproben (Kendalls Konkordanzkoeffizient) durchgeführt. War dieser Test signifikant, so wurde mit dem Wilcoxon-Test paarweise verglichen. Zur Vermeidung eines Fehlers der ersten Art wurde dabei immer das Signifikanzniveau nach der Formel

$$1 - (1 - \alpha)^{1/c}$$

abgesenkt, wobei  $\alpha$  für ein Signifikanzniveau von 5% steht und  $c$  für die Anzahl der durchgeführten Tests.

**TERMINOLOGIE:** Wie schon erwähnt, verwenden wir für die untersuchten Konstellationen statt linguistisch exakter Bezeichnungen abkürzende, die zugleich die Faktoren und ihre Klassen im varianzanalytischen Design bezeichnen; sie sind im weiteren Text in diesem Zusammenhang mit Großbuchstaben geschrieben:

MODUS (AUSSAGE, KONTRAST, EXKLAMATIV);  
 FOKUS (FRÜH, SPÄT);  
 POSITION (SEHR FRÜH, FRÜH, MITTE, SPÄT);  
 HALBTON (1, 3, 5, 7, 9, 11).

Die Tests werden ebenfalls von nun an mit einer abkürzenden Schreibweise bezeichnet: 'IT' steht für 'Identifikationstest', 'DT' für 'Diskriminationstest' und 'KT' für 'Kontexttest'.

### 3. DIE EXPERIMENTE

An Exp.1 und Exp.3 sowie an Exp.2 und Exp.4 nahmen je 20 Versuchspersonen teil, die auf zwei Gruppen aufgeteilt wurden; der einen Gruppe wurde zuerst der IT, dann der KT (aufgeteilt auf zwei Sitzungen) dargeboten, die andere durchlief die Tests in umgekehrter Reihenfolge. Die Tests fanden in drei aufeinanderfolgenden Wochen zum gleichen Termin statt.

#### *EXPERIMENT 1*

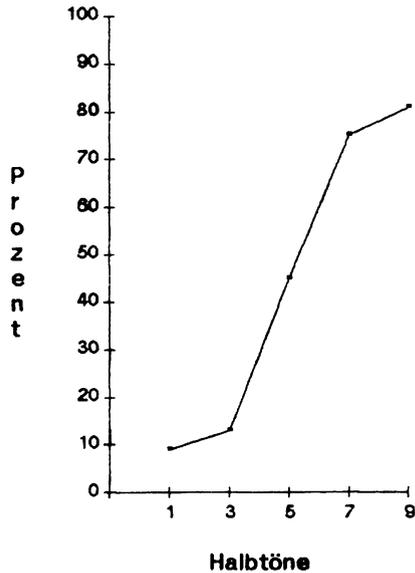
**FRAGESTELLUNG:** Es sollte untersucht werden, ob eine Gipfelerhöhung ohne gleichzeitige Dehnung im IT eine (kategoriale) Grenze zwischen AUSSAGE und EXKLAMATIV ergibt.

**MATERIAL UND DESIGN:** Die fünf Stimuli aus dem Kontinuum I mit einem MITTLEREN Gipfel von einem bis neun Halbtönen ohne kovariierende Dauervariation wurden in einem IT je zehnmal randomisiert dargeboten.

**ERGEBNIS:** Fig.2 zeigt die Zuweisungen zum EXKLAMATIV, wobei auf der y-Achse die Antworten in Prozent, gemittelt über die Versuchspersonen, und auf der x-Achse die Stimuli aufgetragen sind. Die 50%-Grenze liegt im Bereich des Stimulus mit einem Gipfel von 5 Halbtönen; dieser Stimulus bildet deshalb einen 'Schnittstimulus' (der Identifikationskurve mit der 50%-Grenze). Die Eckstimuli liegen bei 9% bzw. 80.5%. Die Identifikation kann also als schwach kategorial bezeichnet werden. Eine ausgeprägte Kategorialität wäre gegeben, wenn die Kurve steiler verlaufen würde und zumindest die Eckstimuli nahe bei 0% bzw. 100 % liegen würden. (Zum Begriff 'Kategorialität' vgl. Repp 1984; wir verwenden hier diesen Begriff allerdings in einer nicht strikt definierten Weise, vgl. die Diskussion weiter unten.) Zwei systematische Erklärungen für den relativ niedrigen Wert des 9-HALBTON-Stimulus bieten sich an: zum einen ist auch eine

emphatische Aussage mit einem hohen Gipfel innerhalb unseres Satzmodussystems möglich, zum anderen mag der 9-HALBTON-Stimulus, da er nicht zusätzlich gedehnt ist, noch nicht im typischen Exklamativ-Bereich liegen.

**Fig.2: Identifikationstest:  
Exklamativzuweisung**

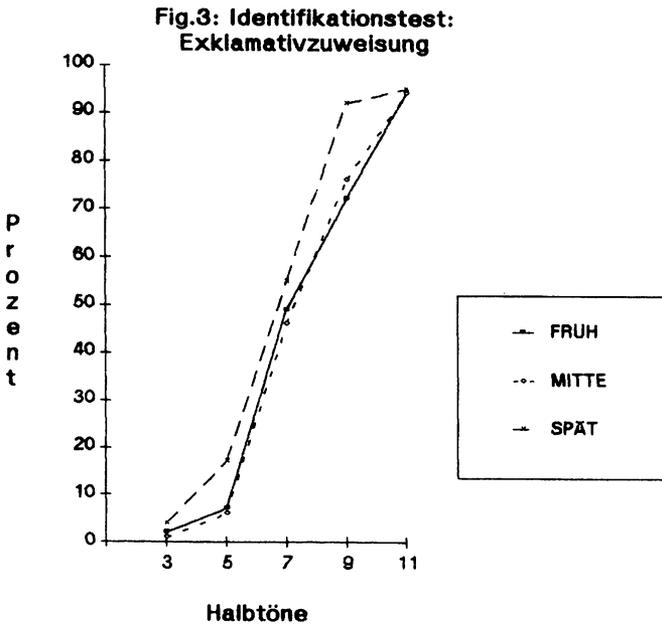


## **EXPERIMENT 2**

**FRAGESTELLUNG:** Die Fragestellung von Exp.1 wurde erweitert und modifiziert: anhand von drei unterschiedlichen Gipfelpositionen soll untersucht werden, welche dieser Positionen eher exklamativtypisch ist. Die Kontinua sollten durch eine Verschiebung im 'Range'-Bereich (3-11 HALBTÖNE) und eine gleichzeitige kovariierende Dehnung den EXKLAMATIV besser indizieren können: wie aus Tab.2 ersichtlich ist, geht eine Erhöhung um zwei HALBTÖNE immer einher mit einer Dehnung um 30 msec.

**MATERIAL UND DESIGN:** Analog zu Exp.1 wurden nun die Stimuli aus den Kontinua III, IV und V mit einem Gipfel von 3 bis 11 HALBTÖNEN und FRÜHER, MITTLERER und SPÄTER Gipfelposition in einem IT je fünfmal dargeboten. Wir erwarteten dabei zumindest beim Eckstimulus mit dem SPÄTEN Gipfel von 11 HALBTÖNEN eine eindeutige und insgesamt beim SPÄTEN Kontinuum eine frühere Exklamativzuweisung.

**ERGEBNIS:** Fig.3 zeigt für die drei Positionen grundsätzlich den gleichen Verlauf, wobei die Eckstimuli diesmal eindeutiger, nämlich mit 1%, 2% und 4% bzw. 94%, 94% und 95 % als EXKLAMATIV kategorisiert werden. Kendalls Konkordanzkoeffizient zeigt für die Positionen einen signifikanten Unterschied bei den Stimuli mit einem Gipfel von 5 und 9 HALBTÖNEN. Ein paarweiser Vergleich mit einem Wilcoxon-Test ergab für das abgesenkte Signifikanzniveau von .014 bei einseitiger Fragestellung und dem 9-HALBTON-Gipfel einen signifikanten Unterschied zwischen dem SPÄTEN und dem MITTLEREN Gipfel ( $p < .003$ ) und zwischen dem SPÄTEN und dem FRÜHEN Gipfel ( $p < .006$ ). Damit ist unsere Hypothese, daß ein SPÄTER Gipfel eher Exklamativität indiziert, für den Bereich von 9 HALBTÖNEN bestätigt. Zwei Unterschiede zu Exp.1 können beobachtet werden:



(1) Die Eckstimuli werden eindeutiger beurteilt - sie entsprechen auch eher den beiden unterschiedlichen Modi: die 3-HALBTON-Stimuli in Exp.2 sind um 30 msec kürzer als ihr Pendant in Exp.1 und die 11-HALBTON-Stimuli sind 90 msec länger und zwei HALBTÖNE höher als der entsprechende Eckstimulus in Exp.1. Bei den Eckstimuli ist kein Unterschied in der Gipfelposition zu beobachten. Offensichtlich kennzeichnet hier die Gipfelhöhe allein schon hinreichend die Modi.

(2) Die 50%-Grenze liegt für alle drei Positionen im Bereich des 7-HALBTON-Stimulus - obwohl die kovariierende Dauermanipulation einen früheren Schnittpunkt erwarten ließe. Diese Frage wird weiter unten in Exp.5 aufgegriffen.

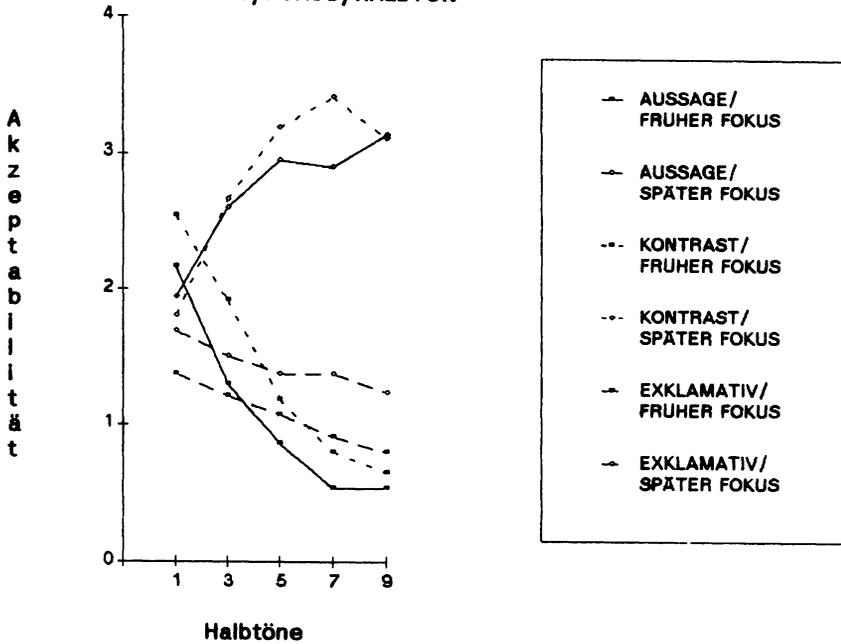
### **EXPERIMENT 3**

**FRAGESTELLUNG:** Es soll untersucht werden, ob sich im KT eine zum IT analoge Kategoriengrenze ergibt und ob sich die Modi bei der Fokusindizierung unterscheiden. Der Aussagemodus wurde zusätzlich unterteilt in normale Aussage (AUSSAGE) und Aussage mit Kontrastakzent (KONTRAST), um herauszufinden, ob der Kontrastakzent anders bewertet wird als der normale Hauptakzent.

**MATERIAL UND DESIGN:** Paare aus allen sechs Kontextsätzen (vgl. Tab.1) und den Stimuli aus Kontinuum I wurden in einem KT je viermal randomisiert dargeboten. Damit ergab sich ein Versuchsplan mit den Faktoren MODUS (AUSSAGE, KONTRAST und EXKLAMATIV), FOKUS (FRÜH und SPÄT) sowie HALBTON (1, 3, 5, 7, 9). Wir erwarteten, ausgehend von der in Batliner (1988) beschriebenen Analyse, bei AUSSAGE und KONTRAST klare, bei EXKLAMATIV nur rudimentäre Fokusedifferenzierung.

**ERGEBNIS:** In Fig.4 sind auf der x-Achse die Stimuli und auf der y-Achse die Mittelwerte der Akzeptabilitätsurteile aufgetragen; '0' steht für 'passen sehr gut zusammen', '1' für 'gut', '2' für 'noch gut', '3' für 'schlecht' und '4' für 'sehr schlecht'. Die Varianzanalyse ergab signifikante Haupteffekte und Interaktionen mit Ausnahme der Interaktion MODUS mit HALBTON (vgl. Tab.3). Für die Interpretation kann also auf die einfachen Haupteffekte zurückgegriffen werden. D.h. daß z.B. der Faktor MODUS nicht gemittelt über die beiden FOKI und alle fünf HALBTON-Höhen betrachtet wird, sondern für jeden FOKUS und jede HALBTON-Höhe getrennt. Das gleiche gilt für die beiden anderen Faktoren. In Tab.3 sind die signifikanten Ergebnisse für die einfachen Haupteffekte bei MODUS und FOKUS mit einem Stern gekennzeichnet. Die HALBTÖNE unterscheiden sich signifikant bei AUSSAGE und KONTRAST, nicht aber beim EXKLAMATIV.

Fig.4: Kontexttest:  
MODUS/FOKUS/HALBTON



Tab.3: Haupteffekte, Interaktionen und einfache Haupteffekte von Exp.3

Faktoren/Interaktionen:	df	F	Sig.			
MODUS	2,18	12.33	.000			
FOKUS	1,19	65.09	.000			
HALBTON	4,16	6.98	.002			
MODUS mit FOKUS	2,18	30.15	.000			
MODUS mit HALBTON	8,12	2.54	.071			
FOKUS mit HALBTON	4,16	18.50	.000			
MODUS mit FOKUS mit HALBTON	8,12	11.07	.000			
einfache Haupteffekte:						
	bei HALBTON	1	3	5	7	9
FOKUS bei AUSSAGE		*	*	*	*	*
FOKUS bei KONTRAST		*	*	*	*	*
FOKUS bei EXKLAMATIV		*	*		*	*
MODUS bei FRÜHEM FOKUS		*	*			
MODUS bei SPÄTEM FOKUS			*	*	*	*

**Fokusedifferenzierung:** Bei AUSSAGE und besonders bei KONTRAST indiziert ein Gipfel von 3 oder mehr HALBTÖNEN einen Fokus auf *Leo*. Offenbar genügt ein Gipfel von 5 HALBTÖNEN für die Fokusindizierung, da sich bei den höheren Gipfeln die Kurvenverläufe nicht mehr wesentlich ändern. Die Verläufe von AUSSAGE und KONTRAST sind sehr ähnlich; das entspricht unseren Erwartungen, da es sich in beiden Fällen um denselben Modus handelt. Für den Kontrastakzent wird üblicherweise eine lokal stärker ausgeprägte Markierung angenommen (vgl. Bannert 1985:300). Anhand unseres Ergebnisses läßt sich der Kontrastakzent vom normalen Hauptakzent nur beim Stimulus mit einem Gipfel von einem Halbton unterscheiden, der bei der AUSSAGE indifferent ist, beim KONTRAST aber signifikant eher einen SPÄTEN FOKUS auf *säuft* indiziert. Beim EXKLAMATIV verlaufen die Kurven für FRÜHEN und SPÄTEN FOKUS ganz anders als bei den Aussagen, nämlich parallel und leicht abfallend. Der SPÄTE FOKUS wird durchgehend schlechter bewertet. Dieses Ergebnis stimmt mit Batliner (1988) überein, wo sich auch nur eine rudimentäre Fokusedifferenzierung beim EXKLAMATIV zeigt. Die Differenzierung sieht so aus, daß grundsätzlich bei einem Kontext, der einen Fokus auf *säuft* erwarten läßt, alle Stimuli mit ihren weniger oder mehr ausgeprägten Akzenten auf *Leo* schlechter bewertet werden. Die Differenzierung führt aber nicht dazu, daß in so einem Fall ein deutlicherer Gipfel auf *Leo* wie bei den Aussagen eine entsprechend schlechtere Bewertung hervorruft.

**Modusedifferenzierung:** Auf die Differenzierung **innerhalb** der Aussagemodi (AUSSAGE vs. KONTRAST) wurde schon im letzten Abschnitt eingegangen; es gibt keinen Unterschied im grundsätzlichen Verlauf der Kurven. Anders bei der Differenzierung **zwischen** den Modi: beim FRÜHEN FOKUS setzt die EXKLAMATIV-Kurve deutlich tiefer an, d.h. diese Kombination wird besser bewertet als die entsprechenden Aussagekombinationen; die Kurve fällt dann aber nicht so steil ab und kreuzt sogar die der Aussagen. Die Verbesserung mit zunehmender Gipfelhöhe ist also beim EXKLAMATIV bei weitem nicht so ausgeprägt wie bei den Aussagen – ein Ergebnis, das auf den ersten Blick kontraintuitiv aussieht, da wir ja davon ausgehen, daß sich der EXKLAMATIV durch ein 'Mehr', also z.B. durch ein 'Mehr' an Gipfelhöhe auszeichnet. Wir nehmen zur Erklärung zwei Faktoren an, deren jeweiligen Stellenwert wir im Augenblick allerdings nicht gegeneinander abwägen können:

(1) Der EXKLAMATIV ist weniger lokal als vielmehr **global** markiert. Die Exklamativität, der Ausdruck des Erstaunens, muß nicht auf dem Testsatz liegen, sondern wird auch schon im Kontext durch das *Mann*- indiziert; der Testsatz kann also auch nur noch einen intonatorischen 'Nachlauf' ohne besondere

Hervorhebung bilden. Dieser Fall kann bei den Stimuli mit einem niedrigen Gipfel vorliegen. (Es ließe sich nun einwenden, daß in einem solchen Fall der Testsatz eben kein Exklamativsatz sei, sondern ein Aussagesatz. Der Testsatz wäre aber in diesem Kontext und ohne exklamativindizierende intonatorische Hervorhebung statt mit Verb-Zweit-Stellung auch mit Verb-Erst-Stellung denkbar: (*Mann - ) Säuft der Leo.* Konsequenterweise wäre dieser Satz dann als Fragesatz zu betrachten, da er aufgrund der Verb-Stellung nur Exklamativsatz oder Fragesatz sein kann - eine etwas seltsam anmutende Lösung. Möglicherweise ergibt sich für dieses Phänomen eine allseits zufriedenstellende Erklärung erst dann, wenn man als Beschreibungseinheit nicht den Satz, sondern die Äußerung ansetzt.)

(2) Wir befinden uns nicht in einem für den EXKLAMATIV typischen Parameterbereich, da die Dauer ja gleich blieb.

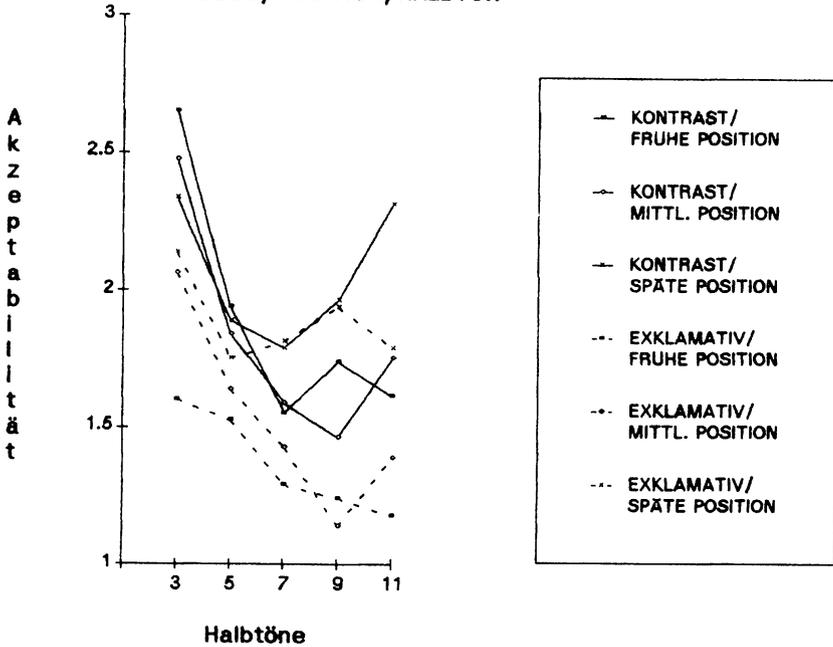
Die Globalität des EXKLAMATIVS ist auch dafür verantwortlich, daß man keinen direkten Vergleich zwischen dem IT und dem KT ziehen kann - man sucht ja vergeblich nach einem Pendant zum Schnittstimulus von ca. 5 HALBTÖNEN im Exp.1. Im IT gab es nur ein Indiz, nur einen Parameter, der verschiedene Werte annimmt: die Gipfelhöhe. In einem solchen Fall wird natürlich dieses 'Mehr' an Ausprägung in einem 'forced choice test' auch mit einem 'Mehr' an Exklamativität gleichgesetzt. Beim KT kann aber auch der Kontext schon eine exklamativindizierende Funktion übernehmen.

#### **EXPERIMENT 4:**

**FRAGESTELLUNG:** Es soll untersucht werden, ob sich die ausgeprägtere Kategorialität der Differenzierung zwischen AUSSAGE und EXKLAMATIV sowie die bessere Indizierung des EXKLAMATIVS durch einen SPÄTEN Gipfel nicht nur im IT, sondern auch im KT nachweisen läßt.

**MATERIAL UND DESIGN:** Die beiden Kontexte KONTRAST/FRÜHER FOKUS und EXKLAMATIV/FRÜHER FOKUS wurden mit den Kontinua III, IV und V gepaart und in einem KT je viermal randomisiert dargeboten. Daraus ergab sich ein Versuchsplan mit den Faktoren MODUS (KONTRAST, EXKLAMATIV), POSITION (FRÜH, MITTE und SPÄT) sowie HALBTON (3,5,7,9,11). Die Kontinua decken einen anderen Bereich als das Kontinuum I ab, damit auf alle Fälle auch exklamativtypische Konturen enthalten sind; dem diente auch die kovariierende Dauermanipulation (Dehnung in 30 msec-Schritten von 270 bis 390 msec auf [eo]).

**Fig.5: Kontexttest:  
MODUS/POSITION/HALBTON**



**ERGEBNIS:** Der erste Eindruck von Fig.5 mag täuschen, da hier, um die Unterschiede deutlicher zu machen, nur eine Teilskala von 1-3 für die y-Achse gewählt wurde. Die Urteile sind also indifferenter, es fehlen 'sehr gute' und 'sehr schlechte' Bewertungen. Auf der einen Seite hat es sich zwar schon in Exp.3 angedeutet, daß beim MODUS keine so ausgeprägten Differenzierungen zu erwarten sind wie beim FOKUS. Auf der anderen Seite ist es aber doch unerwartet, daß die Stimuli mit einem Gipfel von 9 und 11 HALBTÖNEN, die durch die kovariierende Dauermanipulation noch exklamativtypischer werden sollten, mit dem Exklamativkontext relativ schlecht beurteilt wurden, schlechter als der Eckstimulus in Exp.3. U.E. handelt es sich hierbei um einen kombinierten Effekt der spezifischen Art des Tests und der Art der Materialmanipulation und -zusammenstellung: Es wurde ein zwar resynthetisierter, aber sehr gut klingender Kontextsatz gekoppelt mit einem schematisch manipulierten Testsatz; damit ergab sich vor allen Dingen bei den hohen Gipfeln mit starker Dehnung eine gewisse Diskrepanz im globalen Zeitbereich. Da den Versuchspersonen bei dieser Art des Tests natürlich freisteht, bestimmte Kombinationen grundsätzlich weniger akzeptabel zu finden - aber nicht aus den vom Versuchsleiter intendierten

Gründen -, so können, wie in diesem Fall, weniger trennscharfe Kurven entstehen. Entscheidend ist aber, ob sich trotzdem signifikante und sinnvoll interpretierbare Unterschiede ergeben. Tab.4 zeigt das Ergebnis der Varianzanalyse: wieder sind alle Haupteffekte und zwei der Interaktionen signifikant; wir greifen also auf die einfachen Haupteffekte zurück. Wie bei Exp.3 sind die signifikanten Ergebnisse für die einfachen Haupteffekte bei MODUS und POSITION mit einem Stern gekennzeichnet. Die HALBTÖNE unterscheiden sich beim KONTRAST signifikant in allen drei POSITIONEN, beim EXKLAMATIV nur in der MITTLEREN. Wir beschränken uns im folgenden auf eine Interpretation der Fig.6-8, anhand derer die wichtigsten Punkte gut herausgearbeitet werden können.

**Tab.4: Haupteffekte, Interaktionen und einfache Haupteffekte von Exp.4**

Faktoren/Interaktionen:	df	F	Sig.			
MODUS	1,19	10.79	.004			
POSITION	2,18	7.78	.004			
HALBTON	4,16	6.60	.002			
MODUS mit POSITION	2,18	6.67	.007			
MODUS mit HALBTON	4,16	2.68	.069			
POSITION mit HALBTON	8,12	3.85	.018			
MODUS mit POSITION mit HALBTON	8,15	1.95	.142			
einfache Haupteffekte:						
	bei HALBTON	3	5	7	9	11
MODUS bei FRÜHEM GIPFEL		*	*		*	
MODUS bei MITTLEREM GIPFEL					*	*
MODUS bei SPÄTEM GIPFEL						*
POSITION bei KONTRAST						*
POSITION bei EXKLAMATIV		*		*	*	*

Fig.6 zeigt die Mittelwerte für die Interaktion POSITION mit MODUS bei kombiniertem HALBTON. Die FRÜHE POSITION ist beim EXKLAMATIV deutlich besser als die anderen; wie Fig.5 zeigt, sind dafür besonders die niedrigeren Gipfelhöhen verantwortlich. Die SPÄTE POSITION wird am schlechtesten bewertet. Das gilt auch für den KONTRAST, wobei sich hier FRÜHER und MITTLERER Gipfel nicht so stark unterscheiden.

Fig.7 zeigt die Interaktion HALBTON mit MODUS bei kombinierter POSITION. Beim EXKLAMATIV ist der Verlauf ähnlich wie in Exp.3 für beide Fokuspositionen, nämlich leicht fallend. Beim KONTRAST ist er, wieder analog zu Exp.3, anfangs stärker fallend; dann folgt aber ein Wendepunkt beim Stimulus mit dem Gipfel

von 7 HALBTÖNEN. Bei sieben HALBTÖNEN liegt aber auch der Schnittstimulus von Exp.2, d.h. hier zeigt sich auch im KT die Grenze zwischen KONTRAST und EXKLAMATIV, wenn Gipfelerhöhung und -dehnung zusammen eingesetzt werden. Bei solchen Stimuli werden die höheren Gipfel als für den KONTRAST nicht mehr adäquat bewertet.

Fig.8 zeigt die Interaktion HALBTON mit POSITION bei kombiniertem MODUS. Ab 7 HALBTÖNEN ist der Verlauf für den SPÄTEN Gipfel schlechter. Die Verschlechterung ist am deutlichsten beim KONTRAST, vgl. Fig.5. Das wurde erwartet, da der SPÄTE Gipfel Exklamativität indizieren soll. Aber auch beim EXKLAMATIV wird der SPÄTE Gipfel deutlich schlechter bewertet als die beiden anderen Positionen. Dieses Ergebnis steht in Widerspruch zu Exp.2, bei dem der SPÄTE Gipfel signifikant früher Exklamativität indiziert als die beiden anderen. U.E. liegt die Erklärung dafür in einer Diskrepanz im Zeitbereich, wie sie oben für die weniger ausgeprägte Trennschärfe in diesem Experiment überhaupt gegeben wurde, allerdings in einer spezielleren Diskrepanz, nämlich in einem schlechten 'timing' bei der Manipulation des Gipfels im Verhältnis zu den beiden vokalischen Elementen [e] und [o]. Damit wird bewirkt, daß der SPÄTE Gipfel per se als unnatürlich eingestuft wird und das natürlich besonders dann, wenn dieser Gipfel prominent ist, wie bei den höheren Gipfeln im Kontinuum. (Wir können noch nicht angeben, welches 'timing' nötig wäre, um einen besser klingenden Stimulus zu erzeugen, wollen dieser Frage aber in einer eigenen Teiluntersuchung nachgehen.) Der vermeintliche Widerspruch zu Exp.2 erklärt sich also wieder aus der unterschiedlichen Aufgabenstellung von IT und KT: bei einem 'forced choice' muß auch ein schlechter Stimulus akzeptiert, d.h. einer der beiden Kategorien zugeordnet werden. Entsprechend unserer Hypothese wird dann der SPÄTE Gipfel eher dem EXKLAMATIV zugeordnet. Bei einem KT steht es den Versuchspersonen frei, die Kombination zu bewerten, und wenn eine Stimulusserie weniger gut generiert ist, so wird sie einfach durchgehend schlechter bewertet, egal, ob sie mit einem KONTRAST-Kontext oder einem EXKLAMATIV-Kontext gepaart ist.

Fig.6: Haupteffekte bei  
komb. HALBTON

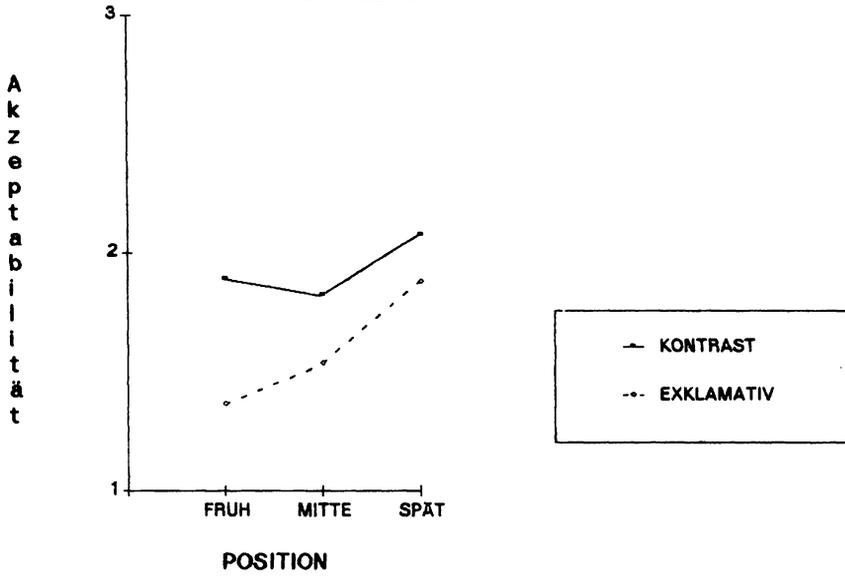


Fig.7: Haupteffekte bei  
komb. POSITION

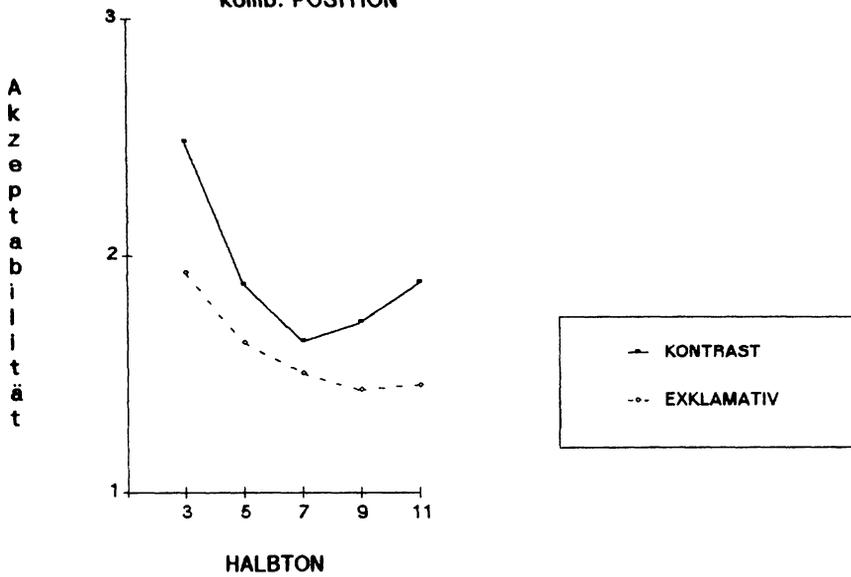
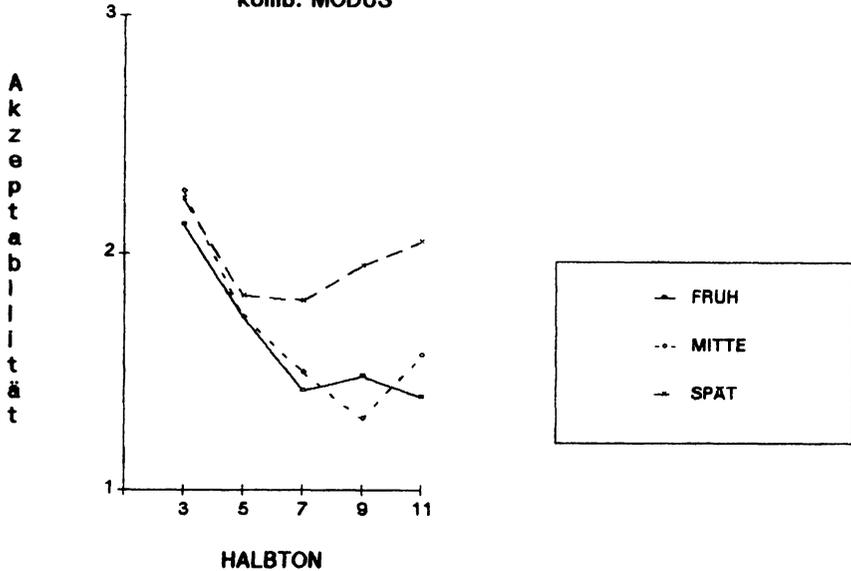


Fig.8: Haupteffekte bei komb. MODUS



#### 4. FOLGEEXPIMENTE

Mit anderen Gruppen von Versuchspersonen wurden weitere Experimente durchgeführt, um bestimmte Fragen, die sich aus Exp.1-4 ergaben, klären zu können.

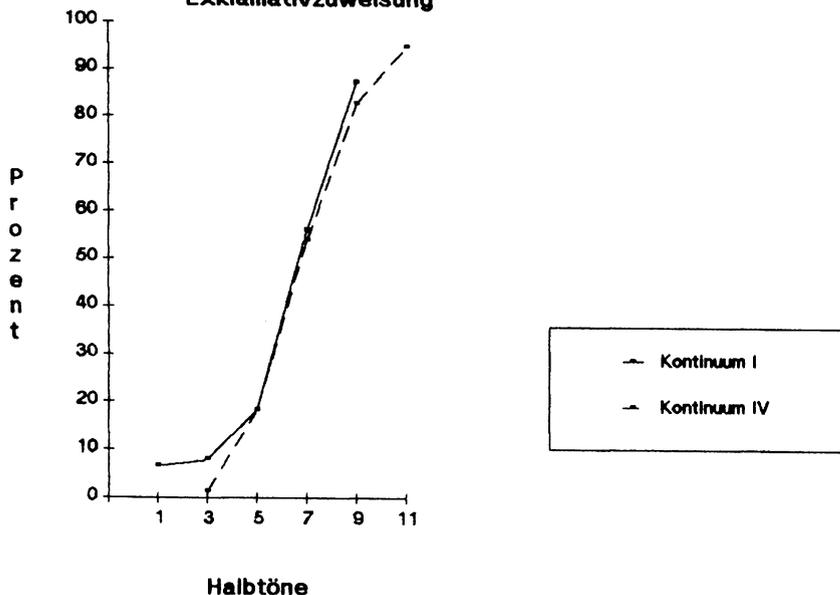
##### EXPERIMENT 5

**FRAGESTELLUNG:** Dieses Experiment sollte überprüfen, ob sich unterschiedliche Schnittstimuli für die Kontinua I und IV im IT ergeben, wenn beide Kontinua zusammen getestet werden.

**MATERIAL UND DESIGN:** 15 Versuchspersonen nahmen an einem IT teil, bei dem das Kontinuum I und das Kontinuum IV mit 10 randomisierten Wiederholungen pro Stimulus dargeboten wurden. (Da der Bezugsstimulus beiden Kontinua angehört, kam er zwanzigmal vor.) Beim Kontinuum I wird nur der Gipfel angehoben, die Dauer bleibt unverändert. Beim Kontinuum IV werden Gipfelhöhe und Dauer variiert. Die 3-, 7- und 9-HALBTON-Stimuli unterscheiden sich also nur in der Dauer, der 5-HALBTON-Stimulus gehört zu beiden Kontinua.

**ERGEBNIS:** In Fig.9 bestätigt der geringe Unterschied zwischen den Kontinua bei einer Gipfelhöhe von 5, 7 und 9 HALBTÖNEN, daß zumindest für diese Schrittgrößen die Gipfelerhöhung relevanter ist als die Dehnung. (Die Dauerdifferenz beträgt zwischen 0 und 60 msec, vgl. Tab.2). Bei beiden Kontinua liegt der Schnittstimulus kurz unterhalb des 7-HALBTON-Stimulus. Ein Designeffekt der Art, daß die Versuchspersonen die Kategoriengrenze in die Mitte zwischen den Eckstimuli, also zwischen die 5- und 7-HALBTON-Stimuli legen, ist nicht auszuschließen. Wenn wir weiter annehmen, daß der Fo-Range und damit auch in gewissem Rahmen die Grenze zwischen AUSSAGE und EXKLAMATIV bei der Produktion sprecherspezifisch variieren können und daß der Hörer - wieder in einem gewissen Umfang - sich bei der Kalibrierung der intonatorischen Parameter darauf einstellt, so läßt sich zwischen dem Designeffekt und den Bedingungen der natürlichen Kommunikation eine Verbindung herstellen: in beiden Fällen gibt es keine exakte Grenze in absoluten Werten, sondern nur eine relativ zum dargebotenen Fo-Range.

**Fig.9: Identifikationstest:  
Exklamativzuweisung**



**EXPERIMENT 6 UND 7**

**FRAGESTELLUNG:** Der wesentliche Punkt bei Überlegungen zur Kategorialen Wahrnehmung von Intonationsphänomenen läßt sich wie folgt skizzieren: Im festen Paradigma der Kategorialen Wahrnehmung wird davon ausgegangen, daß der Mensch nur dann einen Unterschied zwischen zwei Stimuli wahrnimmt, wenn sie verschiedenen Kategorien angehören. Damit wird implizit angenommen, daß die Antworten der Versuchspersonen im Experiment ihre **Wahrnehmung** wiedergeben und nicht ihre **Klassifizierung** der Stimuli. Je mehr Zeit zwischen Präsentation der Stimuli und Reaktion der Versuchspersonen verstreicht, desto unwahrscheinlicher ist diese Annahme - bewiesen werden im eigentlichen Sinne konnte sie auch für ganz kurze Stimuli und Reaktionszeiten nicht. Sie ist im Gegenteil auch bei diesen Fällen fragwürdig (vgl. im einzelnen Massaro/Cohen 1983). Bei unserem Material vergehen zwischen Präsentation der Stimuli und Antwort der Versuchsperson mindestens zwei Sekunden - ergo wird die Annahme für solche Intonationsphänomene noch unwahrscheinlicher. Eine eingehende Auseinandersetzung mit dem Paradigma der Kategorialen Wahrnehmung findet sich in Schiefer/Batlner (1988).

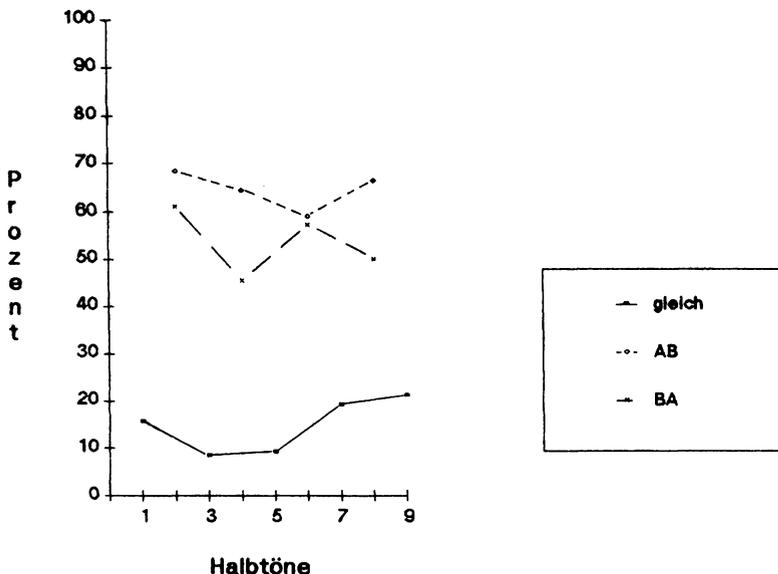
Auf der anderen Seite bedeuten diese Argumente gegen das strikte Paradigma natürlich nicht, daß man den Begriff 'Kategorialität' abschafft. Auch wenn man annimmt, daß in solchen Tests immer klassifiziert und nie nur wahrgenommen wird, daß aber dann, wenn die beiden Stimuli unterschiedlichen Kategorien angehören, die Diskrimination erleichtert wird, kann es sinnvoll sein, einen DT zur Überprüfung dieser Annahme durchzuführen.

**MATERIAL UND DESIGN:** Um nun zu überprüfen, ob sich eine Kategoriengrenze zwischen AUSSAGE und EXKLAMATIV auch bei der Diskrimination zeigt, wurde getrennt ein DT mit den Kontinua I und IV und jeweils fünfmaliger randomisierter Wiederholung pro Anordnung (AB, BA, AA, BB) durchgeführt, an dem 11 (Kontinuum I) bzw. 13 (Kontinuum IV) Versuchspersonen teilnahmen. Da die Paare im 'Einerschritt' zusammengestellt waren, gab es innerhalb der Anordnung AB bzw. BA immer eine Differenz von 2 HALBTÖNEN; hinzu kam beim Kontinuum IV die Dauerdifferenz.

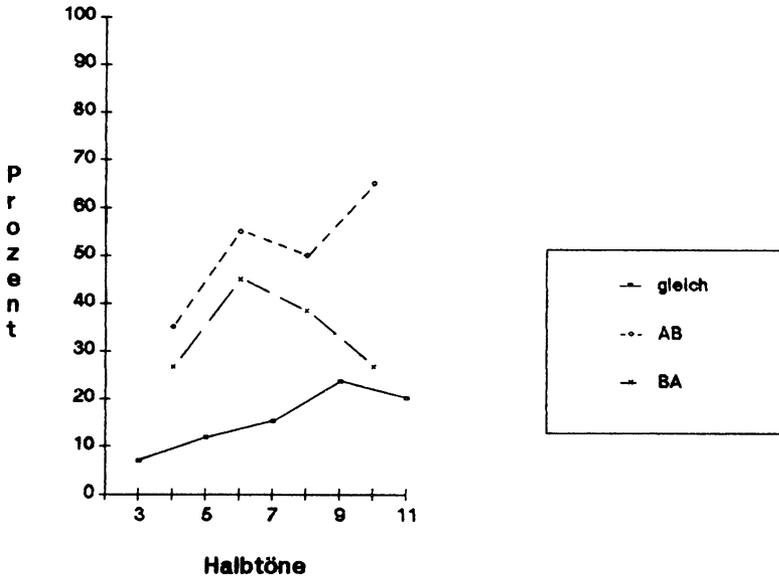
**ERGEBNIS:** Während sich in Fig.10 beim Kontinuum I für die Stimuli ohne kovariierende Dauervariation kein Diskriminationsgipfel zeigt, liegt in Fig.11 für das Kontinuum IV mit Dauervariation bei der Anordnung BA ein Diskriminationsgipfel zwischen dem 5- und dem 7-HALBTON-Stimulus vor. Bei dieser Anordnung folgt der Stimulus mit dem niedrigeren  $F_0$ -Gipfel im Paar dem Stimulus

mit dem höheren Fo-Gipfel. Dieser Gipfel liegt beim Schnittstimulus aus den Exp.1, 2 und 5. Bei der anderen Anordnung AB zeigt sich zwar auch ein solcher Gipfel, aber zugleich ein weiterer, noch nicht sinnvoll erklärbarer Gipfel am rechten Ende des Kontinuums, der sich bei der Anordnung BA nicht wiederfindet. Eine aus den beiden Anordnungen AB und BA gemittelte Kurve würde denn auch keinen ausgeprägten Gipfel zeigen. Wir können dieses Ergebnis hier nicht eingehender diskutieren; in Schiefer/Batliner (1988) vermuten wir, daß die Anordnungen AB und BA den gleichen Status wie die Schrittgröße zwischen den Stimuli haben. D.h. wir können in diesem Fall nur die weniger prominente Anordnung BA diskutieren, die einen deutlicheren Gipfel aufweist. Man kann also sagen, daß wir im DT beim Kontinuum IV eine, wenn auch ebenso wie im IT schwach ausgeprägte, Kategorialität von AUSSAGE und EXKLAMATIV wiederfinden - vorausgesetzt, wir begnügen uns introspektiv mit der Übereinstimmung des Diskriminationsgipfels mit dem Schnittstimulus der Identifikation und versuchen nicht, die Kategorialität der Wahrnehmung im Sinne des Paradigmas nachzuweisen. Der fehlende Gipfel bei den Diskriminationskurven in Fig.10 ist ein weiterer Hinweis darauf, daß für einen kategorialen Übergang von AUSSAGE zu EXKLAMATIV eine Dehnung der Hauptakzentsilbe unerlässlich ist.

**Fig.10: Diskriminationstest:  
Kontinuum I**



**Fig.11: Diskriminationstest:  
Kontinuum IV**



##### 5. EIN WEITERES KONTINUUM: DER 'SEHR FRÜHE' GIPFEL

Kohler (1987) hat auf der Grundlage des natürlich produzierten Satzes *Sie hat ja gelogen* ein Kontinuum mit 11 Stimuli erzeugt, bei dem der Gipfel von einer prä vokalischen Position im [l] zu einer späten Position im [o:] von *gelogen* wandert. Er schließt aus den mit diesem Kontinuum durchgeführten Experimenten Folgendes (Kohler 1987:136):

Das Gehör spricht ganz offensichtlich kategorial auf die Positionierung des F<sub>0</sub>-Gipfels relativ zur Grenze zwischen Silbenperipherie und Silbenkern [...] an. Und diese kategoriale Differenzierung wird im Deutschen semantisch genutzt für eine "begründende Äußerung am Ende einer Argumentationskette" gegenüber dem "Beginn einer Argumentation" diesseits und jenseits der Grenze in der 11-er Serie von links nach rechts. Eine weitere Verschiebung nach rechts erhöht den Grad der Emphase/Empörung.

Auch wenn bei Kohler die Gipfelhöhe immer konstant blieb und sowohl seine linguistischen Vorgaben als auch seine Art der Manipulation nur bedingt mit unserem Vorgehen vergleichbar sind, so bildet doch die Verschiebung des Gipfels aus einer frühen in eine späte Position ein *tertium comparationis* - allerdings

ein unvollständiges, da Kohlers früher Gipfel zwischen Konsonant und Vokal in unseren bisherigen Kontinua nicht auftritt: unser FRÜHER Gipfel liegt schon 100 msec im Vokal. Damit wäre auch unser weniger eindeutiges Ergebnis bei Exp.4 zu erklären, wenn man annimmt, daß im Material keine Kategoriengrenze enthalten ist. Um diese Frage zu überprüfen, wurde ein zusätzliches Kontinuum (II in Tab.2) generiert, bei dem der Gipfel der Fo-Kontur noch im [l] liegt und [eo] einen Fall aufweist.

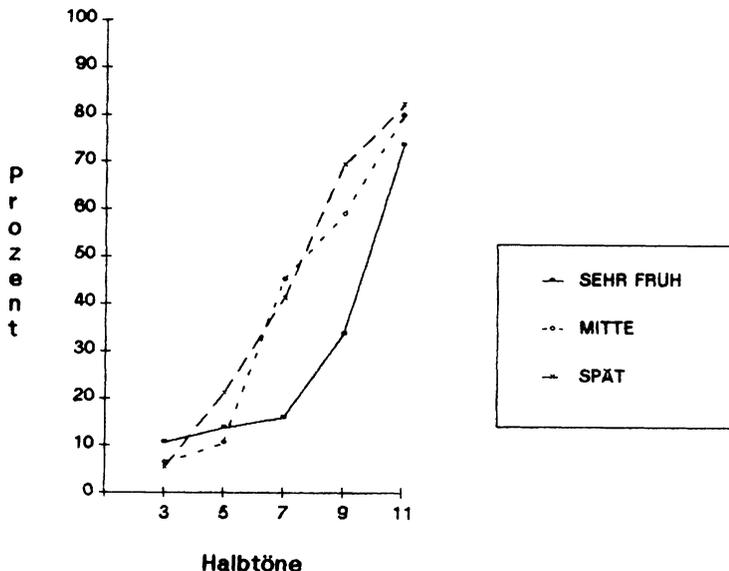
### **EXPERIMENT 8**

**FRAGESTELLUNG:** Es sollte untersucht werden, ob sich die Annahme, daß ein SEHR FRÜHER Gipfel eher eine AUSSAGE als einen EXKLAMATIV indiziert, im IT bestätigt.

**MATERIAL UND DESIGN:** Die Stimuli mit einem SEHR FRÜHEN, MITTLEREN und SPÄTEN Gipfel (Kontinua II, IV und V in Tab.1) wurden in einem IT je fünfmal randomisiert 19 Versuchspersonen dargeboten. Sie sollten entscheiden, ob es sich bei den Stimuli um eine Aussage oder einen Ausruf handelt.

**ERGEBNIS:** Kendalls Konkordanzkoeffizient zeigt für das in Fig.12 dargestellte Ergebnis signifikante Unterschiede bei den Stimuli mit einem Gipfel von 7 und 9 HALBTÖNEN. Für den paarweisen Vergleich nach Wilcoxon bei einer einseitigen Fragestellung wurde das Signifikanzniveau auf .017 abgesenkt. Der SEHR FRÜHE Gipfel unterschied sich bei beiden Gipfelhöhen signifikant vom MITTLEREN und SPÄTEN Gipfel. Für 7 Halbtöne galt: SEHR FRÜH vs. MITTE:  $p < .010$ , SEHR FRÜH vs. SPÄT:  $p < .017$ . Für 9 Halbtöne galt: SEHR FRÜH vs. MITTE:  $p < .016$ , SEHR FRÜH vs. SPÄT:  $p < .003$ . Wie bei Exp.2 zeigte sich bei dem höchsten und den niedrigen Gipfeln kein signifikanter Unterschied. Wir vermuten, daß bei den niedrigen Gipfeln der Unterschied in der POSITION perzeptiv weniger auffällig ist und daß beim höchsten Gipfel mit 11 HALBTÖNEN die Gipfelhöhe gegenüber der POSITION dominiert. Der MITTLERE und der SPÄTE Gipfel liegen dabei mit 80% bzw. 82.2% Exklamativzuweisung um gut 10% unter den Werten aus Exp.2. 'Störfaktor' sind hier einzelne Versuchspersonen, die - im Sinne unserer Annahmen - 'irregulär' urteilten. Wir haben aber grundsätzlich davon abgesehen, solche Versuchspersonen unberücksichtigt zu lassen.

**Fig.12: Identifikationstest:  
Exklamativzuweisung**



#### **EXPERIMENT 9**

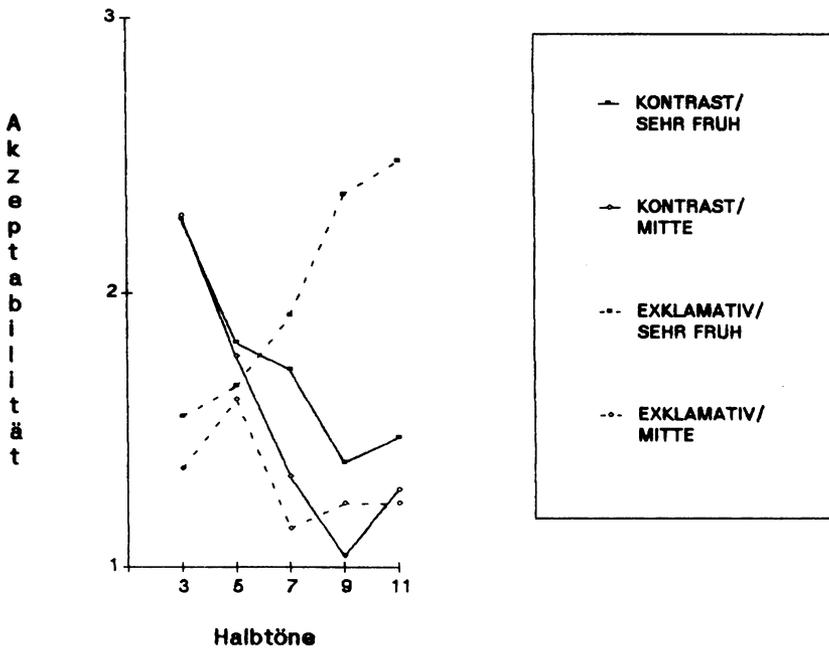
**FRAGESTELLUNG:** Es sollte untersucht werden, ob sich die schlechtere Exklamativzuweisung an die Stimuli mit einem SEHR FRÜHEN Gipfel auch im KT replizieren läßt.

**MATERIAL UND DESIGN:** 13 Versuchspersonen wurden in einem KT die Stimuli aus den Kontinua II und IV, also die SEHR FRÜHEN und MITTLEREN Gipfel, gepaart mit den Kontextsätzen KONTRAST, FRÜHER FOKUS und EXKLAMATIV, FRÜHER FOKUS (vgl. Tab.1) je dreimal randomisiert dargeboten.

**ERGEBNIS:** Kendalls Konkordanzkoeffizient ergibt für die Verläufe in Fig.13 einen signifikanten Unterschied bei den Stimuli mit einem Gipfel von 3 und 9 HALBTÖNEN. Da im Anschluß daran 12 paarweise Vergleiche nach Wilcoxon durchgeführt wurden, mußte das Signifikanzniveau für die einseitige Fragestellung auf .008 abgesenkt werden. Dabei unterschied sich der SEHR FRÜHE Gipfel mit 9 Halbtönen in der Kombination mit dem Exklamativkontext signifikant vom MITTLEREN Gipfel mit Kontrastkontext ( $p < .008$ ) und Exklamativkontext ( $p < .004$ ). Dieses

Ergebnis entspricht der Annahme, daß der SEHR FRÜHE Gipfel exklamativuntypisch ist. Die geringe Anzahl der Versuchspersonen, die Streuung der Antworten und die notwendige Absenkung des Signifikanzniveaus dürften weitere signifikante Unterschiede verhindert haben. Es mag aber doch erlaubt sein, die Kurvenverläufe insgesamt zu interpretieren: mit Ausnahme des SEHR FRÜHEN Gipfels im KONTRAST-Kontext gleichen sie den Verläufen der SPÄTEN und MITTLEREN Gipfel von Exp.4 in Fig.5. Im Gegensatz zum SPÄTEN Gipfel haben wir es aber hier nicht mit einer Verschlechterung in beiden Kontextkonstellationen zu tun. D.h. es liegt nicht an einer mangelhaften Qualität der Stimuli, sondern die Verschlechterung beim Exklamativkontext ist wirklich durch die Unverträglichkeit des SEHR FRÜHEN Gipfels mit dem EXKLAMATIV bedingt.

**Fig.13: Kontexttest:  
MODUS/POSITION/HALBTÖN**



Die Ergebnisse von Exp.8 und 9 stützen die Annahme von Kohler, daß ein prävokalischer Gipfel eher eine abgeschlossene Aussage und ein Gipfel im Vokal eher Emphase (und damit cum grano salis Exklamativität in unserem Sinne) indiziert. Allerdings haben wir kein Pendant zu seiner Unterscheidung zwischen einem kategorialen und einem allmählichen Übergang beim Wechsel von der prävokalischen Position in eine Position im Vokal einerseits und bei einer weite-

ren Verschiebung nach rechts andererseits gefunden. Möglicherweise ist das durch die unterschiedliche Materialauswahl und Fragestellung bedingt. Unabhängig davon zeigen aber unsere Ergebnisse, daß die Gipfelhöhe ein Hauptmerkmal, die Gipfelposition dagegen ein Nebenmerkmal sein dürfte, da die Position nur in einem begrenzten Fo-Range-Bereich wirksam wird.

## 6. DISKUSSION

**CAVEAT:** Die üblichen Schlüsse aus solchen Experimenten mit einem beschränkten Material auf 'die deutsche Sprache' können natürlich nur vorsichtig gezogen werden und sollten gestützt werden durch konvergierende Ergebnisse mit anderem Material und anderen Methoden. Ebenso nötig ist es, die 'Ontologie' der Begriffe zu überprüfen: Ist z.B. die Gipfelposition auf der Hauptakzentsilbe das entscheidende Merkmal, oder ist es nicht doch der Position dieser Hauptakzentsilbe im Satz untergeordnet? Dieses zweite Merkmal wurde in unseren Tests nicht berücksichtigt. Ein Modell, das lokale Akzentveränderungen über einen globalen Intonationsverlauf der Äußerung legt, wird eher die Relevanz dieses Merkmals annehmen; vgl. auch die Überlegungen weiter unten.

**METHODE:** Es hat sich als sinnvoll erwiesen, alle drei verschiedenen Arten von Tests einzusetzen, wobei wir auf alle Fälle für die kombinierte Verwendung von IT und KT plädieren. Der IT ist trennschärfer bei der Entdeckung von relevanten Merkmalen, der KT setzt einen natürlicheren Rahmen und kann dabei helfen, unnatürliche Kombinationen bzw. Stimuli zu entdecken.

**DIE INTONATORISCHE KENNZEICHNUNG VON AUSSAGE UND EXKLAMATIV:** Ein nach rechts verschobener Gipfel begünstigt eine Interpretation als EXKLAMATIV, ein nach links (aus dem vokalischen Bereich heraus) verschobener Gipfel spricht eher für einen Nicht-Exklamativ. Bei der Gipfelhöhe deutet alles darauf hin, daß wir es mit einem breiten Übergangsbereich zu tun haben, wobei die Kategoriegrenze bei einem Gipfel von etwa 7 HALBTÖNEN liegt. Der Nicht-Exklamativ, insbesondere der KONTRAST, kann natürlich auch einen höheren Gipfel als 7 HALBTÖNE aufweisen, er überschreitet aber nicht einen gewissen Dauerbereich. Exklamativtypisch ist also auch - und damit bestätigen wir Bekanntes - die Kombination von hohem Gipfel und Dehnung. Da wir die Dehnung nur kovariierend zum Gipfel untersucht haben, können wir für sie keine Grenzzone - etwa einen Verhältniswert von Dauer der Hauptakzentsilbe zur durchschnittlichen Dauer der übrigen Silben - angeben. Die stabilste Unterscheidung zwischen

AUSSAGE und EXKLAMATIV findet sich, um unseren Begriff aus der Eingangsdiskussion wieder aufzunehmen, wohl nicht bei einem 'Mehr' oder 'Weniger' an intonatorischer Ausprägung, sondern dort, wo gar nichts ist, nämlich bei der rudimentären oder fehlenden Fokusedifferenzierung des EXKLAMATIVS. Beim EXKLAMATIV ist also die Akzentuierung freier als bei der AUSSAGE; er besitzt deshalb keinen Fokusakzent, sondern nur einen Hauptakzent. Damit erhält man eine 'negative' Definition des EXKLAMATIVS, die auch operationalisiert werden kann: Wenn die Akzentuierung einer Äußerung in scheinbarem Widerspruch zur Fokusindizierung durch den Kontext steht, so muß es sich um einen EXKLAMATIV handeln - sieht man von schlichten Fehlakzentuierungen ab. Kennzeichnend für den EXKLAMATIV ist auch, daß andere sprachliche Mittel als die Intonation der betreffenden Teiläußerung den Ausdruck der Exklamativität übernehmen können - z.B. ein vorhergehender Kontext. Diese Ansicht setzt natürlich voraus, daß man ein intonatorisches 'Mehr' zwar als prototypisch, aber nicht als konstituierend für den EXKLAMATIV ansieht.

**DIE FRAGE NACH DEM TIEFEREN GRUND:** Es mag am Ende erlaubt sein, etwas zu spekulieren. Wir haben ganz zu Anfang die These aufgestellt, daß sich der Exklamativ gegenüber der Aussage durch ein 'Mehr' an intonatorischer Ausprägung auszeichnet. (Die Modi werden im folgenden wieder klein geschrieben, da diese Überlegungen über die speziellen Konstellationen unserer Experimente hinausreichen.) Die übliche Annahme, daß ein höherer Gipfel und eine größere Dauer exklamativtypisch sind, hat sich denn auch in unseren Experimenten bestätigt. Was haben aber ein fehlender Fokusakzent, das gänzliche Fehlen von intonatorischen Ausprägungen und der SPÄTE Gipfel (bzw. ein Gipfel, der nicht SEHR FRÜH in unserem Sinn positioniert ist) mit einem 'Mehr' zu tun? Ohne eine Erklärung dieser Ergebnisse müßte die Antwort auf die im Titel dieses Beitrags formulierte Frage ein entschiedenes, aber unbefriedigendes 'sowohl - als auch' enthalten: Der Exklamativ zeichnete sich dann sowohl durch ein 'Mehr' aus als auch gelegentlich durch etwas anderes oder durch gar nichts.

Eine Lösung dieses Dilemmas liegt u.E. darin, daß man nicht automatisch ein 'Mehr' an intonatorischer Ausprägung mit einem 'Mehr' an Information gleichsetzt. Nehmen wir an, daß der markierte Modus Exklamativ sich durch ein solches 'Mehr' auf der funktionalen Seite (Ausdruck des Erstaunens, größeres Engagement, usw.) auszeichnet. Wir können nun diesem Exklamativ den unmarkierten Normalfall einer 'unauffälligen' Aussage gegenüberstellen, die einen leicht fallenden Deklinationsverlauf und einen normalen, mit dem Kontext verträglichen Fokusakzent ohne allzu große Hervorhebungen aufweist. Wenn wir annehmen, daß

ein Abweichen von dieser Normalstruktur markiert und deshalb per se geeignet zur Kennzeichnung des Exklamativs ist, so kann auch ein 'Weniger' an intonatorischer Ausprägung ein 'Mehr' an Information bedeuten. Das Fehlen des normalen Fokusakzents bzw. eine auf den ersten Blick nicht mit der durch den Kontext indizierten verträgliche Akzentuierung bedeutet ein 'Mehr' an Information, ebenso eine Abweichung von der fallenden Deklination. Ein SPÄTER Gipfel, der absolut gesehen die gleiche Höhe wie ein früherer Gipfel aufweist, ist natürlich relativ zur Deklinationslinie prominenter.

Diese letzte Feststellung gilt dann, wenn man den Gipfel in bezug setzt zur **globalen Fo-Kontur**. Betrachtet man die Verhältnisse **innerhalb** der Silbe, so kann ein weiterer Effekt vorliegen. Der FALL ist beim SEHR FRÜHEN und beim SPÄTEN Gipfel gleich groß, im signifikanten Bereich von Exp.8 und 9 hat aber der RISE auf dem betonten Element beim SPÄTEN Gipfel einen Umfang von 9 HALBTÖNEN, wogegen der RISE beim SEHR FRÜHEN Gipfel nur einen Umfang von 2 HALBTÖNEN aufweist. (Ein ähnlicher Unterschied ist auch bei den Stimuli von Kohler 1987 zu beobachten.) Bei insgesamt gleicher Dauer der Silbe wird also beim SPÄTEN Gipfel ein bedeutend längerer und stärkerer RISE produziert. Es ist nachgewiesen (vgl. Sundberg 1979), daß die Produktion eines RISE länger dauert als die Produktion eines FALL gleichen Umfangs. Das könnte auch dazu führen, daß in der Perzeption bei physikalisch gleicher Dauer beim SPÄTEN Gipfel der Eindruck einer größeren Dauer entsteht (vgl. Ohala 1978:30) – und die Dehnung ist ja ein stabiles Kennzeichen des Exklamativs. Damit ließe sich der arbiträr aussehende Unterschied der Gipfelpositionen doch einem 'Mehr' oder 'Weniger' an intonatorischer Ausprägung unterordnen.

Es muß dahingestellt bleiben, welche dieser Erklärungsmöglichkeiten zutrifft. Sie schließen sich natürlich nicht gegenseitig aus, und sie sind alle gleichermaßen attraktiv, weil sie einen gemeinsamen Nenner bilden für die Antwort auf die Frage, warum denn gerade so unterschiedlich aussehende Merkmale wie Gipfelposition, fehlende Fokuszuweisung und größere intonatorische Ausprägung den Exklamativ kennzeichnen. In allen diesen Fällen ist mit einem 'Mehr', aber auch mit einem 'Weniger' an intonatorischen Ausprägung immer ein 'Mehr' an Information verbunden.

## LITERATUR

- Altmann, Hans (1987): Zur Problematik der Konstitution von Satzmodi als Formtypen. In: Meibauer, Jörg (Hg.): Satzmodus zwischen Grammatik und Pragmatik. Tübingen, Niemeyer, 22-56. (= Linguistische Arbeiten 180).
- Bannert, Robert (1985): Fokus, Kontrast und Phrasenintonation im Deutschen. Zeitschrift für Dialektologie und Linguistik, 52, 289-305.
- Batliner, Anton (1988): Modus und Fokus als Dimensionen einer Nonmetrischen Multidimensionalen Skalierung. (In diesem Band).
- Kohler, Klaus J. (1987): Funktionen von F<sub>0</sub>-Gipfeln im Deutschen. In: Tillmann, Hans G. / Willée, Gerd (Hgg.): Analyse und Synthese gesprochener Sprache. Hildesheim u.a., Olms, 133-140.
- Massaro, Dominic W. / Cohen, Michael M. (1983): Categorical or Continuous Speech Perception: A New Test. Speech Communication, 2, 15-35.
- Näf, Anton (1987): Gibt es Exklamativsätze? In: Meibauer, Jörg (Hg.): Satzmodus zwischen Grammatik und Pragmatik. Tübingen, Niemeyer, 140-160. (= Linguistische Arbeiten 180).
- Oden, Gregg C. (1979): Fuzzy Propositional Approach to Psycholinguistic Problems: An Application of Fuzzy Set Theory in Cognitive Science. In: Gupta, Madan M. / Ragade, Rammohan K. / Yager, Ronald R. (Hgg.): Advances in Fuzzy Set Theory and Applications. Amsterdam u.a., North-Holland Publishing Company, 409-420.
- Ohala, John J. (1978): Production of Tone. In: Fromkin, Victoria A. (Hg.): Tone. A Linguistic Survey. New York u.a., Academic Press, 5-39.
- Repp, Bruno H. (1984): Categorical Perception: Issues, Methods, Findings. In: Lass, Norman J. (Hg.): Speech and Language. Orlando u.a., Academic Press, 243-335.
- Schiefer, Lisa / Batliner, Anton (1988): Intonation, Ordnungseffekt und das Paradigma der Kategorialen Wahrnehmung. (In diesem Band).
- Sundberg, Johan (1979): Maximum Speed of Pitch Changes in Singers and Untrained Subjects. Journal of Phonetics, 7, 71-79.